

# Modalité épistémique et discours scientifique

*Une étude contrastive des modalisateurs épistémiques dans des articles de recherche français, norvégiens et anglais, en linguistique et médecine*

**Eva Thue Vold**



Thèse pour le degré de *philosophiae doctor* (PhD)

Université de Bergen

2008

Bergen, Norvège 2007

# Modalité épistémique et discours scientifique

*Une étude contrastive des modalisateurs épistémiques dans des articles de recherche français, norvégiens et anglais, en linguistique et médecine*

**Eva Thue Vold**



Thèse pour le degré de *philosophiae doctor* (PhD)

Universitetet i Bergen  
Institutt for fremmedspråk  
Det humanistiske fakultet

2008



## ABSTRACT

Epistemic modality markers are linguistic expressions that explicitly qualify the truth value of a proposition, by marking the informational content as either certain or uncertain. The present thesis focuses on epistemic modality markers indicating uncertainty, and explores the use of such markers in academic discourse. The material used is compiled within the larger KIAP-corpus and is a selection of research articles written in three different languages (French, Norwegian and English) and belonging to two different disciplines (linguistics and medicine). Carried out within the framework of the Norwegian KIAP project (Kulturell Identitet i Akademisk Prosa (Cultural Identity in Academic Prose)), the study adopts a doubly contrastive approach, focusing on variation across languages as well as across disciplines. Gender differences are also examined.

The first part of the thesis forms the theoretical basis for the analyses and is to a large extent devoted to a discussion of the concept of modality and to relevant previous research. It is argued that the linguistic category of modality is traditionally quite poorly defined in the literature. Consequently, there is not always coherence between its definition and the linguistic items actually included in it. Considerable emphasis is put on the delimitation of the category of epistemic modality and the ways in which it differs from and overlaps with related categories such as alethic modality, sporadicity and hedging.

The second part of the thesis explores the frequency (ch.6) and pragmatic functions (ch.7) of a selection of epistemic modality markers in the corpus. The study shows that French-speaking authors use significantly fewer such markers than do English-speaking and Norwegian authors, suggesting that language background has a large influence on the authors' use of hedges. As far as frequency is concerned, no major differences were observed between the disciplines, nor between male and female authors. However, the factors of discipline and language seem to be interrelated, in the sense that differences between languages are larger within linguistics than within medicine, thus illustrating the fact that medicine is a more internationalised discipline than linguistics. Moreover, the two disciplines seem to prefer different types of markers, and there are notable differences when it comes to the pragmatic functions the markers tend to have. In the medical texts, the occurrences most often have a content-oriented function, i.e. they are used to indicate hypotheses, to hedge conclusions or to signal methodological limitations. These usages, except for the latter (which seems to be typical of experimental articles), are also frequent in the linguistic texts, but in addition the linguists make use of more interpersonal functions. For example, they use epistemic modality to mitigate criticism put forward against the work of others and to signal precaution while interpreting other researchers' texts or findings. These observations reflect the fact that the norms for argumentation differ between disciplines. Overt and direct argumentation is more recurrent in linguistics than in medicine, and thus the use of epistemic modality or hedging as a politeness strategy is more frequent in linguistic articles than in medical ones.

The findings are relevant for teachers and students of academic writing as well as for anyone involved in cross-cultural communication between researchers.



## REMERCIEMENTS

Je tiens à remercier en tout premier lieu Mme Kjersti Fløttum, qui a dirigé ce travail de thèse avec autant d'enthousiasme que de compétence. C'est elle qui m'a encouragée à poursuivre mes études au niveau du doctorat, et sans elle cette thèse n'aurait jamais vu le jour. Je la remercie pour le soutien infailible qu'elle m'a apporté tout au long du projet, aux meilleurs moments comme aux plus difficiles, ainsi que pour sa disponibilité constante, y compris durant les périodes chargées. Je la remercie également pour les entretiens stimulants que j'ai eus avec elle et qui m'ont beaucoup appris, ainsi que pour ses remarques et conseils toujours judicieux, qui ont été d'une grande utilité pour la rédaction de la présente thèse. Je lui suis gré aussi du temps qu'elle a consacré à lire mes écrits, non seulement pour la thèse, mais aussi pour les communications et articles, et la remercie enfin pour la confiance qu'elle m'a accordée en m'accueillant au sein de son équipe KIAP, à laquelle je tiens aussi à exprimer ma vive gratitude. Mes remerciements vont à tous les membres de l'équipe KIAP pour nos discussions stimulantes, nos réunions agréables et fructueuses et pour avoir lu et commenté mes textes – leurs remarques m'ont été précieuses. J'adresse un remerciement particulier à Torodd Kinn pour son aide dans le choix et la réalisation des tests statistiques.

Ma gratitude va également à l'école doctorale de sciences du langage et de philologie à l'Université de Bergen et à ses membres pour toutes les réunions que nous avons eues et pour l'appui que nous nous apportons mutuellement. Les séances où j'ai pu présenter mes travaux m'ont été particulièrement utiles – et je tiens à ce propos à remercier tous ceux qui m'ont aidée par leurs remarques précieuses.

J'exprime aussi ma reconnaissance à M. Francis Grossman et son équipe au laboratoire LIDILEM à l'université Stendhal Grenoble 3 pour m'avoir si chaleureusement accueillie lors de mon stage de 15 jours. Ce stage bref, mais fort stimulant, m'a permis de mettre à l'épreuve mes idées dans un milieu de linguistes et de didacticiens qui travaillent sur la langue française. J'adresse un remerciement spécial à Fanny Rinck pour avoir fait de mon stage un temps très réussi et pour avoir relu et corrigé mon manuscrit.

Je tiens enfin à remercier Eivind Kolflaath pour avoir toujours pris le temps de répondre à mes questions et pour les discussions fructueuses que nous avons eues sur la modalité.





# SOMMAIRE

<b>1 Introduction</b> .....	9
1.1 Objet et objectifs .....	9
1.2 La modalité épistémique et l'atténuation : une présentation préliminaire .....	13
1.3 Questions de recherche et déroulement de l'étude .....	14
1.4 Plan de la thèse .....	16

## Première partie : théorie

<b>2 Genre étudié et approche adoptée</b> .....	21
2.1 Le genre de l'article de recherche .....	21
2.1.1 <i>L'article de recherche en général</i> .....	21
2.1.2 <i>Les articles de recherche en médecine et en linguistique</i> .....	24
2.1.3 <i>Recherches antérieures sur l'article de recherche</i> .....	26
2.2 Approche adoptée .....	35
2.2.1 <i>Une double perspective comparative</i> .....	35
2.2.2 <i>Quelques remarques sur les difficultés liées à la perspective comparative</i> .....	36
<b>3 La modalité épistémique</b> .....	41
3.1 La modalité .....	42
3.1.1 <i>Circonscrire la catégorie de la modalité</i> .....	42
3.1.1.1 Introduction et définition de départ .....	42
3.1.1.2 Quelques précisions terminologiques .....	44
3.1.1.3 La modalité : une catégorie sémantique .....	47
3.1.1.4 Conceptions divergentes de la modalité .....	51
3.1.2 <i>Types de modalité</i> .....	62
3.1.2.1 La modalité épistémique .....	62
3.1.2.2 La modalité aléthique .....	67
3.1.2.3 La modalité existentielle .....	71
3.1.2.4 La modalité déontique .....	73
3.1.2.5 La modalité dynamique .....	74
3.1.3 <i>Bilan</i> .....	74

3.2 L'épistémicité .....	76
3.2.1 Définir la catégorie de l'épistémicité .....	76
3.2.2 Conceptions différentes .....	78
3.2.3 Evidentialité et marqueurs mixtes .....	80
3.3 La réalisation linguistique de la modalité épistémique .....	85
3.3.1 Moyens grammaticaux et moyens lexicaux .....	85
3.3.2 La réalisation linguistique de la modalité épistémique en français, anglais et norvégien .....	88
3.3.2.1 Moyens grammaticaux .....	88
3.3.2.2 Moyens lexicaux .....	98
3.3.3 La combinaison de moyens grammaticaux et lexicaux .....	100
3.3.4 Bilan comparatif .....	101
<b>4. Recherches antérieures sur le phénomène d'atténuation .....</b>	<b>103</b>
4.1 Le concept d'atténuation ('hedging') .....	103
4.2 Quelques résultats antérieurs .....	110

## Seconde partie : analyses

<b>5. Matériaux et méthode .....</b>	<b>121</b>
5.1 Présentation des matériaux .....	121
5.2 Méthode d'analyse .....	124
<b>6. Analyse quantitative .....</b>	<b>129</b>
6.1 Hypothèses et problématique de la recherche .....	129
6.2 La sélection des marqueurs : qu'est-ce qu'un marqueur épistémico-modal ? .....	131
6.2.1 Modalité épistémique et atténuation .....	131
6.2.1.1 Modalité épistémique et adverbess « modificateurs » .....	132
6.2.1.2 Modalité épistémique et approximateurs .....	133
6.2.1.3 Modalité épistémique et adverbess de fréquence (sporadicité) .....	133
6.2.1.4 Modalité épistémique et compléments adverbiaux d'opinion .....	134
6.2.2 Marqueurs de certitude, marqueurs « neutres » et marqueurs d'incertitude .....	134

6.2.3 Modalisation rapportée ou rapport modalisé : la question de savoir qui modalise ...	137
6.3 Analyse exploratoire et premières observations .....	140
6.4 Analyse quantitative .....	151
6.4.1 Méthode .....	151
6.4.1.1 Critères de sélection .....	151
6.4.1.2 Les recherches sur ordinateur .....	152
6.4.1.3 Le classement des attestations .....	154
6.4.2 Résultats .....	160
6.4.2.1 Fréquence .....	164
6.4.2.2 Types de marqueurs .....	173
6.5 Bilan .....	179
<b>7. Analyses qualitatives .....</b>	<b>183</b>
7.1 Fonctions pragmatiques de modalisateurs épistémiques .....	183
7.2 ‘Sembler’ et ‘paraître’ .....	191
7.2.1 Fonctionnement syntaxique .....	191
7.2.2 Sémantisme .....	194
7.2.3 ‘Sembler’ et ‘paraître’ – marqueurs polyphoniques .....	197
7.2.4 La fréquence de ‘sembler’ et de ‘paraître’ dans le corpus .....	200
7.2.5 Le cotexte immédiat .....	202
7.2.5.1 Répartition des constructions syntaxiques .....	203
7.2.5.2 L’emploi de modalisateurs accumulés .....	208
7.2.5.3 Le rôle du datif à la première personne .....	208
7.2.6 Les fonctions pragmatiques de ‘sembler’ et de ‘paraître’ .....	211
7.2.7 Bilan .....	219
7.3 ‘Synes’ et ‘se ut (til/som)’ .....	221
7.3.1 Fonctionnement syntaxique et sémantisme .....	221
7.3.2 La fréquence de ‘synes’ et de ‘se ut (til/som)’ dans le corpus .....	226
7.3.3 Le cotexte immédiat .....	227
7.3.3.1 Répartition des constructions syntaxiques .....	227
7.3.3.2 L’emploi de modalisateurs accumulés .....	231
7.3.3.3 Le rôle du datif à la première personne .....	232
7.3.4 Fonctions pragmatiques de ‘synes’ et de ‘se ut (til/som)’ .....	233
7.3.5 Bilan .....	236

7.4 'Pouvoir' et 'kunne' .....	238
7.4.1 Fonctionnement syntaxique de 'pouvoir' / 'kunne' .....	239
7.4.2 Contenu sémantique .....	240
7.4.3 'Pouvoir' / 'kunne' épistémiques – critères de classification .....	244
7.4.4 La fréquence de 'pouvoir' / 'kunne' dans le corpus .....	246
7.4.5 Le cotexte immédiat .....	252
7.4.6 Fonctions pragmatiques de 'pouvoir' / 'kunne' épistémiques .....	256
7.4.7 Bilan .....	263
7.5 Bilan .....	265
7.6 Quatre études de cas .....	268
7.6.1 L'article Noling05 .....	268
7.6.2 L'article Nomed08 .....	280
7.6.3 L'article Frling15 .....	285
7.6.4 L'article Frmed11 .....	290
7.6.5 Bilan .....	295
<b>8. Conclusion</b> .....	297
8.1 Synthèse .....	297
8.2 Pistes pour des recherches ultérieures.....	303
8.3 Applications pratiques .....	304
<b>Bibliographie</b> .....	307
<b>Appendice 1</b> .....	317
<b>Appendice 2</b> .....	327

# 1. INTRODUCTION

## 1.1 Objet et objectifs

La recherche portant sur le discours scientifique<sup>1</sup> a connu un intérêt croissant ces dernières années. Cet essor va de pair avec l'expansion de l'idée que le discours scientifique est un genre rhétorique plutôt qu'un simple moyen de transmettre des résultats de recherches scientifiques. Contrairement à la conception générale et traditionnelle, selon laquelle le discours scientifique se caractérise par son ton neutre et objectivant, sa non-subjectivité, les chercheurs qui s'intéressent à l'analyse du discours scientifique s'accordent aujourd'hui sur le fait que ce genre de discours vise non seulement à informer, mais aussi à convaincre (voir par exemple Markkanen et Schröder 1997 : 9). On peut dire, avec Breivega (2003 : 11), que les textes scientifiques ne servent pas à transmettre des « vérités scientifiques », mais plutôt à les discuter et les promouvoir. Le discours scientifique n'est plus appréhendé comme un objet indépendant d'un sujet écrivain, mais comme un discours dont les procédés rhétoriques employés dépendent d'un ici, un maintenant et un sujet locuteur (cf. Beaufrère-Bertheux 1997), c'est-à-dire de la situation d'énonciation. Dans cette optique, le discours scientifique devient un objet d'analyse intéressant et légitime, pour les linguistes du texte, les sociologues et rhétoriciens qui cherchent à comprendre les mécanismes, les normes (souvent implicites, voire tacites) et les motivations qui influencent la nature de ce type de discours ainsi qu'à décrire ses traits caractéristiques et leurs effets communicationnels. Dans le courant rhétorico-linguistique, l'analyse du discours scientifique constitue un champ de recherche qui s'intéresse avant tout à la manière de transmettre le contenu, par exemple à travers la structure textuelle, les stratégies rhétoriques, les choix lexicaux etc.

La présente étude s'inscrit dans cette orientation et en particulier, dans la tradition relativement récente d'études comparatives sur le discours scientifique. Alors que le discours scientifique anglais fait depuis longtemps l'objet d'une littérature abondante, cela ne vaut pas pour d'autres langues, comme le français et le norvégien, et par ailleurs, d'éventuelles

---

<sup>1</sup> Le concept de *discours scientifique* tel qu'il sera employé dans le présent travail, désigne le discours (écrit ou oral) qui a lieu entre académiciens et chercheurs dans les milieux universitaires et qui porte sur les questions d'ordre professionnel, indépendamment de l'appartenance disciplinaire. Le terme *scientifique* n'est pas restreint aux « sciences » dans le sens strict de sciences de la nature, mais inclut toute branche académique. Notre concept de *discours scientifique* sera donc employé comme l'équivalent du terme anglais « *academic discourse* ». Nous nous intéresserons ici au discours scientifique *écrit*, et nous nous intéressons à la communication qui a lieu entre experts du même champ scientifique et non au discours pédagogique (communication entre enseignants et étudiants) ou à la vulgarisation (communication visant à diffuser les résultats des recherches à un plus grand public) (cf. Poudat 2006).

différences entre l'anglais et d'autres langues sont demeurées peu explorées jusqu'à ces dernières années. Malgré l'apparition d'un certain nombre de travaux qui se proposent d'étudier le discours scientifique à travers les langues (voir par exemple Mauranen 1993, Vassileva 1998, Ventola et Mauranen 1996 et Fløttum, Dahl et Kinn 2006), il apparaît que plusieurs questions restent en suspens. C'est vrai notamment du phénomène linguistique qui sera investigué dans la présente étude, à savoir la modalité épistémique, phénomène qui a été peu étudiée dans une perspective comparative. Le recours à la modalité épistémique, ou au marquage d'incertitude, constitue une stratégie d'atténuation fréquemment employée dans le discours scientifique et jugée essentielle pour les auteurs de textes scientifiques spécialisés. L'atténuation en général représente un phénomène qui a été abondamment étudié dans le discours scientifique anglais, mais par rapport auquel il existe très peu d'études comparatives. Bien qu'on ait vu récemment la publication de quelques études qui comparent l'emploi d'atténuateurs entre disciplines (Varttala 2001, Hyland 1999 a et b, 2004), le nombre d'études comparant différentes langues reste très limité. Afin de contrebalancer quelque peu cet état de choses, la présente étude adoptera une perspective doublement comparative, au niveau des langues et des disciplines, en se proposant d'examiner l'emploi de modalisateurs épistémiques dans un corpus d'articles de recherche tirés des disciplines de médecine et de linguistique et rédigés en français, norvégien et anglais. La comparaison donnera lieu à une approche quantitative et à une approche qualitative, mais on se centrera sur le français et le norvégien pour l'approche qualitative. Ce sont donc le français et le norvégien qui jouent les rôles principaux dans la présente étude.

Etant donné que les conventions de rédaction scientifique – y compris les conventions d'atténuation – varient selon les langues et les branches professionnelles, une approche contrastive est jugée particulièrement fructueuse dans la recherche sur la rédaction scientifique (voir Markkanen et Schröder 1997 : 12-15). Comme nous l'avons dit, les connaissances actuelles sur d'éventuelles variations entre langues et disciplines dans le domaine de la modalité épistémique restent assez limitées, notamment en ce qui concerne le paramètre des langues. Par son approche interlinguistique, la présente étude vise à contribuer au comblement de cette lacune. Il importe de combler cette lacune, car un manque de connaissances en matière de conventions culturelles ou linguistiques pourrait mener à des obstacles et des malentendus dans une situation de communication interculturelle. La présente étude vise à fournir des connaissances sur l'usage effectif d'atténuateurs dans le discours scientifique des trois langues, et par cela contribuer à une sensibilisation augmentée des conventions culturelles partagées ou divergentes qui existent dans ce domaine. En ce qui

concerne le paramètre des disciplines, il entend contribuer à l'ensemble de recherches qui peuvent augmenter nos connaissances sur les variations disciplinaires. Comme il a déjà été mentionné, cet ensemble reste assez limité, et cela vaut en particulier pour les variations disciplinaires dans d'autres langues que l'anglais<sup>2</sup>. Dans l'ensemble, l'examen de l'emploi d'atténuateurs peut nous dire quelque chose sur la manière dont sont transmis les résultats de recherches et les idées de l'auteur, ce qui à son tour nous aide à développer une compréhension de ce que c'est la science ; comment se fait la construction des connaissances et comment se fait la transmission de ces connaissances ?

La mondialisation dans le secteur académique exige qu'un grand nombre de chercheurs s'expriment par écrit ou oralement dans une langue qui ne soit pas leur propre langue, ce qui explique que le manque d'études sur la variation interlinguistique soit fort regrettable. En particulier, le grand nombre de chercheurs non-anglophones qui désirent ou doivent publier ou communiquer en anglais pourraient profiter de telles études pour perfectionner leur anglais langue de spécialité. Étant donné que l'anglais est devenu dans une large mesure la *lingua franca* du discours scientifique, la maîtrise de l'anglais comme langue de spécialité devient de plus en plus importante pour les chercheurs du monde entier, et cette compétence ne requiert pas seulement une compétence linguistique, mais aussi une compétence culturelle : pour réussir, il faut connaître les normes propres au discours scientifique anglais aussi bien que les règles grammaticales de la langue générale. Une conscience explicite de la manière dont ces normes diffèrent des normes du discours scientifique de leur propre langue maternelle pourrait faciliter la tâche de rédaction d'article ou de communication dans un anglais approprié. Soulignons que la comparaison anglais – français – norvégien entreprise dans cette étude se veut utile non seulement pour les non-anglophones qui doivent rédiger leurs textes scientifiques en anglais, mais aussi, et peut-être avant tout, pour les Norvégiens (étudiants, doctorants, enseignant-chercheurs) qui écrivent en français.

Le présent travail est effectué dans le cadre du projet KIAP ([www.uib.no/kiap](http://www.uib.no/kiap)) – un projet linguistique lancé à l'université de Bergen en 2002 dans le but d'étudier le genre de l'article de recherche à travers langues et disciplines. La double perspective comparative adoptée dans la présente étude (comparaison entre disciplines et entre langues) est donc issue de ce projet<sup>3</sup>.

---

<sup>2</sup> Rappelons que les études mentionnées plus haut (Varttala 2001, Hyland 1999 a et b, 2004) portent toutes sur le discours scientifique anglais.

<sup>3</sup> Pour les résultats du projet KIAP, voir Fløttum, Dahl & Kinn 2006.

À travers l'analyse d'une série de phénomènes linguistiques, le projet KIAP cherche à examiner dans quelle mesure on peut parler d'identités culturelles différentes dans le discours scientifique, et si ces identités dépendent plus de la discipline scientifique pour laquelle on écrit ou de la communauté langagière auquel appartient l'auteur (voir Fløttum 2003a : 8). Pour ce faire, le projet KIAP a établi une large base électronique comprenant 450 articles de recherche tirés de trois disciplines différentes (médecine, linguistique et économie politique) et écrits en trois langues différentes (anglais, français et norvégien). Les articles sont tous tirés de revues reconnues qui sélectionnent leurs articles à partir d'une validation par un comité scientifique de lecture, souvent constitué par deux rapporteurs. Les données sur lesquelles se fondent les analyses de la présente étude sont tirées du corpus KIAP. Plus précisément, nos données consistent en une sélection de 120 articles de recherche, extraits de la première partie de ce corpus (voir [http://www.uib.no/kiap/corpus\\_1en.htm](http://www.uib.no/kiap/corpus_1en.htm)) et équitablement répartis en fonction des disciplines et des langues étudiées. Pour des raisons pratiques, la présente étude se restreint à deux disciplines, à savoir la médecine et la linguistique, et l'économie politique n'est donc pas retenue dans ce travail. La médecine et la linguistique ont été retenues parce qu'elles représentent deux domaines différents des sciences : elles reflètent l'opposition traditionnelle entre sciences naturelles et sciences humaines. Dans une telle perspective, l'économie politique – en tant que science sociale – se situe entre ces deux domaines<sup>4</sup>.

Les analyses du projet KIAP sont centrées autour de trois questions principales, formulées dans Fløttum (2003a : 8) et reprises ici :

- 1) Dans quelle mesure et de quelle manière l'auteur se manifeste-t-il ?
- 2) Dans quelle mesure les points de vue (ou voix) d'autres chercheurs se manifestent-t-ils ?
- 3) Comment l'auteur « vend-il » sa recherche ?

La présente étude vise à apporter des éléments de réponses à toutes ces questions. La première parce qu'un modalisateur épistémique est une marque de subjectivité – il exprime le jugement du locuteur par rapport à la valeur de vérité d'un énoncé –, la deuxième parce que les modalisateurs épistémiques peuvent transmettre une prédiction ou une évaluation de points de vue soutenus par d'autres chercheurs, et la troisième parce que l'auteur peut employer les modalisateurs épistémiques pour des raisons de persuasion et de négociation, c'est-à-dire en vue de faire accepter sa recherche aux membres de la communauté scientifique en question. Pour réussir, il faut que l'auteur soit conçu comme un chercheur sérieux et sincère qui a de la

---

<sup>4</sup> Pour une discussion de la division catégorique entre ces deux domaines, voir *infra*, section 2.2.2



crédibilité, et entre autres procédés, les atténuateurs peuvent contribuer à une telle image de l'auteur (cf. Markkanen et Schröder 1997 : 10).

La problématique principale de la présente étude peut se résumer dans les questions suivantes, formulées à l'exemple de celles posées par le projet KIAP :

- Dans quelle mesure y a-t-il des différences et des ressemblances entre les disciplines et les langues examinées en ce qui concerne l'emploi de modalisateurs épistémiques, et en quoi consistent ces différences et ressemblances ?
- Dans quelle mesure les différences reflètent-elles des identités culturelles différentes, et est-ce que celles-ci sont liées principalement à l'appartenance langagière ou disciplinaire ?

Avant de formuler des questions de recherche plus précises sur ce sujet, une présentation préliminaire du phénomène étudié, la modalité épistémique, est de mise.

## **1.2 La modalité épistémique et l'atténuation : une présentation préliminaire**

Les notions de *modalité épistémique* et d'*atténuation* ('*hedging*' en anglais) sont souvent mal distinguées dans la littérature (cf. *infra*, ch. 4) et parfois utilisées de façon interchangeable. Personnellement, nous opterons pour une différenciation entre les deux termes (qui néanmoins sont intimement liés), et afin d'éviter la confusion fréquemment observée sur ce point, cette thèse s'efforce de préciser les relations et les distinctions qui existent, à notre avis, entre ces termes.

Selon l'optique adoptée dans la présente étude, l'emploi de modalisateurs épistémiques exprimant l'incertitude constitue une stratégie d'atténuation, c'est-à-dire que les modalisateurs épistémiques d'incertitude constituent un type parmi d'autres d'atténuation. L'atténuation est un phénomène dit caractéristique du discours scientifique (voir par exemple Lewin 2005 : 164 et Hyland 2000 : 179), et le concept renvoie à des marqueurs, des tournures ou des stratégies discursives dont l'auteur se sert pour exprimer des précautions ou des incertitudes par rapport à ce qu'il dit. L'emploi de tels marqueurs et tournures peut être motivé par un véritable manque de savoir définitif, mais aussi simplement par un désir de répondre aux conventions de l'écriture scientifique, étant donné que l'atténuation est considérée comme un phénomène plus ou moins routinier de ce type d'écriture (Salager-Meyer 1997).

Les modalisateurs épistémiques sont quant à eux des marqueurs linguistiques qui qualifient de façon explicite la valeur de vérité d'un énoncé. Telle qu'elle sera définie ici, la catégorie de la modalité épistémique englobe des expressions de certitude aussi bien que des expressions d'incertitude, mais ce sont uniquement ces dernières qui seront examinées dans le présent travail. Il traite des modalisateurs épistémiques en tant qu'atténuateurs, et les expressions de certitude sont donc exclues. Voici quelques exemples de modalisateurs épistémiques d'incertitude<sup>5</sup> :

1. Bien qu'à première vue ce phénomène **paraisse** lié au nombre de lits disponibles, car la durée de séjour **semble** être inversement proportionnelle au nombre d'admissions, il est **plus vraisemblablement** causé par l'augmentation du nombre d'hospitalisations de jeunes enfants à l'automne, clientèle dont le séjour moyen est plus court. (frmed02)

Les marqueurs en gras dans l'exemple ci-dessus expriment une évaluation épistémique de la part du locuteur, c'est-à-dire un jugement sur la fiabilité de l'information transmise. Comme c'est le cas pour les atténuateurs en général, les modalisateurs épistémiques s'utilisent tantôt pour exprimer une incertitude réelle, tantôt comme un moyen stratégique employé pour des raisons d'ordre interpersonnel ou interactionnel, par exemple dans le but de convaincre, d'aller au devant d'une critique anticipée, d'apparaître modeste etc. La qualification épistémique constitue une dimension fondamentale de toute communication scientifique, étant donné le rôle essentiel des académiciens dans la construction et la transmission des connaissances. Il est donc communément admis que la qualification épistémique représente une stratégie rhétorique fondamentale pour les auteurs d'articles de recherche scientifique (voir par exemple Hyland 1996 a et b, 1998).

### 1.3 Questions de recherche et déroulement de l'étude

Comme évoqué plus haut, le but principal de cette étude est d'examiner dans quelle mesure il y a des différences et des ressemblances entre les différentes langues et disciplines en ce qui concerne l'emploi de modalisateurs épistémiques et de préciser en quoi consistent ces différences et ces ressemblances. Plusieurs facettes de l'emploi de modalisateurs épistémiques seront examinées, que l'on peut regrouper sous deux aspects principaux : la fréquence d'un

---

<sup>5</sup> Tout au long de cette thèse, les modalisateurs épistémiques dans les exemples seront typographiquement marqués par des caractères gras. En outre, les exemples tirés de notre corpus sont tous suivis d'un code qui indique de quel article ils sont tirés. Ce code associe la langue (fr = français, no = norvégien, eng = anglais), la discipline (ling = linguistique, med = médecine) et un numéro. Le code frmed02 renvoie ainsi à l'article numéro 2 du sous-corpus « articles de médecine écrits en français ». On peut se référer à l'appendice 1 pour retrouver les références bibliographiques complètes.

côté (l'aspect quantitatif) et les fonctions pragmatiques de l'autre (l'aspect qualitatif). Tout au long de l'étude, l'approche adoptée sera doublement comparative : les variations langagières aussi bien que disciplinaires seront investiguées. En ce qui concerne l'aspect quantitatif, les trois langues (ainsi que les deux disciplines) seront comparées. Quant à l'aspect qualitatif, l'anglais sera mis à côté, étant donné que le domaine fonctionnel d'atténuateurs a été abondamment étudié pour le discours scientifique anglais (voir par exemple Hyland 1998, Salager-Meyer 1994, 1997, Varttala 1999, 2001, Myers 1989, Markkanen and Schröder 1997). Pour ce deuxième aspect, la comparaison entre langues sera donc restreinte au français et au norvégien.

Plusieurs questions se posent tant au niveau de la fréquence des modalisateurs qu'au niveau de leurs fonctions. En ce qui concerne le niveau de fréquence, les questions suivantes seront étudiées :

- En se basant sur des analyses du discours scientifique anglais, bon nombre de chercheurs ont conclu que l'atténuation constitue un trait caractéristique du discours scientifique. Dans quelle mesure une telle observation est-elle valable aussi pour les discours scientifiques français et norvégien ?
- Quels sont les moyens linguistiques les plus utilisés pour exprimer la modalité épistémique ? Dans quelle mesure les différents groupes utilisent-ils différents types formels (moyens grammaticaux, moyens lexicaux) et différents types sémantiques de marqueurs (verbes de cognition, semi-auxiliaires de perception etc.) ?

Au niveau du fonctionnement pragmatique, les questions suivantes seront étudiées :

- Quel est le rôle des modalisateurs épistémiques dans les textes, c'est-à-dire quelles sont leurs visées communicatives ? Dans quelle mesure servent-ils les mêmes fonctions pragmatiques dans les différentes langues et disciplines ?

Pour répondre à ces questions, une analyse en trois temps sera entreprise, avec une combinaison de méthodes quantitatives et qualitatives :

Premièrement, c'est l'aspect quantitatif qui retiendra notre attention : est-ce que les modalisateurs épistémiques varient en fréquence selon les langues et les disciplines ? Si oui, quelle peut être l'explication de telles variations ?

Une des questions relatives à cette partie de l'analyse concernera la fréquence de différents types de marqueurs : la distribution de différents types formels de marqueurs pourra nous dire quelque chose sur les moyens linguistiques dont se servent les différentes langues pour exprimer la modalité, et la distribution de différents types sémantiques de marqueurs peut contribuer à éclairer la question de savoir dans quelle mesure la source de la qualification épistémique est explicitée ou « objectivée ».

En ce qui concerne l'aspect quantitatif, nous incluons aussi un troisième facteur, celui du sexe de l'auteur. Ce facteur de variation ne fait pas partie des variables principales explorées par le projet KIAP, mais il paraît particulièrement intéressant à propos du phénomène de l'atténuation ; en effet, l'atténuation, y compris la modalisation épistémique, a souvent été considérée comme une caractéristique du discours féminin (cf. la discussion menée par Markkanen et Schröder (1997 : 8-9), qui renvoient, à ce propos, aux études de Robin Lakoff (1975) et de Preisler (1986)). Le bien-fondé de cette corrélation supposée entre sexe et modalisation épistémique ne semble cependant pas validé par des études empiriques. La présente étude pourra contribuer à déterminer la validité de cette hypothèse<sup>6</sup>.

Deuxièmement seront effectuées des analyses approfondies de certains marqueurs spécifiques, sélectionnés en premier lieu sur la base de leur fréquence. Ces analyses aborderont l'aspect qualitatif : à travers l'analyse contextuelle des occurrences des marqueurs sélectionnés, nous chercherons à déterminer quelles sont les fonctions pragmatiques que ces marqueurs remplissent dans le texte. De même, une étude du cotexte<sup>7</sup> immédiat doit permettre d'élucider dans quelle mesure et de quelle manière l'auteur se manifeste dans son texte.

Troisièmement, quatre études de cas illustreront de façon supplémentaire cet aspect pragmatique-communicationnel : l'examen d'articles individuels permet d'analyser de manière plus fine comment un auteur, en se servant de modalisateurs épistémiques dans le texte, donne de lui l'image d'un chercheur honnête et sérieux.

## 1.4 Plan de la thèse

Cette étude se scinde en deux parties majeures : la première partie présente le fondement théorique du présent travail et le situe dans un cadre de recherche plus large. Cette partie s'ouvre par un chapitre qui traite du genre de l'article de recherche (2.1) et de l'approche comparative adoptée (2.2). Y sont signalées les différences les plus importantes entre la

---

<sup>6</sup> Pour une présentation des différences de sexe par rapport à d'autres phénomènes linguistiques, voir Fløttum et al. 2006.

<sup>7</sup> Dans la présente étude, nous utiliserons le terme *cotexte* pour désigner le contexte linguistique, c'est-à-dire le texte autour d'un marqueur ou d'un énoncé. Par *contexte* nous entendons la situation du discours.

rédaction d'articles de médecine et celle d'articles de linguistique. Un bref aperçu de recherches antérieures pertinentes sur l'article scientifique est également inclus, ce qui nous fournit des éléments pertinents pour avancer quelques hypothèses générales. Vient ensuite une présentation de la double perspective comparative, s'intéressant notamment aux défis d'une telle approche. Le chapitre 3 est consacré à la description théorique des catégories linguistiques qui font l'objet de cette étude, à savoir la modalité (3.1) et l'épistémicité (3.2). Les multiples approches qui existent en linguistique au sujet de ces catégories appellent une clarification de ces notions. C'est notamment le concept de modalité qui est problématique, étant donné les divergences dans la manière dont il est conçu. Dans la section 3.1, nous présenterons certaines approches de ce concept avant de retenir celle qui servira de base pour la présente étude. Pour clore ce chapitre théorique sur la notion de modalité épistémique, un inventaire des différents moyens linguistiques susceptibles d'exprimer la modalité épistémique est établi dans la section 3.3. Le chapitre 4 est consacré au phénomène d'atténuation en général : il s'ouvre par une présentation théorique de la notion elle-même, et met l'accent sur ses relations et divergences avec la notion de modalité épistémique ; vient ensuite un compte-rendu de quelques résultats pertinents issus d'études antérieures sur l'atténuation dans le discours scientifique. Ces résultats fournissent une base de comparaison pour nos propres analyses ainsi qu'une assise permettant de postuler des hypothèses plus concrètes sur les possibles variations interlinguistiques et interdisciplinaires dans ce domaine.

Une fois le cadre théorique posé on passera aux analyses du corpus. Le chapitre 5 donne une présentation des matériaux et trace les grandes lignes de la méthodologie adoptée. Dans les chapitres 6 et 7 seront respectivement exposés les résultats de l'analyse quantitative et qualitative. Le chapitre 6 s'ouvre par un résumé des hypothèses étudiées et des questions de recherche (6.1) ; nous présenterons ensuite les principes suivis lors de la sélection de marqueurs (6.2), avant de passer aux analyses mêmes. Commencerons par une analyse exploratoire (6.3), qui a pour but de servir de base à la sélection de marqueurs à étudier dans l'analyse quantitative ; nous passerons alors à l'analyse quantitative, sa méthodologie et ses résultats (6.4). Le chapitre 7 traite de l'aspect pragmatique des modalisateurs épistémiques. Il s'ouvre par une présentation générale des fonctions pragmatiques des atténuateurs (7.1) ; ensuite nous étudierons de plus près l'emploi de certains marqueurs sélectionnés à l'aide d'analyses cotextuelles (7.2-7.5). Quatre études de cas portant sur des articles closent ce chapitre (7.6). La conclusion (chapitre 8) comprend une synthèse des résultats principaux et propose quelques pistes à suivre pour des travaux ultérieurs.

Au cours de la thèse nous nous référons fréquemment à des exemples ; ceux-ci sont numérotés consécutivement, mais la numérotation recommence à chaque section principale (par exemple à 3.1, 3.2 et 3.3). Comme nous l'avons mentionné plus haut, les exemples qui relèvent de notre corpus sont tous suivis d'une parenthèse indiquant l'article dont l'exemple est tiré.

# PREMIÈRE PARTIE

## THÉORIE





## **2. GENRE ÉTUDIÉ ET APPROCHE ADOPTÉE**

### **2.1 Le genre de l'article de recherche**

Dans cette section, nous allons d'abord faire quelques remarques sur les caractéristiques du genre de l'article de recherche en général (2.1.1), avant de signaler certaines spécificités distinctives des articles médicaux et linguistiques (2.1.2). Ensuite sera donné un aperçu de recherches antérieures sur le genre de l'article de recherche (2.1.3).

#### **2.1.1 L'article de recherche en général**

L'étude du discours scientifique peut se réaliser à travers plusieurs sous-genres, écrits (monographies, articles de recherche, rapports, résumés) et oraux (communications de conférence, séminaires). Le présent travail se base, comme les autres études liées au projet KIAP, sur un corpus électronique composé d'articles de recherche. Le genre de l'article de recherche est celui qui a reçu le plus d'attention dans les travaux existants, ce qui s'explique par le fait qu'il représente un moyen essentiel de communication entre chercheurs. L'article de recherche possède un statut particulier dans le monde de la recherche, à la fois par sa profusion et son importance, et il occupe une position centrale dans la construction et la communication des savoirs au sein de la communauté scientifique. De plus, il est plus facilement abordable que certains des autres genres écrits, comme les monographies, à la fois par sa longueur relativement limitée et par son accessibilité. Tout cela fait du genre de l'article de recherche un choix privilégié pour celui qui s'intéresse au discours scientifique écrit.

L'article de recherche constitue un genre relativement « fermé », en ce sens qu'il s'adresse à un public restreint et spécialisé, le plus souvent aux confrères de l'auteur, c'est-à-dire à d'autres spécialistes de la même branche scientifique (Swales 2002). Il est peut-être le moyen principal dont disposent les chercheurs pour échanger des idées et des résultats scientifiques. Selon Breivega (2003 : 19), les articles de recherche se distinguent d'autres textes spécialisés en ce qu'ils cherchent à documenter une activité qui est gouvernée par des exigences scientifiques particulières. Ce n'est pas uniquement le processus de recherche qui est gouverné par de telles règles, car il y a également des exigences (souvent assez rigoureuses) liées à l'expression écrite elle-même : l'article de recherche représente un genre fortement structuré et régulé qui obéit à des contraintes plus ou moins rigides tant au niveau de la structure qu'au niveau stylistique. Cependant, la rigueur de ces contraintes varie

considérablement selon les domaines scientifiques ; l'article de recherche des sciences naturelles est par exemple beaucoup plus standardisé que celui des sciences humaines (Breivega 2003). Certaines caractéristiques sont toutefois communes à tous les domaines. Par exemple, étant donné l'idéal scientifique de s'exprimer de façon claire, précise et correcte, les articles de recherche se caractérisent par une terminologie hautement spécialisée dont la compréhension suppose souvent des connaissances étendues dans le domaine de recherche en question. L'article de recherche constitue donc un exemple type de communication de spécialistes à spécialistes.

Le genre de l'article de recherche trouve son origine dans les lettres informatives que les hommes de sciences se sont toujours écrits les uns aux autres, mais par l'établissement de périodiques spécialisés au cours du XVII<sup>ème</sup> siècle, il s'est transformé peu à peu en l'article de recherche tel qu'on le connaît aujourd'hui (Swales 2002 : 110). L'objectif de l'article de recherche a toujours été d'informer, de transmettre des données et des connaissances, et son style est traditionnellement supposé objectif, neutre et préférablement dépourvu de traces énonciatives telles que par exemple les pronoms personnels de la première personne. En témoigne le grand nombre de guides de rédaction scientifique qui déconseillent l'emploi de traces subjectives et favorisent les structures impersonnelles et la voix passive.

Néanmoins, les recherches sur l'écriture scientifique qui ont été effectuées ces dernières décennies ont montré que les textes scientifiques spécialisés ne sont pas nécessairement des textes rédigés dans un style objectif et impersonnel<sup>1</sup> et dont le seul but est d'informer, mais qu'au contraire, il s'agit souvent de textes empreints de subjectivité et porteurs de divers procédés rhétoriques qui visent à persuader. Cet aspect persuasif de l'article scientifique va de pair avec un aspect interactionnel : par son texte, l'auteur négocie avec son public et cherche à lui faire accepter les prémisses et les idées avancées dans l'article. Le texte est dialogique (au sens bakhtinien) ou polyphonique (voir par exemple Nølke, Fløttum et Norén 2004 sur la théorie ScaPoLine) au sens où il est attentif aux lecteurs et leurs éventuelles remarques critiques : en imaginant le flux de questions et de remarques que peuvent poser les lecteurs, l'auteur essaye d'aller au devant de la critique en formulant des réponses appropriées avant même que les questions ne soient posées. Ainsi, il tente de convaincre son public et de justifier sa recherche. L'article de recherche est aussi interactionnel dans un sens plus concret, dans la mesure où l'auteur fait sans cesse référence à d'autres travaux et d'autres chercheurs

---

<sup>1</sup> Selon Lundquist (1990 : 71-72), les styles objectif et impersonnel vont souvent de pair et se caractérisent par l'absence de marques de personne et d'expressions évaluatives. Les tournures impersonnelles, les constructions indéfinies et l'emploi du passif sont privilégiés.

dans le même domaine spécialisé. Cependant, dans la présente étude nous nous intéresserons moins à ces traces évidentes d'autres voix qu'aux traces subtiles d'une polyphonie plus implicite.

On pourrait donc conclure que l'article de recherche se caractérise par au moins trois aspects, théoriquement différents et pourtant intimement liés, qui peuvent être plus ou moins saillants dans différents textes. Le premier est l'aspect informatif, car bien qu'on mette ces dernières années de plus en plus l'accent sur la persuasion, la rhétorique et le besoin du chercheur de se positionner dans un milieu compétitif, l'objectif fondamental de tout article de recherche est d'informer la communauté scientifique concernée de nouvelles connaissances, issues d'études qui ont été effectuées en vue de faire progresser le champ concerné. Le deuxième est l'aspect persuasif : par son texte, l'auteur cherche à persuader le lecteur du bien-fondé de ses points de vue, il cherche à positionner son étude par rapport à d'autres et à faire accepter ses idées par le lecteur. Le troisième est l'aspect interactif : il est d'une importance capitale pour le chercheur de montrer qu'il fait partie d'une communauté, ce qui requiert l'inclusion (explicite et/ou implicite) d'autres voix scientifiques dans le texte.

Ces trois aspects sont présents dans les articles des sciences naturelles aussi bien que dans ceux de sciences humaines. L'objectif est d'informer *et* de faire accepter les idées émises dans l'article, c'est-à-dire de convaincre le lecteur. L'article de recherche, quelle que soit la discipline, peut par conséquent être considéré comme un genre rhétorique (Breivega 2003 : 10, Hyland 1998 : 16).

Il ne s'agit pourtant pas d'un genre homogène. Le style d'un texte scientifique spécialisé est déterminé par les conventions et les pratiques discursives propres au contexte socioculturel ou socioprofessionnel dans lequel il est produit. Ces conventions et pratiques sont susceptibles de varier selon les disciplines académiques, les pays et les langues. En ce qui concerne la question de disciplines, on distingue généralement, parmi les multiples champs scientifiques qui existent, deux domaines majeurs : les sciences naturelles et les sciences humaines, qui seront toutes deux représentées ici, par les disciplines de médecine et de linguistique. Comme nous l'avons mentionné plus haut, il y a des différences évidentes entre les pratiques d'écriture de ces deux domaines. Dans les sciences naturelles, la structuration de l'article de recherche est beaucoup plus standardisée et l'exigence d'un ton impersonnel est beaucoup plus forte que dans les sciences humaines. La traditionnelle conception de l'article de recherche comme un genre objectif et neutre s'est sans doute basée sur la pratique d'écriture propre aux sciences naturelles. Ce style est en quelque sorte devenu un idéal, un style auquel aspirent certaines branches des sciences sociales et humaines. Mais on aurait tort

de penser que le style impersonnel caractéristique des articles des sciences naturelles équivaut nécessairement à l' « objectivité » et la « neutralité ». Selon Swales (2002 : 112), ce style impersonnel fut développé dans le but de *faire croire* au destinataire qu'il n'y a pas de rhétorique, qu'il n'y a pas de persuasion de la part de l'auteur, mais au contraire, que les faits parlent d'eux-mêmes. On peut en conclure que si l'argumentation et la persuasion se font de diverses manières dans les différentes disciplines, il y a toujours une dimension rhétorique (cf. Breivega 2003 : 11).

Pour résumer, on peut dire que malgré la multiplicité des cultures scientifiques existantes et leurs différences, les articles de recherche de diverses communautés scientifiques ont certains traits en commun qui justifient de les traiter comme un genre<sup>2</sup> ; en particulier, ils documentent une activité pour laquelle il existe des exigences scientifiques rigoureuses, ils visent à informer et persuader les lecteurs, ils s'adressent aux experts dans un domaine particulier et ils font usage d'une terminologie spécialisée.

### **2.1.2 Les articles de recherche en médecine et en linguistique**

Les articles spécialisés de différentes disciplines possèdent aussi, évidemment, des traits caractéristiques spécifiques. Une série de différences distinguent ainsi les articles de médecine et ceux de linguistique. Nous avons déjà mentionné le style impersonnel caractéristique du discours des sciences naturelles qui se fait valoir également en médecine (cf. Gilhus 2003). Une deuxième différence qui se manifeste d'emblée entre l'article de médecine et l'article de linguistique est celle de la structure : alors que les articles de médecine suivent rigoureusement une structure standardisée, à savoir le schéma IMRED<sup>3</sup>, les articles de linguistique ne sont pas structurés selon ce même format imposé. Les différences au niveau de la structure sont aisément identifiables et sont peut-être celles qui distinguent le plus nettement les articles issus de la recherche en sciences naturelles et ceux issus de la recherche en sciences humaines. Or, le format IMRED gagne du terrain, et s'est répandu dans plusieurs branches des sciences humaines et sociales, parmi lesquelles la linguistique, sans toutefois s'établir comme une structure obligatoire (Breivega 2003 : 22). Dans la discipline de linguistique, le format IMRED est devenu de plus en plus courant, au point que certaines

---

<sup>2</sup> On pourrait cependant dire que ce genre se subdivise en plusieurs sous-genres : Swales (2004) propose de diviser le genre de l'article de recherche en trois sous-genres séparés : l'article expérimental, l'article théorique et l'article de synthèse.

<sup>3</sup> Acronyme pour Introduction, Matériel et Methodes, Résultats Et Discussion, plus connu sous la forme anglaise IMRAD. Toutes ces rubriques, à l'exception de l'introduction, sont souvent subdivisées à leur tour en plusieurs sous-sections qui portent, elles aussi, des titres assez standardisés.

guides de rédaction pour les étudiants présentent cette structure comme une norme (voir par exemple Björk et Räisänen 2003 : 275). Toutefois, aucun des articles de linguistique dans notre corpus ne suit rigoureusement ce schéma, bien qu'il y en ait quelques-uns qui s'en rapprochent, plus précisément certains rapports de recherches expérimentales.

Le plan IMRED est beaucoup plus qu'une chronologie imposée de rubriques obligatoires. Chaque rubrique remplit des fonctions rhétoriques particulières (Swales 2002) et pour chaque rubrique, il y a des règles bien précises à respecter. L'Introduction, par exemple, comporte obligatoirement une formulation des objectifs de la recherche et une justification de l'étude effectuée. La rubrique Méthode doit fournir une description pas à pas du processus suivi, détaillé de manière à ce qu'il soit théoriquement possible, pour d'autres, de le reproduire et d'obtenir les mêmes résultats. Les résultats sont présentés dans la section intitulée Résultats, mais *sans commentaire* de la part de l'auteur. On y trouve des chiffres et de la statistique, mais aucune interprétation, car celle-ci doit indispensablement être remise à la section suivante. Une fois les résultats présentés, l'auteur peut, dans la section Discussion, évaluer et interpréter les résultats et tirer des conclusions sur leurs implications (cf. Breivega 2003 : 22-24, voir aussi Skelton 1997). La section Discussion est donc résolument la partie la plus « subjective » d'un article de recherche des sciences naturelles, et elle est, comme nombre d'études l'ont montré (Salager-Meyer 1994, Varttala 1999, 2001 : 165), celle qui contient le plus grand nombre d'atténuateurs. C'est souvent dans cette section que les aspects persuasif et interactif sont les plus saillants.

La structure d'un article linguistique n'est pas standardisée au même point. Les articles de linguistique sont constitués de composantes facultatives qui suivent le plus souvent une organisation thématique où les sous-titres sont motivés par le contenu des sections plutôt que par la fonction que celles-ci remplissent (voir Rinck 2006). Vu le degré relativement faible de standardisation (par rapport à l'article de médecine), le linguiste est plus libre en composant son texte, bien qu'il doive, évidemment, observer lui aussi certaines normes. La tâche de rédiger un article scientifique se présente peut-être comme plus difficile pour un étudiant ou un jeune chercheur en linguistique que pour un étudiant ou jeune chercheur en médecine, car le médecin-chercheur a recours à des règles bien explicites, alors que le linguiste se trouve face à une série de normes tacites (cf. Breivega 2003 : 242). De plus, ce dernier a probablement reçu moins de conseils sur ce sujet au cours de sa formation que son collègue en médecine. Ce n'est que ces dernières décennies que l'on commence à enseigner la rédaction scientifique aux étudiants dans les disciplines des sciences humaines, et bien que cette pratique se soit instituée dans des pays anglo-américains et nordiques, elle reste peu

répandue en France (Poudat 2006 : 24, 66). Cela rend particulièrement importante l'étude des pratiques discursives et textuelles propres aux diverses sciences humaines, étant donné que c'est là où le besoin d'explicitation semble le plus grand.

Une troisième différence concerne la longueur des articles. Les articles en médecine sont généralement beaucoup plus brefs que les articles de linguistique. Selon Gilhus (2003 : 22), l'idéal dans la rédaction médicale est d'écrire d'une manière aussi brève et concise que possible, et les revues les plus prestigieuses tendent à contenir les articles les plus courts. Un tel idéal n'existe pas en linguistique. Cette différence de longueur d'article se manifeste très clairement dans notre corpus, où les articles de linguistique constituent deux tiers de l'ensemble des mots.

Il convient aussi de mentionner que les contraintes éditoriales pour les articles en médecine sont souvent strictes et très détaillées. Les éditeurs de certaines revues médicales exigent par exemple que la section Résultats soit écrite au passé (Skelton 1997 : 130). Les linguistes doivent aussi tenir compte de contraintes éditoriales, mais elles ne sont probablement pas aussi détaillées.

La principale différence entre les articles de recherches en sciences naturelles et les articles en sciences humaines est donc que les premiers sont soumis à des contraintes formelles plus rigides. Il s'ensuit qu'on peut raisonnablement s'attendre à une plus grande homogénéité dans le corpus d'articles médicaux que dans celui d'articles linguistiques. Or, il existe des variations culturelles : les contraintes semblent être beaucoup plus strictes pour les articles destinés à un public international et publiés en anglais dans les grandes revues internationales que pour les articles qui s'adressent à une communauté scientifique d'une taille plus restreinte et qui sont publiés dans des revues nationales (voir Poudat 2006).

### **2.1.3 Recherches antérieures sur l'article de recherche**

Si le discours scientifique constitue un champ d'étude relativement récent, les recherches dans ce domaine ont pris un essor considérable à partir des années 1980, et beaucoup de chercheurs se sont intéressés au genre de l'article de recherche ces dernières années. Comme nous l'avons dit, la recherche porte avant tout sur le discours scientifique anglais, pour lequel les études sont si nombreuses et forment un tissu si dense de correspondances qu'on parle d'un courant, le courant de l'*English for Specific Purposes*, voire de l'*English for Academic Purposes*. Ici ne sera donné qu'un petit aperçu de ceux des travaux qui nous semblent les plus pertinents pour la présente étude. Il s'agit, d'un côté, d'études générales qui décrivent le genre et qui ont ainsi contribué à la compréhension de sa complexité et de ses multiples aspects, et,

de l'autre côté, d'études plus spécifiques (souvent dédiées à un nombre limité de phénomènes isolés et visant une comparaison de deux ou plusieurs langues et /ou disciplines) dont les résultats peuvent nous servir de base dans l'élaboration d'hypothèses. Notons que cette section traite des recherches sur divers aspects de l'article scientifique ; les recherches qui portent directement sur l'usage de l'atténuation en général ou de la modalité épistémique en particulier et qui peuvent ainsi être mises en perspective avec nos propres résultats, seront abordées plus en détail dans le chapitre 4.

Parmi les premières études qui ont mis en relief la nature rhétorique des articles de recherche, on trouve celle de Bazerman (*Shaping Written Knowledge*, 1988) et celle de Swales (*Genre Analysis*, 2002 [1990]). Bazerman met en évidence que les stratégies rhétoriques et persuasives sont importantes dans des textes scientifiques et que le choix de ces stratégies varie selon la nature de l'objet d'étude et par conséquent, selon les disciplines. Il déplore le manque d'étude sur la rhétorique des sciences et préconise avec insistance la réalisation de recherches dans ce domaine, principalement à des fins didactiques (1988 : 332). Aujourd'hui, le courant rhétorico-linguistique *English for Academic Purposes*, exauce ce souhait.

L'ouvrage très influent de Swales (2002) montre comment les introductions des articles de recherches expérimentales suivent un schéma particulier composé de plusieurs stades ou « démarches » qui remplissent des fonctions bien précises : établir un domaine de recherche digne d'intérêt, se faire une niche, c'est-à-dire adresser des questions jusque-là peu ou pas explorées, et ensuite occuper cette niche en présentant son propos et ses objectifs (Swales 2002 : 141). Par son analyse, Swales met en évidence l'importance d'éléments rhétoriques dans les introductions des articles de recherche : il s'agit de justifier et de promouvoir sa recherche, c'est-à-dire que l'article de recherche a un but persuasif aussi bien qu'informatif.

Dans son article datant de 1989, Myers met en relief la nature interactive du discours scientifique et montre comment différentes stratégies de politesse (y compris l'atténuation) s'utilisent dans des articles de recherche, signalant, de façon plus ou moins directe, l'interaction entre auteur et lecteur. Il montre comment l'auteur d'un article de recherche doit être sensible au public pour lequel il écrit, à ses points de vue, ses éventuelles réactions et son estime de soi ; de fait, tout élément de critique doit être accompagné d'une stratégie discursive qui permette de préserver la face des tenants des points de vue critiqués.

Par leur mise en relief des aspects persuasif et interactif de l'article de recherche, ces études sont parmi les premières à mettre en question la conception traditionnelle de l'article



de recherche comme neutre, objectif et exempt de toute dimension rhétorique. Par ses études sur la présence de l'auteur et d'autres voix dans les articles de recherche, le projet KIAP met en évidence et réaffirme à la fois la nature persuasive et la nature interactive de l'article de recherche, en concluant que l'aspect interactif est tout aussi important que celui de persuasion (Fløttum et al. 2007 : 36), bien que ce soit peut-être ce dernier qui ait le plus attiré l'attention des chercheurs.

Ce sont les articles rédigés en anglais qui ont été le plus amplement étudiés. En ce qui concerne le discours scientifique en norvégien, les études sont peu nombreuses, mais Breivega (2003) a effectué un travail pionnier dans le domaine en comparant la composition textuelle d'articles norvégiens en histoire, médecine et linguistique. Breivega s'intéresse surtout aux stratégies argumentatives employées dans les différentes disciplines, et suite à ses analyses, elle conclut que les trois disciplines font usage de différentes stratégies argumentatives, mais que les articles des linguistes et des historiens sont très proches les uns des autres. Dans les articles de linguistique et d'histoire, les auteurs font usage de stratégies argumentatives ouvertes et explicites : l'auteur se manifeste clairement dans son texte, il prend explicitement en charge ses propositions et les qualifications évaluatives et épistémiques sont nombreuses (ibid. : 242). Breivega suggère que cette pratique d'écriture pourrait être considérée comme une pratique d'écriture caractéristique des sciences humaines, étant donné qu'elle se manifeste avec autant de force dans deux disciplines de sciences humaines aussi différentes que la linguistique et l'histoire (ibid. : 245). Les médecins-chercheurs, de leur côté, font preuve de stratégies argumentatives plus couvertes. Si ces stratégies sont moins aisément identifiables que les traces énonciatives dans les articles de linguistique et d'histoire, elles n'en sont pas moins présentes. De plus, Breivega note une grande différence entre les différentes sections de l'article médical ; dans la section Discussion, les traces énonciatives, y compris les qualifications évaluatives et épistémiques, sont souvent nombreuses, alors qu'une telle argumentation ouverte est inconcevable dans les sections plus descriptives (Méthode, Résultats).

Par sa perspective comparative anglais-français-norvégien, les études de KIAP (sur lesquelles nous reviendrons) contribuent aussi à la description et la compréhension du norvégien comme langue de spécialité. Les études comparatives de ce type représentent une approche récente dans la recherche sur le discours scientifique, mais elles continuent à s'accroître en nombre. Vassileva (1998), par exemple, a comparé l'emploi des pronoms sujets de la première personne dans des articles de linguistique écrits en cinq langues : l'anglais, l'allemand, le français, le russe et le bulgare. L'analyse révèle que la présence de l'auteur est



la plus manifeste dans les articles anglais, alors que les auteurs français sont les moins « visibles ». Cela corrobore l'observation de Loffler-Laurian (1980) qui parlait de « le moi haïssable » après avoir constaté que le « je » était totalement absent de son corpus d'articles scientifiques français tirés des disciplines de physique et de chimie – en dépit du fait que certains de ces articles étaient rédigés par un seul auteur. Elle souligne pourtant que l'auteur d'un article en sciences humaines peut intervenir dans son discours plus facilement que les chercheurs en sciences dures.

Ces tendances sont également confirmées par les études de Fløttum (2003b) et de Fløttum et al. (2006), qui comparent l'emploi de pronoms personnels dans des articles de recherche à travers diverses langues (anglais, français, norvégien) et disciplines (économie, linguistique, médecine). L'analyse montre que les auteurs anglais sont ceux qui utilisent le plus les pronoms sujets de la première personne, alors que les auteurs français les utilisent le moins. Quant aux disciplines, ces pronoms s'emploient majoritairement dans les textes linguistiques et économiques, et le moins dans les textes médicaux.

La présence de l'auteur telle qu'elle se manifeste par l'emploi de pronoms personnels, est probablement l'aspect le plus examiné dans les recherches comparatives sur le discours scientifique écrit. Cependant, la présence de l'auteur dans son texte ne se limite pas à ces marques les plus évidentes, et certains d'autres aspects, comme la rhétorique de la critique (Salager-Meyer et al. 2003), ont aussi attiré l'intérêt des chercheurs. Salager-Meyer et al. (2003) ont relevé des différences culturelles dans le discours médical en français, anglais et espagnol, ce qui montre que même au sein d'un discours aussi standardisé que l'est le discours médical, il existe des variations culturelles considérables et des traits propres à certaines langues ou cultures. Plus spécifiquement, l'étude de Salager-Meyer et al. a montré que la fréquence de remarques critiques ouvertes et directes était beaucoup plus élevée chez les auteurs français et espagnols que chez les auteurs anglais (ibid. : 232). Ces derniers ont tendance à présenter la critique sous une forme atténuée et plus couverte.

Mauranen (1993) adopte une approche plus holistique en ce qu'elle ne s'intéresse pas à un phénomène isolé, mais à l'organisation générale des textes scientifiques. A travers une analyse de textes scientifiques écrits en anglais et en finnois, elle compare et contraste les pratiques rhétoriques des deux mondes scientifiques. La conclusion est que la pratique d'écriture anglo-américaine est plus explicitement orientée vers le lecteur que la pratique finnoise, en ce sens que les auteurs anglo-américains utilisent plus d'éléments métatextuels pour guider le lecteur et ils annoncent clairement dès le début les principaux points de vue soutenus dans l'article. Mauranen caractérise le style finnois comme plus indirect et implicite

(1993 : 252 sqq), ce qui demande au lecteur de faire un effort plus grand pour saisir le texte, mais en même temps cela lui permet davantage une part interprétative. Le style anglo-américain peut donc être caractérisé comme ‘*reader-friendly*’<sup>4</sup> (voir Clyne 1998) ou ‘*writer responsible*’, alors que le style finnois se caractérise plutôt comme ‘*reader-responsible*’ (pour reprendre les termes de Hinds 1987).

Des différences culturelles similaires ont été observées par Lundquist (1998). Par le biais d’une analyse textuelle de deux jugements écrits<sup>5</sup>, elle compare les discours juridiques français et danois. L’analyse montre que le texte français est d’une complexité et d’une densité extrêmes, alors que le texte danois est beaucoup plus facilement accessible. Le texte français exige donc du lecteur un effort beaucoup plus grand que ne le fait le texte danois.

Les résultats des études menées par l’équipe KIAP – qui adopte, rappelons-le, une perspective doublement comparative de langues et disciplines – corroborent l’impression que la pratique d’écriture anglo-américaine favorise un style dialogique, alors que la pratique d’écriture française semble moins orientée vers le lecteur (Fløttum et al. 2006 : 168-169, 264, Dahl 2004). Les investigations du projet KIAP révèlent également que la présence de l’auteur telle qu’elle se manifeste à travers les phénomènes linguistiques examinés est la plus saillante dans les articles de linguistique et la plus faible dans les articles de médecine. Au niveau des langues, elle est la plus nette dans les textes anglais, qui sont immédiatement suivis des textes norvégiens, alors que les textes français font preuve d’un taux beaucoup plus bas de traces de l’auteur (Fløttum et al. 2006 : 260). La linguistique se présente comme plus ouvertement argumentative que les autres disciplines (ibid. : 263), qui ne sont pas aussi polémiques. En ce qui concerne la question de savoir lequel des deux facteurs de variation culturelle est le plus important, l’équipe KIAP conclut que pour la majorité des phénomènes investigués, c’est la discipline qui l’emporte sur la langue. En d’autres termes, les conventions d’écriture déterminées par une discipline semblent plus homogènes ou plus rigides que celles déterminées par une culture nationale ou une communauté linguistique<sup>6</sup>. Ainsi, on peut dire, en généralisant, que les médecins-chercheurs, par exemple, rédigent leurs articles selon les normes prescrites par la tradition d’écriture médicale, quelle que soit la communauté

---

<sup>4</sup> Traduit en français par le terme *style dialogique* (Poudat 2006 : 62).

<sup>5</sup> L’étude de Lundquist ne porte pas sur les articles de recherche, mais sur un autre type de textes spécialisés, à savoir les jugements juridiques.

<sup>6</sup> La juxtaposition des termes *culture nationale* et *communauté linguistique* ne doit absolument pas entraîner l’impression que ces deux termes désignent la même chose. La relation entre ces termes tient simplement à ce que les langues sont les reflets de traditions culturelles et intellectuelles (cf. Poudat 2006 : 48), et l’étude de langues peut donc aussi nous dire quelque chose sur les cultures qu’elles représentent.

linguistique à laquelle ils appartiennent. Pour certains des phénomènes investigués, le facteur de langue s'avère pourtant assez pertinent<sup>7</sup>.

Dans leur étude sur les résumés scientifiques, Melander, Swales et Fredrickson (1997) adoptent aussi une perspective doublement comparative, en analysant des résumés écrits en suédois et en anglais, et relevant des disciplines de la médecine, la biologie et la linguistique. Leur conclusion générale est similaire à celle du projet KIAP : le facteur disciplinaire joue un rôle plus important pour les pratiques d'écriture que le facteur langue (1997 : 267). Cependant, l'importance du facteur langue varie selon les disciplines : en linguistique, Melander et al. observent une grande variation interlinguistique (mais aussi intralinguistique, du moins pour les résumés écrits en suédois), alors qu'il y a une plus grande homogénéité dans les deux autres groupes, notamment celui de la biologie. La question de savoir dans quelle mesure les disciplines disposent d'un ensemble de pratiques standardisées au niveau international varie donc considérablement d'une discipline à l'autre.

De ce parcours rapide des recherches antérieures, on peut conclure qu'alors que le discours scientifique anglo-américain est largement décrit, les discours scientifiques français et norvégien demeurent peu étudiés. Il est pourtant clair que les pratiques d'écriture varient selon les contextes culturels, disciplinaires autant que langagiers. Si le style anglo-américain se répand et s'est même instauré comme une norme dans beaucoup de disciplines (cf. Poudat 2006 : 52), ces études montrent que le discours scientifique dépend toujours du contexte culturel et langagier dans lequel il est produit, et qu'il respecte les normes propres à ces cultures.

Finalement, il convient de noter un point soulevé par Salager-Meyer et al. dans leurs études diachroniques. Ces études (Salager-Meyer et Defives 1998, Salager-Meyer 2000, Salager-Meyer et al. 2003) nous rappellent que les normes propres à une communauté scientifique particulière évoluent au cours des années – elles ne sont pas stables, mais dynamiques et en constant développement. Cette dynamique prend probablement de plus en plus d'importance au fur et à mesure que les champs scientifiques s'internationalisent et par conséquent s'influencent réciproquement. Cependant, la présente étude adopte une perspective synchronique et c'est le discours scientifique contemporain qui sera étudié.

---

<sup>7</sup> Ce résumé succinct des principaux résultats du projet KIAP est évidemment bien trop schématisé et ne rend pas justice aux fines nuances repérées ; il n'y a cependant pas lieu ici de fournir autre chose qu'une vue d'ensemble des tendances générales qui se sont dessinées. Pour une présentation détaillée des résultats issus de ce projet, nous renvoyons à Fløttum et al. 2006.

Comment tous ces résultats peuvent-ils être rapportés à la question des modalisateurs épistémiques? Premièrement, les études comparatives évoquées plus haut indiquent que les normes concernant l'usage de traces personnelles dans un texte scientifique (notamment les pronoms personnels de la première personne, qui sont le plus amplement investigués) varient selon les cultures. En ce qui concerne les cultures langagières, on peut dire que la culture anglo-américaine accepte, voire préconise, l'usage de telles traces, alors que la culture française le déconseille (Vassileva 1998, Loffler-Laurian 1980, Fløttum 2003b). En ce qui concerne les cultures disciplinaires, un tel usage est attendu dans la tradition des sciences humaines, tandis qu'il est évité dans les sciences naturelles (Fløttum 2003b, Breivega 2003). Dans la mesure où les marqueurs épistémico-modaux sont aussi des traces personnelles de l'auteur, on pourrait s'attendre aux mêmes tendances quant au phénomène de la modalité épistémique, c'est-à-dire qu'il y aurait plus de marques dans les articles écrits en anglais et dans les articles de linguistique que dans les articles écrits en français et dans les articles de médecine. Cependant, il faut rappeler que les marques linguistiques de la modalité épistémique ne vont pas nécessairement de pair avec les pronoms personnels de la première personne, et il existe d'autres alternatives, d'apparence moins subjectives, pour exprimer des jugements épistémiques que les *I suggest* et *I assume*<sup>8</sup> si plébiscités parmi les linguistes anglais (cf. Fløttum et al 2006 : 93).

Deuxièmement, l'observation de Salager-Meyer et al. (2003), selon laquelle les auteurs anglais ont tendance à présenter la critique sous une forme atténuée et plus couverte, pourrait indiquer que dans la culture scientifique romane, on a tendance à s'exprimer d'une manière plus directe que dans la culture anglo-américaine, et que, par conséquent, les atténuateurs seraient plus nombreux dans les articles anglais que dans les articles français.

Troisièmement, le style dialogique favorisé par la culture anglo-américaine implique un plus grand degré d'interaction entre auteur et lecteur, entre autres en ce que l'auteur s'efforce d'anticiper les réactions du lecteur, ce qui pourrait mener à un emploi plus extensif d'atténuateurs.

Quatrièmement, le fait que les linguistes adoptent un style plus ouvertement argumentatif que les médecins-chercheurs pourrait porter à croire que ces premiers font un usage plus fréquent d'atténuateurs, dans la mesure où l'atténuation représente une stratégie d'argumentation (cf. *infra*, ch. 4).

---

<sup>8</sup> Notons que les séquences *I suggest* et *I assume* n'expriment pas nécessairement une modalité épistémique : *suggest* et *assume* ont d'autres sens que le sens épistémico-modal (cf. *infra*, section 6.4.1.3).

A partir de ce qui est déjà connu sur la rédaction scientifique médicale et linguistique et sur les discours scientifiques anglais, français et norvégien, on peut donc postuler quelques hypothèses préliminaires au sujet de l'usage des modalisateurs épistémiques. En ce qui concerne l'aspect quantitatif, on pourrait émettre les deux suivantes :

- Il y aura un usage plus fréquent de modalisateurs épistémiques dans les articles anglais que dans les articles français. La position du norvégien est peu claire, mais étant donné la parenté typologique entre les langues norvégienne et anglaise, la grande influence que la culture anglo-américaine exerce sur la culture norvégienne, et le fait que les discours scientifiques anglais et norvégiens se sont avérés assez similaires à bien des égards (Fløttum et al. 2006), il est raisonnable de s'attendre à un plus grand degré de similitude entre l'anglais et le norvégien qu'entre l'anglais et le français ou entre le français et le norvégien.
- Il y aura un usage plus fréquent de modalisateurs épistémiques dans les articles linguistiques que dans les articles médicaux ; cette hypothèse est étayée par l'observation de Hyland (1999 a et b), montrant que la fréquence d'atténuateurs varie beaucoup entre les disciplines des sciences naturelles et les disciplines des sciences humaines et sociales, en ce sens qu'il y a plus d'atténuation dans ces dernières (cf. *infra*, ch. 4). Cependant, il faut admettre que la médecine et la linguistique ne sont pas des représentants typiques de leur domaines scientifiques respectifs (voir Fløttum et al. 2006), et selon Varttala (1999 : 178), la médecine est réputée avoir un emploi d'atténuateurs relativement élevé. Ces observations rendent difficile la formulation d'hypothèses concrètes sur cette question.

Quant à la question de la variation intragroupe, nous pouvons, toujours en nous fondant sur les observations évoquées ci-dessus, émettre l'hypothèse suivante :

- Il y aura un plus grand degré de similitude à l'intérieur du groupe médecine que dans le groupe linguistique, étant donné que la rédaction médicale est fortement standardisée et régulée.

En ce qui concerne l'aspect fonctionnel, il n'est pas possible, à ce stade, de formuler des hypothèses, étant donné que c'est avant tout à partir des exemples attestés dans le corpus que peuvent être réalisées nos descriptions des fonctions pragmatiques des modalisateurs

épistémiques. L'étude de fonctions est donc essentiellement de nature exploratoire, bien que nous soyons partie de certaines catégories principales (cf. *infra*, ch. 7).

Suite à la présentation de recherches antérieures sur l'atténuation, nous chercherons à préciser quelque peu ces hypothèses générales. Mais avant de passer au phénomène de l'atténuation en général, il convient de préciser ce que nous entendons par modalité épistémique, qui est la catégorie linguistique dont fait l'objet ce travail. L'objectif du chapitre suivant (ch. 3) est justement de délimiter la catégorie de modalité épistémique et de rendre compte des moyens linguistiques qui l'expriment. Tout d'abord seront présentées quelques réflexions sur l'approche comparative adoptée dans cette étude (section 2.2).

## 2.2 Approche adoptée

Dans cette section, nous précisons d'abord la position de la présente étude par rapport aux traditions de recherche dans les études comparatives (2.2.1), pour proposer ensuite quelques éléments de réflexion sur les défis qu'une approche doublement comparative peut poser (2.2.2).

### 2.2.1 Une double perspective comparative

Cette étude adopte une perspective doublement comparative, en ce qu'elle compare d'un côté différentes langues et de l'autre côté différentes disciplines. Cette approche est héritée du projet KIAP, dont l'originalité réside en grande partie précisément dans cette double perspective comparative.

Comme il a déjà été mentionné, des analyses transdisciplinaires ont ces dernières années pris leur essor dans la recherche sur le discours scientifique (voir par exemple Hyland et Bondi 2006). Les études interculturelles ou interlinguistiques ont aussi commencé à s'accroître en nombre, bien que les variations langagières à propos du phénomène de l'atténuation n'aient pas reçu beaucoup d'attention. Malgré le fait que ces deux types d'études contrastives continuent de se multiplier, le projet KIAP est l'un des premiers à combiner ces deux axes de comparaison. Dans la mesure où l'on peut parler de différentes identités culturelles propres à différentes communautés scientifiques, l'objectif de la double approche comparative adoptée par le projet KIAP est de déterminer quelle dimension influence le plus cette identité : est-ce que c'est l'appartenance disciplinaire qui situe les chercheurs dans une culture scientifique spécifique ou est-ce que c'est plutôt la communauté linguistique à laquelle ils appartiennent qui décide de leur identité culturelle scientifique ? Il va de soi que la réponse n'est pas tranchée, mais que discipline et langue contribuent toutes deux, dans des mesures différentes, à l'identité scientifique du chercheur, et influencent ainsi sa manière de s'exprimer par écrit dans ses articles. Les études qui combinent les deux axes de comparaison (telles que le projet KIAP) peuvent donc offrir des observations et des résultats susceptibles d'accroître notre compréhension de l'*interaction* des divers facteurs qui interviennent, que ceux-ci relèvent des domaines disciplinaire ou langagier.

Les analyses comparatives et contrastives sont donc centrales dans cette étude, sans pour autant qu'elle s'inscrive dans l'approche théorique appelée *linguistique contrastive* ou *analyse contrastive*, qui traditionnellement aborde notamment des questions d'acquisition d'une langue étrangère et de traduction. L'objectif de la présente étude n'est pas de comparer

et contraster différentes langues en général, mais d'étudier un genre discursif particulier, à savoir l'article scientifique, et ses variations à travers différentes communautés linguistiques, et donc à travers différentes cultures. Ainsi, cette étude est plus proche de la tradition appelée *rhétorique contrastive*, qui se développe depuis la deuxième partie des années 60 et qui s'intéresse aux variations interculturelles dans l'usage de stratégies rhétoriques, notamment dans les *textes* (Connor 2004 : 271). Selon Connor (2004 : 272), le terme de rhétorique contrastive a acquis certaines connotations négatives au cours des années, et il n'y a pas d'unanimité sur l'étendue de ce terme ni sur les types d'études qu'il désigne. Bien que le présent travail ne s'inscrive pas directement dans le cadre théorique de la rhétorique contrastive, il peut sans doute fournir des résultats pertinents pour cette tradition de recherche, par l'accent mis sur la variabilité interculturelle de l'usage d'un outil rhétorique comme l'est la modalisation épistémique. Si l'on rencontre dans cette thèse des termes comme 'analyse comparative' ou 'perspective contrastive', cela réfère tout simplement au fait que des textes issus de différentes cultures ou milieux scientifiques et rédigés en différentes langues sont comparés. Ils ne doivent pas être entendus comme appartenant à une terminologie propre à une approche théorique particulière.

### **2.2.2 Quelques remarques sur les difficultés liées à la perspective comparative**

Le choix d'une approche doublement comparative ne va évidemment pas sans problèmes. Les différents sous-corpus sont-ils effectivement comparables ? Une telle comparabilité suppose que chaque sous-corpus soit suffisamment homogène – parce que les textes d'un sous-corpus doivent avoir quelque chose en commun qui permette de les contraster avec d'autres groupes de textes –, mais aussi qu'il soit suffisamment hétérogène, parce qu'il est censé représenter une population plus large de textes.

Pour aborder d'abord la dimension de la discipline : une comparaison de différentes disciplines ne néglige-t-elle pas les considérables variations qui existent à l'intérieur d'une même discipline ? La discipline de la linguistique par exemple, comprend toute une série de branches diverses, de la linguistique computationnelle et la linguistique diachronique à la linguistique textuelle et appliquée, de la phonologie à la pragmatique, en passant par la morphologie et la sémantique, et ainsi de suite. En outre, la linguistique embrasse un grand nombre d'orientations théoriques divergentes, comme le générativisme et la grammaire fonctionnelle pour ne mentionner que ces exemples. On trouvera une hétérogénéité semblable au sein de la discipline de la médecine, qui englobe des sciences fondamentales telles que la biochimie et la microbiologie ainsi que des sciences cliniques, comme la chirurgie, et



également des branches qui s'orientent vers la santé publique, comme la médecine sociale et la santé au travail. Les disciplines de linguistique et de médecine ont en commun le fait d'englober un grand nombre de sous-disciplines dont certaines sont plus « dures » et d'autres moins. Bien que la médecine soit censée représenter les sciences naturelles, et la linguistique les sciences humaines, il existe des branches médicales « molles » et des branches linguistiques « dures ». Afin de répondre à ce problème, le projet KIAP a élaboré certains critères pour guider la sélection des articles : en ce qui concerne les articles de médecine, ont été sélectionnés principalement des articles portant sur des enjeux de la médecine dite « dure », c'est-à-dire que les branches plus « molles » de la médecine, telle que la psychiatrie et la médecine sociale ont été, en grande partie, évitées. Les articles de linguistique traitent généralement de la langue dans laquelle ils sont écrits, c'est-à-dire que les articles français traitent de linguistique française, les articles norvégiens traitent du norvégien, etc. On a cherché à éviter la linguistique générale proprement dite. Par ailleurs, les revues d'où sont tirés les articles déterminent dans une certaine mesure les orientations théoriques et les sujets thématiques représentés. Cela assure une certaine homogénéité au sein des corpus linguistique et médical. Si les principes évoqués ci-dessus ne sont pas toujours suivis rigoureusement (voir Fløttum et al. 2006 : 8sq), les textes de chacun des sous-corpus ont pourtant suffisamment de traits en commun pour qu'on puisse justifier leur regroupement dans une même discipline scientifique. Il y a évidemment une considérable diversité aussi, une diversité qui est tout à fait voulue, pour des raisons évidentes de généralisabilité des résultats<sup>9</sup>.

La même question de variation *intra*culturelle se pose pour la comparaison de différentes langues. Il existe évidemment différentes variétés des trois langues. Cela vaut du moins pour l'anglais et le français, langues qui sont parlées dans plusieurs pays et dans différentes parties du monde. Si l'on peut établir un lien relativement étroit entre le norvégien et la culture nationale norvégienne, une telle correspondance entre langue et culture nationale ne se fait pas aussi aisément pour ce qui est de l'anglais et du français. Or, les articles anglais du corpus KIAP sont généralement issus de la Grande Bretagne et (dans un moindre degré) des Etats-Unis et on peut donc dire qu'ils représentent la culture anglo-américaine. En outre, il semble clair que la pratique d'écriture scientifique dite anglo-américaine dépasse les frontières nationales en ce sens que les textes scientifiques rédigés en anglais par des locuteurs natifs semblent répondre aux mêmes normes et représentent alors, sinon la même

---

<sup>9</sup> Pour plus d'information sur les critères de sélection qui ont guidé la constitution du corpus KIAP, voir Fløttum et al. 2006 ou la page web <http://www.uib.no/kiap/mdcorpusdescr.htm>.

culture nationale, la même culture d'écriture (cf. Fløttum et al. 2006 : 18)<sup>10</sup>. En ce qui concerne les articles français, ceux-ci sont généralement issus de la France et on peut donc dire qu'ils représentent la culture d'écriture française. Cependant, un nombre limité d'articles sont issus du Canada (la partie francophone), et il ne serait pas déraisonnable de supposer que ceux-ci soient plus influés par la pratique d'écriture anglo-américaine que ceux issus de la France. Si de telles variations potentielles ne font pas l'objet d'études systématiques dans le présent travail, nous prendrons tout de même en compte leur possible influence sur les résultats dans l'analyse de nos observations.

Une troisième question est celle du genre. Est-ce que les articles écrits en norvégien ou en français appartiennent tout à fait au même genre que ceux écrits en anglais ? Dans les sciences naturelles, on n'accorde pas la même valeur à un article rédigé dans la langue nationale et publié dans une revue nationale qu'à un article rédigé en anglais et publié dans une revue internationale. Selon Melander (1998), il peut s'agir là non seulement d'une différence de statut, mais aussi d'une différence de genre, étant donné que – dans certaines disciplines, comme par exemple la médecine – les articles rédigés dans la langue nationale (dans le cas de l'étude de Melander, il s'agit du suédois) n'assument pas les mêmes objectifs communicatifs que ceux écrits en anglais. Melander note que les articles médicaux écrits en suédois s'adressent à une communauté de praticiens plutôt qu'à la communauté scientifique proprement dite, c'est-à-dire aux médecins-chercheurs. Les « vrais » *articles scientifiques* (Melander 1998 : 223, notre traduction) sont rédigés en anglais et publiés dans des revues internationales.

Il est bien possible qu'il y ait une différence entre d'un côté les articles rédigés en anglais, dont une grande partie représentent une recherche internationale extrêmement compétitive<sup>11</sup>, et de l'autre côté les autres articles et notamment ceux écrits en norvégien, étant donné que ceux-ci s'adressent à un petit public national. Néanmoins, la revue d'où sont tirés les articles norvégiens est une revue de qualité, qui suit la procédure normale d'évaluation scientifique par rapporteurs anonymes, et qui se lit extensivement par les

---

<sup>10</sup> Cela ne veut évidemment pas dire qu'il ne peut pas y avoir des différences entre les variétés. Kourilová (1994 : 645) suggère par exemple qu'il y a une différence entre l'anglais américain et l'anglais britannique sur le plan de l'atténuation, en ce sens que la contrainte de s'exprimer avec politesse, prudence et modestie serait moins manifeste dans le discours scientifique des États-Unis que dans celui de la Grande Bretagne. De telles différences éventuelles entre diverses variétés ne seront pas un enjeu dans la présente étude. Le fait qu'il puisse exister de telles différences n'empêche pas que les différentes variétés d'une langue ont quelque chose en commun qui rend légitime le fait de les comparer avec d'autres langues.

<sup>11</sup> Rappelons que l'anglais a ici un double rôle ; il est la *lingua franca* du monde scientifique, mais en même temps, il est la langue nationale et maternelle d'un grand nombre de chercheurs. Tous les articles scientifiques écrits en anglais ne représentent donc pas nécessairement « une recherche internationale extrêmement compétitive ».

médecins-chercheurs en Norvège ainsi que par leurs collègues dans les autres pays nordiques. Elle est d'ailleurs la seule revue scientifique norvégienne en recherche médicale, et si l'on se propose d'étudier le discours scientifique médical en norvégien, elle est incontournable. Il faut pourtant prendre en compte les différences liées à l'envergure du public et aux objectifs communicatifs en interprétant les résultats.

En ce qui concerne la discipline de la linguistique, la problématique évoquée par Melander (1998), qui se rapporte avant tout aux sciences naturelles, semble moins pertinente. Les articles linguistiques dans le corpus KIAP traitent généralement de recherches sur les langues respectives, et les spécialistes des langues nordiques s'adressent naturellement en premier lieu à leurs pairs dans les pays nordiques et se servent donc des voies qui existent dans leur propres pays. L'objectif communicatif est pourtant le même que celui des articles publiés dans des périodiques internationales ; il s'agit de diffuser ses idées et ses observations aux autres membres de la communauté scientifique à laquelle on appartient. La revue de linguistique norvégienne dont sont issus les articles de recherche investigués dans cette étude (*Norsk Lingvistisk Tidsskrift*) doit par conséquent être considérée comme une « vraie revue scientifique » et les articles qui y sont publiés sont des exemples d'une communication de spécialistes à spécialistes.

Il faut pour finir mentionner un problème pertinent pour toute étude comparative : le fameux problème d'équivalence<sup>12</sup> (voir par exemple Johansson 2003). Comment savoir si les modalisateurs épistémiques du français sont comparables à ceux du norvégien ou de l'anglais, par exemple en termes de contenu sémantique ou d'effets pragmatiques ? Comme Johansson (2003) l'indique, ce qui est exprimé par des auxiliaires modaux dans une langue peut être exprimé de manières différentes dans une autre. Afin de répondre à ce problème bien connu, nous avons choisi de ne pas comparer directement des expressions dites « équivalentes », mais de partir d'une catégorie sémantique qui existe dans les trois langues (et qui par cela constitue un *tertium comparationis*), pour ensuite relever et enregistrer les différents moyens d'expression de cette catégorie dans les différentes langues, prises indépendamment les unes des autres. Autrement dit, au départ, c'est la fonction plutôt que les formes qui a été l'aspect central, et la première question a été : par quels moyens linguistiques exprime-t-on une modalité épistémique (côté incertitude) dans les trois langues ? Pour répondre à cette question, une analyse exploratoire d'une partie du corpus a été entreprise. Les expressions ou les formes repérées sont de différents types, mais la fonction est la même : il s'agit d'exprimer

---

<sup>12</sup> 'The problem of equivalence'

une modalité épistémique. De surcroît, cette analyse exploratoire sera précédée d'une présentation des différents moyens linguistiques dont disposent les différentes langues pour exprimer la modalité épistémique. En cela nous suivons le principe évoqué par Nølke (1991 : 202), disant qu'avant d'aborder une analyse contrastive, on doit, dans l'idéal, connaître tous les moyens linguistiques susceptibles d'exprimer la même fonction que les expressions examinées. La liste de ces moyens linguistiques servira donc de base pour les investigations exploratoires qui visent à révéler l'usage effectif des modalisateurs épistémiques dans les articles de recherche.

Le point de départ a donc été la fonction plutôt que les formes, et pourtant les formes qui résultent sont dans une large mesure, comme nous allons le voir, équivalentes dans les trois langues. Cela indique que l'expression d'incertitude se fait à l'aide de procédés similaires dans les trois langues, au moins dans le genre de l'article scientifique. Les procédures et les résultats de l'analyse exploratoire seront plus amplement présentés dans la section 6.3.

### 3. LA MODALITÉ ÉPISTÉMIQUE

Il convient à ce stade de situer notre objet d'étude, *la modalité épistémique*, dans un cadre linguistique plus large, ainsi que de clarifier les relations entre des catégories linguistiques apparentées telles que la modalité, l'évidentialité et l'épistémicité. La nature de ces relations n'est pas évidente – les chercheurs ne s'accordent pas sur la question de savoir quels sont les termes génériques et quels sont les catégories plus spécifiques. Constatons tout d'abord que, dans l'optique qui sera adoptée dans cette étude, la catégorie à étudier – c'est-à-dire la modalité épistémique – relève de deux catégories principales : la modalité et l'épistémicité – d'où l'appellation *la modalité épistémique*. La modalité épistémique se situe donc au carrefour de ces deux catégories, lesquelles regroupent pourtant, chacune de leur côté, d'autres sous-catégories que celle de la modalité épistémique (voir figure 1) :

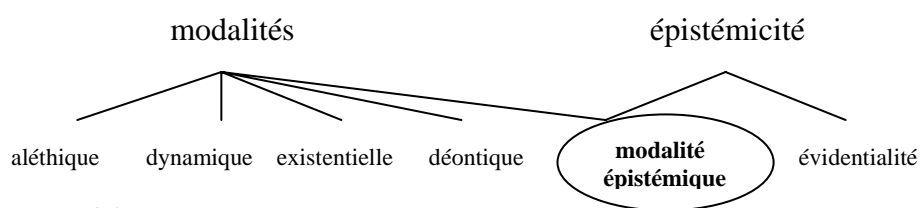


Figure 1 : Modalité et épistémicité

Les sections 3.1 et 3.2 seront consacrées à la définition et la délimitation de ces deux catégories supérieures et à la précision des relations et des distinctions entre les différentes sous-catégories. Pour le moment, retenons les définitions préalables suivantes : La *modalité* réfère aux moyens langagiers qui expriment « l'attitude du locuteur par rapport au contenu propositionnel de son énoncé » (Le Querler 1996 : 14), et l'*épistémicité* réfère aux moyens langagiers dont dispose le locuteur pour s'exprimer sur ses connaissances. On verra par la suite comment ces définitions peuvent être affinées et mises en discussion. Considérons d'abord la catégorie de la modalité.

## 3.1 La modalité

Cette section se scinde en deux parties : tout d'abord, il importe de circonscrire et préciser la notion de modalité, notion polymorphe qui a fait l'objet d'une pléthore d'approches théoriques. Tel sera donc le but de la première partie, alors que la seconde sera consacrée à un exposé des différents types de modalité.

### 3.1.1 Circonscrire la catégorie de la modalité

#### 3.1.1.1 Introduction et définition de départ

Force est de constater que la modalité représente une notion problématique en linguistique. Le Querler (1996 : 49), en reprenant le terme de Meunier (1981), parle de la « nébuleuse » que constituent les théories des modalités et Veters (2001 : 173) parle d'une « situation catastrophique » dans le domaine de la modalité. Il n'existe pas de définition précise unanimement acceptée, et la conception linguistique de la notion de modalité est en général floue, vague et difficile à délimiter. Il est presque de règle dans un ouvrage linguistique portant sur la modalité de souligner les grandes difficultés reliées à la caractérisation de ce terme. Perkins (1983 : 1) écrit à cet égard : « it is by no means easy to find out what modality actually is », et Herslund (1989 : 7) appelle la notion de modalité « this elusive notion » en soulignant que « it is [...] no easy matter to state briefly what modality is », pour ne citer qu'eux.

Nous tenterons dans ce chapitre de mettre un peu d'ordre dans la catégorie floue de la modalité, de sorte qu'on puisse y voir plus clair. Un tel détour est nécessaire afin de pouvoir distinguer d'une manière claire la modalité épistémique d'autres types de modalités et ainsi faciliter le travail consistant à identifier avec précision les véritables modalisateurs épistémiques employés dans le corpus, à l'opposé des marqueurs modaux relevant d'autres domaines de la modalité. Une conscience des différents types de modalités est particulièrement importante quand on a affaire à des marqueurs polysémiques susceptibles d'exprimer plusieurs types de modalités, tels que le verbe *pouvoir* et son correspondant norvégien *kunne*.

Avant d'aborder les divers types de modalité, nous allons faire le tour de différentes conceptions et traditions et voir comment elles convergent et divergent. Le propos ici n'est pas de donner une présentation exhaustive de toutes les théories de la modalité qui ont été proposées au fil des années – un objectif aussi ambitieux serait voué à l'échec – mais de présenter certaines des conceptions les plus habituelles en signalant les aspects fructueux

aussi bien que les côtés plus problématiques. En guise de point de départ, reprenons la caractérisation proposée par Le Querler (1996 : 14), selon laquelle la modalité est « l'expression de l'attitude du locuteur par rapport au contenu propositionnel de son énoncé ». Cette caractérisation correspond à une conception très habituelle en linguistique, selon laquelle la modalité est conçue comme une sorte de commentaire de la part du locuteur sur un contenu énonciatif. Nous verrons par la suite comment cette définition de départ peut être mise en discussion, notamment en l'opposant à la conception originelle de la modalité telle qu'elle est définie dans la tradition philosophico-logique. Pour le moment, constatons que, selon la définition de Le Querler, une assertion simple telle que :

(1) Il est venu

est considérée comme modalement neutre, puisqu'elle ne nous dit rien sur l'attitude du locuteur<sup>1</sup>. Par l'ajout d'un *marqueur modal* (par exemple l'adverbe *peut-être* ou le verbe modal *devoir*), le locuteur révèle son attitude à l'égard de ce qu'il dit :

(2) Il est peut-être venu

(3) Il *doit* venir (c'est absolument nécessaire qu'il soit présent)

A part l'information proprement dite, ces énoncés laissent transparaître *la manière* dont le locuteur envisage le contenu de son assertion, et nous avons donc à faire à des énoncés modalisés. Le premier énoncé est un exemple de la modalité épistémique, qui concerne nos connaissances du monde, les convictions et les présomptions du locuteur, et le deuxième énoncé est un exemple de la modalité déontique, qui concerne notamment les notions de permission et d'obligation, mais aussi, dans un sens plus étendu, les notions de souhait et de volonté (voir par exemple Lyons 1994 [1977] : 825). Ces deux types de modalités sont les

---

<sup>1</sup> Une telle position peut pourtant être contestée. On pourrait bien argumenter qu'en présentant quelque chose comme une simple constatation, le locuteur exprime, sans avoir recours à des marqueurs explicites, qu'il tient cette information pour vraie. Il exprime donc son attitude par rapport au contenu informatif, mais plutôt que de le faire par le biais de marqueurs modaux, il le fait par le biais d'une *absence* de marqueurs. Bien qu'on restreigne traditionnellement en linguistique le concept de modalité aux marqueurs *explicites* (voir par exemple Le Querler 1996 : 61), une telle modalisation assertorique ou affirmative est prise en compte par certains chercheurs qui se situent dans une tradition kantienne plutôt qu'aristotélicienne, notamment Kronning (1996 : 37, 2003 : 138) et Gosselin (2000 : 58). Kronning appelle ce type de modalisation la *modalisation simple*. Il convient de noter qu'on peut trouver, chez d'autres, l'appellation *modalisation zero* pour désigner ce même phénomène (voir Le Querler 1996 : 41), terme qui est utilisé dans un sens différent chez Kronning.

seuls à être solidement établis dans la tradition linguistique (voir par exemple Breivega 2003 : 151 sqq ; Lyons 1994 ; Palmer 1986 ; Kronning 2001b : 98).

### 3.1.1.2 Quelques précisions terminologiques

Avant de poursuivre, il faut apporter certaines précisions au sujet des termes employés dans la définition citée ci-dessus. Pour commencer par la fin : le terme *énoncé* n'est pas explicitement défini par Le Querler, mais on s'accorde généralement à envisager l'énoncé comme l'objet linguistique qui résulte d'un acte d'énonciation, c'est-à-dire qu'il est le produit d'un événement unique qu'est l'énonciation (Maingueneau 1993 : 1, 4 ; Benveniste 1966). En cela il s'oppose à la phrase, qui est une entité abstraite qui peut être réutilisée dans plusieurs contextes différents d'énonciation. Depuis Bally (1965), on considère souvent l'énoncé comme se composant de deux éléments : un élément dit objectif, le *dictum*, qui est une représentation (construite par le locuteur) d'un aspect réel, et un élément subjectif, le *modus*, qui exprime la réaction d'un sujet (le plus souvent celle du locuteur) par rapport à cette représentation. C'est sur cette distinction que se basent les définitions linguistiques traditionnelles de la modalité.

Le concept de *contenu propositionnel* semble être plus difficile à définir d'une manière exacte, sans doute parce que ce terme s'utilise souvent sans définition précise et de façon interchangeable avec des termes plus ou moins analogues comme *proposition*, *prédication* et *dictum*. Kronning (1996, 2005, à paraître) argumente en faveur d'une distinction claire entre dictum et contenu propositionnel, où le dictum se compose d'un contenu propositionnel d'un côté et d'opérateurs de temps et d'aspect ainsi que de certains opérateurs de modalité de l'autre. Le dictum constitue la partie *véridicible* de l'énoncé (c'est-à-dire la partie qui peut être discutée en termes de vrai et de faux), qui est à son tour qualifiée par un modus (implicite ou explicite). Cette distinction entre contenu propositionnel et dictum semble correspondre à bien des égards à la distinction que font les linguistes norvégiens Åfarli et Eide (2003 : 179) entre proposition (*'proposisjon'*) et assertion (*'påstand'*), où une proposition se compose d'un sujet et d'un prédicat (dépourvu de valeur temporelle) et une assertion est une proposition équipée d'une temporalité. Le Querler, de son côté, identifie le contenu propositionnel dans l'exemple (1) comme *lui – venir* (Le Querler 1996 : 9) et elle parle souvent de « *la séquence contenant le contenu propositionnel* » (ibid. : 68, nos italiques), ce qui suggère qu'elle aussi fait une distinction similaire. Or, elle semble ailleurs utiliser le concept de contenu propositionnel de façon interchangeable avec les termes de *dictum* (ibid. : 44-45, 81), de *proposition* (ibid. : 14-15, 44), et même parfois d'*énoncé* (ibid. :



61, 125). Un tel usage de ces termes ne témoigne pas nécessairement d'une confusion terminologique, mais peut simplement indiquer qu'une distinction nette entre ces différents termes n'est pas toujours requise ou pertinente. Cependant, afin d'éviter toute confusion, nous adopterons dans la présente étude une distinction semblable : on emploiera le terme de *contenu propositionnel* pour référer au contenu sémantique d'un énoncé qui reste stable indépendamment de changements du temps, de l'aspect, du mode, de la voix etc., c'est-à-dire un réseau sémantique qui indique comment les diverses unités sémantiques sont organisées entre elles<sup>2</sup>, par exemple *lui – venir*. Le terme de *proposition* réfèrera à un tel réseau sémantique équipé d'opérateurs précisant la nature de la relation entre ces entités sémantiques, y compris par exemple les opérateurs temporeux. Le contenu propositionnel de l'énoncé (1) est donc *lui – venir* et la proposition est *il vient*. Une proposition à ceci de caractéristique qu'elle porte toujours une valeur de vérité (dont on peut discuter, bien sûr). Une distinction nette entre ces termes peut être utile quand on parle de la modalité, car bien qu'on puisse dire des opérateurs modaux qu'ils qualifient le contenu propositionnel d'un énoncé, il serait fautif de dire que tout opérateur modal qualifie une proposition : comme nous allons le voir, il y a plusieurs expressions modales qui font partie intégrante de la proposition. Nous parlerons aussi du *contenu informatif*, là où la distinction entre contenu propositionnel et proposition n'est pas pertinente.

Le concept de *locuteur* renvoie à celui qui énonce. Cependant, celui qui énonce n'équivaut pas toujours à celui qui modalise : il existe bien entendu des cas de modalisation rapportée où le locuteur n'est pas la source de la modalisation, comme dans *Jaques aimerait que tu viennes* (Le Querler 1996 : 63). Dans cet énoncé, ce n'est pas forcément l'attitude du locuteur qui est exprimée, mais l'attitude de Jaques (ou, pour être plus précis, l'attitude que le locuteur attribue à Jaques (voir Picavez 2003 : 43-50)). Ce type d'énoncé rentre donc difficilement dans une définition traditionnelle de la modalité<sup>3</sup>. De tels énoncés sont toutefois couramment considérés comme modaux (voir Picavez 2003 : 49-50). Pour distinguer clairement entre les deux entités, nous appellerons, avec Picavez (ibid. : 43-50) celui qui énonce *le locuteur* et la source de la modalisation soit *le support modal* (terme que Picavez emprunte à Meunier 1990) soit, avec Kronning (2003 : 137) *l'instance modalisante*. Il ressort du propos de cette étude que nous nous intéressons surtout aux cas où le locuteur coïncide

---

<sup>2</sup> Cette définition s'inspire d'une définition donnée à la page web [www.fltr.ucl.ac.be/FLTR/GERM/GramConc.htm](http://www.fltr.ucl.ac.be/FLTR/GERM/GramConc.htm), visitée le 13.03.06.

<sup>3</sup> Il faut souligner que la définition proposée par Le Querler et dont nous nous servons comme point de départ n'est qu'une parmi de nombreuses caractérisations de la modalité en linguistique où le locuteur est l'entité centrale.

avec l'instance modalisante, ce qui implique que dans la majorité des cas, le choix de l'un ou l'autre terme importe peu. Or, pour anticiper un peu l'analyse du corpus, il existe des cas qui peuvent être perçus soit comme une modalisation rapportée, soit comme un rapport modalisé où le locuteur adjoint un marqueur modal à une assertion qui relève de quelqu'un d'autre. Ces cas seront discutés dans la section 6.2.3.

Il faut maintenant expliciter le terme *attitude*, qui semble être le terme le plus souvent employé pour décrire le rapport entre le locuteur et l'information communiquée. Certains parlent de *commentaire* (Vion 2001) ou de *prise en charge* (Tutescu 1997 : 387) à la place. Il est possible d'envisager au moins trois types d'attitude : une attitude envers la valeur de vérité de l'information transmise (modalité épistémique), une attitude envers l'exécution du procès dénoté (modalité déontique) et une attitude affective, c'est-à-dire une appréciation sur l'information assertée (modalité axiologique ou évaluative, voir la section 3.1.1.4 ci-dessous). Il existe pourtant d'autres types de modalité qui ne se laissent pas décrire si aisément en termes d'*attitude* ou de *commentaire* (cf. section 3.1.1.4).

Finissons par un bref commentaire sur le terme *expression*. Le locuteur peut naturellement exprimer son attitude par toute une série de moyens : en principe, un regard ou un geste peuvent exprimer les trois types d'attitude. Ici, il s'agira évidemment de l'expression *linguistique* d'une attitude, c'est-à-dire un mot, un groupe de mots, un affixe grammatical etc. qui véhicule une valeur modale, laquelle, à son tour, représente une attitude chez le locuteur ou, plus rarement, chez un autre être discursif<sup>4</sup>.

Ayant ainsi précisé les termes utilisés dans notre définition de départ, il nous reste à distinguer les termes *modalité* et *modalisation*. Ces deux termes peuvent dans beaucoup de contextes s'utiliser comme des quasi-synonymes. On peut pourtant dire, avec Picavez (2003 : 32-36), que la modalisation désigne le processus alors que la modalité renvoie à une réalité statique (ibid. : 33), c'est-à-dire que la modalisation renvoie au processus par lequel le locuteur effectue une qualification modale, et la modalité renvoie au résultat de ce processus ou bien au type de qualification qu'on veut exprimer ; on peut donc dire que pour exprimer une *modalité*, on doit *modaliser* à l'aide de *marqueurs modaux* ou de *modalisateurs*. Cette distinction terminologique peut sembler triviale, mais elle s'impose du fait que le sens accordé à ces termes n'est pas constant. Vion (2004) par exemple, utilise ces termes dans un sens nettement différent : pour lui, la modalité « contribue directement au sémantisme des énoncés auxquels elle participe » (ibid. : 103), alors que la modalisation ne contribue pas

---

<sup>4</sup> Les êtres discursifs sont des « personnages » présentés dans le discours (Nølke et al. 2004 : 37).

directement à ce sémantisme, mais correspond à des commentaires réflexifs portant sur ces énoncés. Cette distinction semble correspondre à celle que Kronning (1996, 2003) fait entre modalités véridicibles et modalités montrables (voir *infra* 3.1.2.1).

Les difficultés liées à la définition et la délimitation de la catégorie de la modalité relèvent d'au moins deux ordres. Premièrement, la modalité est une catégorie sémantique, voire sémantico-pragmatique, et ces catégories tendent à être moins faciles à définir que les catégories morphosyntaxiques (voir Andersen 2003 : 25). Cela s'explique par le caractère abstrait de ces niveaux linguistiques : le sens et les fonctions ne sont pas aussi visibles (ou auditibles) que les éléments phonologiques, morphologiques ou syntaxiques. Bien que la catégorie de la modalité soit définie sémantiquement, elle a, évidemment, des liens avec les autres niveaux linguistiques, en ce sens que divers éléments de différents niveaux linguistiques peuvent exprimer la modalité.

Deuxièmement, différentes disciplines et orientations linguistiques ont depuis longtemps des conceptions divergentes, pourtant reliées, de ce qu'est la modalité. D'un côté, on a la logique modale et diverses conceptions linguistiques d'inspiration philosophico-logique, tournant autour de termes comme *mondes possibles*, *possibilité* et *nécessité* (voir Lyons 1994 : 787). De l'autre côté, on a une série de conceptions linguistiques où la *subjectivité* et l'*attitude / prise de position du locuteur* ont été les concepts centraux dans la définition de la modalité (voir Palmer 1986 : 14-17 et Ducrot 1993). Avant de présenter ces différentes conceptions de façon plus détaillée, il convient de clarifier les liens entre la catégorie sémantique de la modalité et les autres niveaux linguistiques.

### 3.1.1.3 La modalité : une catégorie sémantique

Comme il a déjà été mentionné, nous considérerons la modalité en premier lieu comme une *catégorie sémantique*, qui a toutefois des liens évidents avec la pragmatique, les deux catégories se recouvrant partiellement. Si la notion de modalité elle-même est considérée comme relevant du domaine de la sémantique, les différents marqueurs modaux peuvent être utilisés par le locuteur dans le but d'influencer le destinataire de diverses manières (on s'approche alors du domaine de la pragmatique). L'aspect pragmatique de la modalité épistémique, c'est-à-dire le rôle qu'elle joue dans le discours, sera traité au chapitre 7.

Le fait de définir la modalité comme une catégorie sémantico-pragmatique n'est pourtant pas évident. Palmer (1986 : 7) s'efforce d'établir la modalité comme une catégorie grammaticale, pour éviter le flou qui, selon lui, caractérise les définitions sémantiques. Il

définit la modalité comme « the grammaticalization of speakers' (subjective) attitudes and opinions » (ibid. : 16), c'est-à-dire qu'il étudie uniquement la modalité telle qu'elle s'exprime par la morphologie (par flexions, affixes, modes, etc.) et par les verbes modaux. Cette position le conduit à exclure, du moins en principe, les éléments lexicaux. Il avoue pourtant que « les unités lexicales ne peuvent pas être complètement négligées » (Palmer 1986 : 6, notre traduction), ce qui suggère qu'il y a, au fond, une définition sémantique.

Une position à l'instar de celle de Palmer n'est pas appropriée pour la présente étude, dont l'un des objectifs est de comparer en quoi les chercheurs de différentes cultures académiques modalisent leurs énoncés ; plus précisément il s'agit de voir dans quelle mesure ils expriment des incertitudes par rapport à l'information transmise. Une étude contrastive portant exclusivement sur la modalité grammaticalisée ne pourrait rien dire sur cette question, car les langues varient bien entendu dans le degré de grammaticalisation des expressions<sup>5</sup>. Si une langue n'a que peu ou pas de modalité grammaticale, cela ne veut sans doute pas dire que cette langue manque de moyens pour exprimer les incertitudes, les permissions, les obligations etc., mais simplement que l'expression de ces phénomènes se produit sous d'autres formes. Nous préférons donc considérer la modalité comme une catégorie sémantique, tout en reconnaissant, bien entendu, que les mécanismes linguistiques à l'œuvre peuvent être de caractère morphosyntaxique, voire phonologique aussi bien que de caractère lexical. Nous reviendrons sous peu aux différents moyens linguistiques permettant d'exprimer la modalité (section 3.3).

Pour éviter toute confusion de terminologie, il convient de distinguer la modalité de deux catégories liées, mais distinctes : le *mode*, d'un côté, et les *actes de langages* de l'autre.

### *Mode et modalité*

Le *mode*, contrairement à la modalité, est une catégorie grammaticale. Plus précisément, par *mode* nous entendons ici une catégorie morphosyntaxique exprimée par la flexion des verbes. On tient compte de trois modes au moins en français : l'indicatif, le subjonctif et l'impératif. Il faut citer en outre le conditionnel, qui est tantôt considéré comme un mode, tantôt comme un temps verbal de l'indicatif.

L'indicatif est vu comme modalement *neutre*. L'impératif exprime la modalité déontique, alors que le subjonctif est lié à la modalité épistémique, en ce sens qu'il présente souvent – mais pas toujours – des choses *virtuelles*. Pour cette raison, les modes sont

---

<sup>5</sup> Pour une présentation plus détaillée de ce que nous entendons par *grammaticalisation* et *moyens grammaticaux*, voir la section 3.3.

traditionnellement assimilés aux modalités, ce qui est une simplification excessive (voir Riegel, Pellat et Rioul 1998 [1994] : 287-288). Il n'y a pas de relation univoque entre les différents modes et modalités : comme Palmer le remarque (1986 : 22), « *moods [...] have a whole variety of semantic functions, and [...] the choice between them is determined grammatically more than by modal meanings.* » Ainsi, les subjunctifs dans les exemples 4 et 5 ci-dessous n'expriment rien de « possible » ou de « virtuel », mais des faits irrévocables, et ils ne sauraient être interprétés comme des modalisateurs épistémiques.

(4) Je regrette qu'il soit venu

(5) Bien qu'il soit malade,...

Le mode de l'impératif, en revanche, semble toujours véhiculer une modalité déontique, en ce qu'il a une valeur fondamentalement *directive* (Riegel et al. 1998 : 331). Qu'il exprime un ordre (6), une prière (7) ou une invitation/suggestion (8), on peut le rattacher au domaine du déontique :

(6) Sortez !

(7) Ne me quitte pas

(8) Alors, viens !

Le terme *mode* renvoie donc à une catégorie morphosyntaxique qui s'applique au verbe. Le mode est l'un des multiples moyens linguistiques par lequel on peut exprimer la modalité. Il peut être le seul marqueur modal d'un énoncé – comme dans l'énoncé *viens ici!*, avec comme seul marqueur modal l'impératif, – est il peut se combiner avec d'autres marqueurs modaux, comme dans l'exemple *il est possible qu'il vienne*, où le subjunctif renforce ou reprend la valeur modale du modalisateur *il est possible que*. Le choix entre le mode du subjunctif et le mode de l'indicatif est essentiellement gouverné par des règles grammaticales, mais le rôle de l'aspect sémantique n'est pas négligeable. Ces facteurs interagissent et se renforcent. Par exemple, c'est l'aspect sémantique qui explique pourquoi on aura *il est possible qu'il vienne* mais *il est probable qu'il viendra*, c'est-à-dire deux modes différents dans des contextes syntaxiques identiques.

---

<sup>6</sup> Exemples empruntés à Gosselin 2000 : 75.

Alors que la relation entre le mode de l'impératif et la modalité déontique semble être très forte, la relation entre le subjonctif et la modalité épistémique est, comme nous l'avons vu, beaucoup plus faible.

### *Actes de langage et modalité*

Les actes de langage tels que l'affirmation, l'ordre et l'interrogation sont parfois appelés *modalités d'énonciation* (Riegel et al. 1998 : 385-386, 580) ou *modalités d'assertion* (Pottier 1985 : 169, cf. Le Querler 1996 : 53). Cet usage n'est pourtant pas inclus dans notre approche de la modalité, laquelle est, pour nous, une notion sémantique, alors que l'étude des actes de langage ou des forces illocutoires relève nettement du champ de la pragmatique. Or, comme mentionné plus haut, ces deux niveaux sont étroitement liés, en ce sens que les énoncés modalisés sont utilisés pour effectuer divers actes de langage (Herslund 1989 : 13-14). Par exemple, il y a une connexion évidente entre la modalité déontique et des actes directifs tels que les obligations, les permissions et les suggestions. Pareillement, mais de manière moins évidente, il y a un lien entre la modalité épistémique et les interrogations. Lyons dit, à propos de la modalité épistémique : « [...]subjectively modalized utterances[...] are not acts of telling ; and [...] their illocutionary force is in this respect similar to that of questions [...] » (Lyons 1994 : 799). Une affirmation atténuée par un modalisateur épistémique peut aisément être perçue comme une question. Il est néanmoins important de maintenir la distinction entre ces deux catégories, car il n'y a pas de relation biunivoque entre les modalités et les actes de langage. Ainsi, l'*interrogation* est un acte de langage qui peut se rattacher à la modalité épistémique, comme dans (9), où le locuteur s'interroge sur la valeur de vérité de la proposition *Pierre vient* :

(9) Est-ce que Pierre vient ?

Or, il est moins facile de lier les questions ouvertes à la modalité épistémique, étant donné que celles-ci peuvent difficilement être conçues comme exprimant une attitude envers la valeur de vérité d'une proposition :

(10) Où vas-tu demain?

Si le terme de modalités d'énonciation est parfois utilisé pour désigner les actes de langage, il renvoie plus fréquemment à des qualifications qui portent sur « la manière dont [le locuteur]

énonce son propos » (Le Querler 1996 : 70), c'est-à-dire sur l'énonciation ou l'acte illocutoire lui-même plutôt que sur une proposition (voir Herslund 1989 : 14, Nølke 1993). L'énoncé (11) est un bon exemple : l'adverbe *franchement* ne qualifie pas forcément la proposition *ce roman est excellent*, mais plutôt la manière dont le locuteur pose cette proposition :

(11) Franchement, ce roman est excellent<sup>7</sup>

Dans la tradition linguistique française, on subsume souvent sous le terme de modalité ce type de qualifications. Ainsi, situé dans la tradition linguistique française, Nølke est un des linguistes qui étendent le terme *modalité* jusqu'à inclure les modalisations des actes de langage, tout en distinguant clairement entre les deux types, appelés les modalités d'énonciation et les modalités d'énoncé (Nølke 1993 : 85, 143). Herslund, en revanche, désire restreindre le terme de modalité à « ce qui a lieu dans la phrase » (Herslund 1989 : 14-15, notre traduction), c'est-à-dire aux modalités de l'énoncé. Aussi hésite-il à inclure dans la catégorie de la modalité des exemples tels que (11). Le Querler, de son côté, maintient qu'un adverbe tel que *franchement* dans l'exemple cité ci-dessus porte sur l'énoncé aussi bien que sur l'énonciation (Le Querler 1996 : 70), et l'inclusion dans la catégorie de la modalité s'impose donc de façon évidente. On note ainsi un désaccord par rapport au statut modal de ce type d'exemples, mais dans la présente étude, ce sont ce que Nølke appelle les modalités d'énoncé qui sont pertinentes, et nous n'entrerons pas plus avant dans la discussion sur le statut modal d'énoncés tels que (11).

En résumé, on pourrait conclure que la catégorie sémantique de la modalité se situe entre le morphosyntaxique d'un côté (mode) et la pragmatique de l'autre (actes de langage/forces illocutoires). Dans cette section, nous avons essayé de montrer les liens entre ces catégories tout en les maintenant séparées.

#### 3.1.1.4 Conceptions divergentes de la modalité

Il est maintenant temps de regarder de plus près les diverses conceptions de la modalité. Nous aborderons d'abord la catégorie de la modalité telle qu'elle a été conçue dans la tradition philosophico-logique, pour ensuite continuer avec les conceptions plus spécifiquement linguistiques.

---

<sup>7</sup> Exemple emprunté à Nølke (1989 : 51).

### *Les conceptions philosophico-logiques*

Dans la tradition philosophico-logique, la modalité est conçue comme ce qui nous permet de parler de mondes possibles. La notion de modalité remonte à Aristote et la logique classique, où les analyses tournaient autour du fameux quaterne de la possibilité, la nécessité, la contingence et l'impossibilité (voir Douay 2003). Le terme et les analyses ont continué d'évoluer, et en 1951, von Wright publie *An essay in modal logic*, qui est devenu un classique dans le domaine. Depuis von Wright, on distingue en philosophie quatre (ou cinq) types de modalités (von Wright 1951 : 1-2) : les modalités aléthiques, qui sont les modes de la vérité (*It is possible that Jones is dead*)<sup>8</sup>, les modalités épistémiques, qui sont les modes de la connaissance (*Jones is known to be dead*), les modalités déontiques, qui sont les modes de l'obligation (*it is permitted/obligatory/forbidden to...*) et les modalités dynamiques, qui sont les modes de la capacité (*Jones can speak German*). La cinquième catégorie comporte ce que von Wright appelle les modalités existentielles ou les modes d'existence (*some/all/no things possess this property*). Selon von Wright, ce dernier type n'est pas forcément modal – il n'est en général pas considéré comme une branche de la logique modale – mais il a des ressemblances essentielles avec les autres types de modalité. Les différents types de modalité seront plus amplement présentés vers la fin de ce chapitre.

Les concepts philosophico-logiques de la modalité (représentés ici par l'acception de von Wright) sont étroitement liés aux deux opérateurs modaux : la possibilité et la nécessité. Les marqueurs prototypiques de la modalité sont les verbes modaux *pouvoir* (possibilité) et *devoir* (nécessité). Or, il existe plusieurs *types* de possibilité et de nécessité, d'où le classement en modalités aléthiques, déontiques, épistémiques et dynamiques<sup>9</sup>.

On oppose traditionnellement les modalités *de re* et les modalités *de dicto* ; ces premières portent sur la relation entre le sujet et le prédicat et s'inscrivent dans la proposition même, alors que les secondes portent sur la proposition dans sa globalité. Ainsi, le *devoir* déontique dans l'exemple 3 (*il doit venir*) porte sur la relation entre *il* et *venir*, alors que le modalisateur *peut-être* dans l'exemple 2 (*il est peut-être venu*) porte sur l'ensemble de la proposition *il est venu*.

La conception philosophico-logique diverge de la conception traditionnelle en linguistique (citée au début de ce chapitre) en ce qu'elle n'implique pas nécessairement que la modalisation exprime une *attitude* ou un *commentaire* de la part du locuteur par rapport à

---

<sup>8</sup> Les exemples et les gloses sont tous empruntés à von Wright (1951).

<sup>9</sup> D'autres logiciens ont tenu compte d'autres catégories : Rescher (1968) par exemple, parle de « temporal modalities », « bouloimaic modalities », « evaluative modalities » and « causal modalities », outre les modalités aléthiques, épistémiques et déontiques (voir Palmer 1986 : 12-13).



l'information transmise. Traditionnellement, la logique modale ne s'intéresse pas beaucoup à la question du sujet énonciateur. Parmi les modalités recensées dans la tradition philosophico-logique sont donc inclus des types qui n'expriment pas forcément une attitude de la part du locuteur à l'égard du contenu informatif de son énoncé. Par exemple, la modalité aléthique se distingue de la modalité épistémique, selon le philosophe Kolflaath (entretien personnel, 2003), en ce qu'elle s'inscrit dans le contenu propositionnel même, alors que la modalité épistémique exprime un jugement qui porte sur un contenu propositionnel<sup>10</sup>. La modalité aléthique n'exprime donc pas une attitude envers l'information communiquée, mais elle en fait partie et elle correspond plutôt à une simple affirmation d'une possibilité ou d'une nécessité (cf. *infra* section 3.1.2.2). Il en va de même pour la modalité dynamique, qui prend plutôt la forme d'une simple constatation que d'un jugement du locuteur. Si on dit *les dauphins peuvent nager*, on n'exprime guère son attitude personnelle (cf. Boye 2002 : 7). On pourrait peut-être avancer que par la modalité dynamique, le locuteur exprime un jugement sur, ou une appréciation des qualités du référent (Le Querler 1996 : 134), mais il est clair que le lien entre attitude et modalisation est beaucoup moins évident dans ce cas-là qu'en ce qui concerne la modalité épistémique et dans une certaine mesure la modalité déontique. Ces dernières s'inscrivent plus facilement dans une conception traditionnelle de la modalité en linguistique dans la mesure où il est plus facile de les relier à l'expression d'une attitude, un jugement ou un commentaire de la part du locuteur. La modalité épistémique exprime l'attitude du locuteur à propos d'une proposition alors que la modalité déontique exprime l'attitude du locuteur à propos d'un événement (d'où l'opposition entre « propositional modality » et « event modality » dans Palmer 2001)<sup>11</sup>

### *Les conceptions linguistiques*

En linguistique, les conceptions varient beaucoup. Il est commun de distinguer une conception large et une conception restreinte (voir par exemple Gardies 1983, Herslund 1989, Le Querler 1996 et Douay 2003) ; cette dernière s'inspire de la conception philosophico-logique, alors que la première s'allie plutôt aux théories linguistiques de l'énonciation et cherche à établir une définition de la notion de modalité qui soit spécifique à la discipline de

---

<sup>10</sup> Il nous semble que Kolflaath utilise ici le concept de contenu propositionnel comme équivalent de notre *proposition*.

<sup>11</sup> Comme Boye (2002 : 7-8) l'a montré, il est parfois difficile de relier la modalité déontique à l'attitude du locuteur. Un énoncé comme *Les élèves doivent se présenter en classe à l'heure* ne révèle pas nécessairement l'attitude du locuteur (il n'exprime pas nécessairement un ordre de la part du locuteur), il peut être conçu comme un simple rapport, une simple description d'un état de choses. Cependant, avec un sujet à la deuxième personne, l'énoncé se serait lu plus aisément comme l'expression de l'attitude du locuteur.

la linguistique. Herslund (1989) rattache ces deux points de vue divergents à deux traditions linguistiques différentes : il relie la conception large à la tradition française et la conception restreinte à la tradition anglo-saxonne. Il est peut-être vrai que la conception philosophico-logique de la modalité a eu plus d'influence parmi les linguistes qui se situent dans une tradition anglo-saxonne (considérer par exemples les ouvrages de Papafragou 2000, Perkins 1983 et Lyons 1994) que parmi ceux situés dans une tradition française. Dans la tradition française, il semble y avoir un plus grand scepticisme sur l'applicabilité de la conception philosophico-logique, et le choix de prendre les catégories de la logique modale comme point de départ a été fortement critiqué. Par exemple, Vion (2001 : 216) écrit à cet égard que « la théorie linguistique ne peut être pensée comme l'extension d'une théorisation produite en dehors d'elle », et que « tout emprunt conceptuel doit faire l'objet d'une véritable intégration [...] », ce qui n'est pas le cas, selon lui, dans l'étude des modalités à partir des catégories logiques. Néanmoins, il faut prendre garde à ne pas trop généraliser : beaucoup de linguistes situés dans une tradition française ont été inspirés par une conception logique de la modalité (voir par exemple David et Kleiber 1982 et Martin 1987, dont le concept d'*univers de croyance* complète celui de *mondes possibles* des logiciens), et il y a également des linguistes situés dans une tradition anglo-saxonne qui définissent la notion de modalité à la manière « française » (par exemple Stubbs 1986). Peut-être s'agit-il autant des domaines de la linguistique auxquels on s'intéresse (par exemple la sémantique formelle vs la pragmatique et l'énonciation) que de la tradition scientifique nationale à laquelle on appartient.

Selon une conception large, la notion de modalité regroupe toute trace énonciative, c'est-à-dire qu'elle embrasse « tous les phénomènes linguistiques qui signalent la présence de l'homme dans le langage » (Herslund 1989 : 7, notre traduction). Dans cette optique, la notion de modalité est assimilée à celle de subjectivité (Douay 2003). Cette conception, faisant de la modalité un terme très étendu nommant une catégorie vaste et hétérogène qui se compose potentiellement, entre autre, de déictiques (le *moi-ici-maintenant*), d'adverbes de jugement, d'expressions évaluatives, etc., reste, selon Herslund, la conception dominante dans la tradition linguistique française. Une telle conception permettra de traiter sous le terme de modalité des phénomènes aussi divers que *la temporalité, l'aspect et la négation*<sup>12</sup>. Étant donné que la modalité correspond à « la présence de l'homme dans le langage », les *actes de langage* (questions, demandes etc.) peuvent aussi être abordés sous le terme de modalité

---

<sup>12</sup> Du moins d'un point de vue polyphonique : les assertions négatives ne sont pas considérées comme de simples assertions, mais comme des assertions « doubles » où l'énoncé négatif est un rejet (une qualification) de l'énoncé positif correspondant.

(Herslund 1989). De surcroît, une conception large permet de considérer les différents *types de phrases* (l'interrogation, l'exclamation, voire l'affirmation simple) comme diverses *modalités* (Le Querler 1996 : 51-52), ce qui implique que tout énoncé est modalisé.

Il est vrai que plusieurs des phénomènes mentionnés sont étroitement liés à la modalité telle qu'elle a été définie au début de ce chapitre : par exemple, plusieurs temps grammaticaux se prêtent, dans certains contextes, plus facilement à une lecture modale qu'à une lecture temporelle (par exemple le conditionnel<sup>13</sup> et le futur). Il y a beaucoup à dire sur les relations entre ces phénomènes et la modalité, mais nous ne développerons pas ces idées ici. Constatons simplement que la conception large de la modalité permet d'y inclure d'innombrables phénomènes langagiers. Le recueil de Vogeleer et al. (1999) pertinemment intitulé *La modalité sous tous ses aspects* illustre bien une telle conception. Dans ce recueil sont réunis des articles portant sur des phénomènes linguistiques aussi diversifiés que les modes, les temps (valeurs modales et valeurs temporelles), le gérondif, le mode d'action et l'aspect, ainsi que certains adjectifs et groupes nominaux.

Une conception large de la modalité a le mérite de prendre en compte les relations qui existent entre différents phénomènes linguistiques et entre différents niveaux linguistiques. Mais une telle approche présente le grand inconvénient de mener justement à une « nébuleuse » (cf. section 3.1.1.1) : il est impossible d'y voir clair, il est difficile de trouver le fil conducteur entre les divers phénomènes, et au fond, la notion perd du sens par faute de contenu précis.

C'est la raison pour laquelle d'autres défendent une acception plus étroite du concept de la modalité. Herslund observe que la modalité, dans ce sens, « semble se restreindre à des phénomènes tels que l'expression de *la possibilité, la nécessité, l'obligation et la permission* (les modalités épistémiques et déontiques)» (Herslund 1989 : 8, notre traduction). Dans cette optique, la *modalité épistémique* concerne les notions de possibilité et de nécessité, par exemple :

(12) Il peut être là (je ne sais pas).

(13) Paul n'est pas là. Il doit être parti.

et la *modalité déontique* concerne les notions de permission et d'obligation, par exemple :

---

<sup>13</sup> Le fait que le conditionnel est souvent considéré comme un *mode* (voir par ex. Le Nouveau Petit Robert 1995 : 1420, Vogeleer et al. 1999 : 2 et, pour une analyse critique de cette pratique, Riegel et al. 1998 : 287) montre à quel point les temps verbaux et la modalité sont liés.

(14) Tu peux y aller (je t’y autorise)

(15) Tu dois partir (je te l’ordonne)

Selon la conception restreinte, le terme de modalité est donc réservé aux expressions épistémiques et déontiques. Ces deux phénomènes, qui à première vue peuvent sembler assez différents l’un de l’autre, sont en fait étroitement liés et parfois même à peine distinguables (voir Lyons 1994 : 846 ; Palmer 1986 : 121-126 ; Herslund 1989 : 12 ; Le Querler 1996 : 114). Il est bien connu que les auxiliaires modaux *pouvoir* et *devoir* se prêtent tantôt à une lecture déontique, tantôt à une lecture épistémique, en fonction du cotexte (cf. les exemples ci-dessus). Ce fait nous signale qu’il y a des liaisons étroites entre le déontique et l’épistémique, car cette ambiguïté des verbes modaux ne peut pas, à notre avis, être mise sur le compte d’une homonymie<sup>14</sup>. Ce sont les termes *nécessité* et *possibilité* qui constituent le pivot et le dénominateur commun : ce qui est permis est possible (possibilité déontique) et ce qui est obligatoire est nécessaire (nécessité déontique) (voir Faarlund, Lie et Vannebo 2002 [1997] : 584). Selon la conception étroite, on aura donc deux types de possibilité (déontique et épistémique) et deux types de nécessité (déontique et épistémique) qui ensemble et seuls constitueront le domaine de la modalité.

La conception restreinte telle qu’elle est présentée par Herslund a l’avantage d’être claire et facilement abordable, mais elle se heurte inévitablement à certaines difficultés. Qu’en est-il notamment de la *possibilité* qui n’est ni épistémique, ni déontique ? Les notions de possibilité et de nécessité recouvrent plus que l’épistémique et le déontique ; la possibilité inclut par exemple la *capacité* (*Je peux soulever cette valise*), l’*occasion* (*Il a la possibilité de venir*) ainsi que certains d’autres types de *possibilité* plus difficiles à étiqueter, mais qu’on pourrait peut-être qualifier de *possibilité radicale sous-déterminée* (voir *infra* 3.1.2.2), comme dans *il est possible d’aller plus loin*. Inversement, de nombreux chercheurs maintiennent que la notion du déontique embrasse plus que la possibilité et la nécessité, par exemple les souhaits et la volonté (Lyons 1994 : 825 ; Faarlund et al. 2002 : 580). Une conception très restreinte présenterait donc l’inconvénient d’exclure plusieurs notions sémantiques qu’il pourrait être utile de prendre en compte. C’est l’admission de ce fait qui constitue, semble-t-il, la raison pour laquelle on trouve des caractérisations légèrement contradictoires de la conception restreinte, telles que la suivante: « la *conception minimaliste* consiste à dire qu[e la

---

<sup>14</sup> Certains ont en effet mis cette plurivocité des auxiliaires modaux sur le compte de l’homonymie, par exemple Huot 1974 (voir Kronning 1996 :15).

modalité] regroupe *au moins* deux catégories : le nécessaire et le possible » (Dendale et Auwera (2001 : i), les italiques sont les nôtres).

Il convient de noter que la conception large et la conception étroite ne constituent pas deux alternatives entre lesquelles le linguiste doit trancher ; elles constituent plutôt deux pôles d'une échelle et la plupart des travaux linguistiques portant sur la modalité témoignent d'une conception médiane, c'est-à-dire d'une conception qui se situe quelque part entre les deux extrêmes évoqués par Herslund (1989). Une définition courante de la modalité est celle qui s'appuie sur des concepts tels que *attitude*, *jugement* et *locuteur*. La modalité est donc toujours liée à la présence de l'homme dans le langage, mais il ne suffit pas que le locuteur se manifeste simplement dans son discours, il doit aussi exprimer une attitude ou un jugement, une *évaluation*. Ces définitions s'inspirent néanmoins de l'ancienne dichotomie entre le *modus* et le *dictum* (Bally 1965), selon laquelle un énoncé se compose de deux éléments : un contenu « objectif » d'un côté (le *dictum*) et un jugement subjectif relatif à ce contenu de l'autre (le *modus*)<sup>15</sup>. C'est sur une telle dichotomie que se base Le Querler (1996 : 14) quand elle définit la modalité comme « l'expression de l'attitude du locuteur par rapport au contenu propositionnel de son énoncé ».

Breivega (2003) présente une autre conception intermédiaire, qui néanmoins se base sur les mêmes concepts. Pour elle, la modalité concerne – outre les notions de possibilité et de nécessité – *les attitudes et les opinions* exprimées par le locuteur par rapport à ce qu'il dit. Cela permet à Breivega d'inclure plus que l'épistémique et le déontique : la définition ouvre sur une troisième catégorie qu'elle nomme la *modalité évaluative*. Il s'agit d'évaluations exprimées par le locuteur, de jugements subjectifs qui situent le contenu propositionnel sur un axe *bon-mauvais*, parallèle à l'axe *certain – incertain* de l'épistémique (voir ci-dessous). La modalité évaluative recouvre des expressions telles que : *il est intéressant / plaisant / important...* (Breivega 2003 : 157-158).

La définition de Breivega est d'une ressemblance frappante avec celle de Le Querler. Cependant, les deux positions ne sont pas identiques. Là où Breivega souligne que l'évaluatif n'est pas une catégorie solidement établie dans la tradition, Le Querler l'inclut sans discussion (sous le terme de *modalités appréciatives*). Le Querler inclut en outre d'autres catégories, par exemple ce qu'elle appelle les *modalités implicatives*, qui concernent des phénomènes tels que la concession, la comparaison et la condition. Ce faisant, elle s'approche d'une

---

<sup>15</sup> Beaucoup de définitions s'inspirent d'une telle dichotomie sans pour autant que les termes *modus* et *dictum* ne soient nécessairement pris dans un sens exactement identique à celui que Bally leur a accordés. C'est l'idée qu'un énoncé se compose de deux « composantes » différentes, l'une dite objective et l'autre subjective, qui est récurrente.

conception large, car on s'éloigne ici nettement des termes de possibilité et de nécessité, et il n'est pas facile non plus de voir immédiatement de quels types d'attitude il est question dans ces cas. Cette différence entre deux conceptions médianes n'est pas très surprenante dans la mesure où Breivega se situe dans une tradition linguistique influée par la tradition anglo-saxonne alors que Le Querler se situe dans une tradition française où le concept de modalité a une extension plus large. Un trait commun, par contre, est l'exclusion de traces énonciatives telles que les déictiques, traces qui n'impliquent pas de jugement ou de qualification de la part du locuteur, mais qui signalent tout simplement la présence de celui-ci. Selon les conceptions médianes, la modalité concerne *une manière de concevoir* et les traces qui tout simplement signalent la présence du locuteur ne suffisent pas à être étiquetées *modales* (Breivega 2003 : 151)<sup>16</sup>.

Les conceptions qui lient la notion de modalité à l'expression de l'attitude du locuteur sont résolument les plus fréquentes. Alors que certains linguistes se contentent de signaler qu'il y a une relation entre attitudes et expressions modales, d'autres tentent de préciser en quoi cette relation consiste, avec des approches plus ou moins novatrices. Nølke (1989) propose d'appliquer une perspective polyphonique à l'analyse des expressions modales, ce qui marche très bien pour les marqueurs épistémico-modaux, dont on peut dire qu'ils impliquent toujours, de par leur bilatéralité (voir *infra* 3.1.2.1), deux points de vue différents (soit la proposition qualifiée est vraie, soit elle est fausse). Ainsi, dans un exemple comme *Pierre est peut-être bête, mais il est riche*<sup>17</sup>, le locuteur met en scène au moins deux êtres discursifs ; le premier est responsable de l'assertion *Pierre est bête*, le second met en question cette assertion, à l'aide du modalisateur *peut-être*. Le locuteur s'associe au second être discursif, mais pas nécessairement au premier ; il peut fort bien maintenir que ce n'était pas lui qui disait que Pierre était bête. Vion (2001, 2004) propose une perspective similaire, selon laquelle les expressions modalisantes expriment un commentaire réflexif de la part du locuteur sur l'information transmise. Dans cette optique, la modalisation<sup>18</sup> entraîne un *dédoublé énonciatif* et « contribue à construire l'image d'un sujet dédoublé » (Vion

---

<sup>16</sup> Il convient de noter que les déictiques sont, à notre connaissance, rarement traités sous le terme de modalité, bien que la première définition évoquée (mais pas soutenue) par Herslund (1989) argue en faveur de l'inclusion de ces traces énonciatives. Les partisans d'une conception intermédiaire semblent pourtant utiliser la notion de modalité dans un sens plus déterminé que ne le font les adhérents d'une conception large ; elle acquiert un sens technique qui s'accorde plus ou moins avec le sens du terme '*modality*' en anglais, c'est-à-dire un sens beaucoup plus restreint que le sens multiple que ce terme possède en français, où il fonctionne comme un mot ordinaire (voir Herslund 1989 : 9).

<sup>17</sup> Emprunté à Nølke 1989 : 60.

<sup>18</sup> Rappelons que Vion (2001, 2004) oppose modalisation et modalités ; une distinction qui semble correspondre à l'opposition entre modalités montrables et modalités véridicibles (Kronning 1996, 2003, cf. *infra* 3.1.2.1) et à la distinction que fait Nølke (1989) entre « non-asserted and asserted modalities ».

2004 : 103). Bien que ces perspectives soient aptes à décrire et expliquer la modalité épistémique, leur applicabilité à d'autres types de modalité est plus contestable. Il est peut-être symptomatique de ce fait que les exemples donnés dans Nølke 1989 et Vion 2001, 2004 soient presque exclusivement du type épistémique<sup>19</sup>.

Pour résumer, on peut dire qu'en linguistique, la modalité est très souvent conçue comme quelque chose qui qualifie ou commente un contenu informatif, exprimant l'attitude ou l'avis subjectif du locuteur<sup>20</sup>. Cela vaut non seulement pour les conceptions larges ou intermédiaires, mais aussi pour les conceptions plus restreintes, dans la mesure où les partisans de celles-ci tendent à restreindre, selon Herslund (1989 : 8), la modalité aux expressions épistémiques et déontiques, autrement dit aux types de modalité qui sont censés exprimer un jugement ou une opinion de la part du locuteur. La modalité aléthique et la modalité dynamique, qui n'expriment pas forcément l'attitude du locuteur, sont exclues (voir Faarlund et al. 2002 : 581). Quelle que soit l'acception du terme, nous avons donc vu que la définition de la modalité en linguistique tend à se baser sur le concept de subjectivité (attitude, jugement du locuteur) et sur la dichotomie traditionnelle *modus – dictum* (sous différentes variantes). Cette optique a été amplement critiquée. C'est surtout la notion de subjectivité sur laquelle se base (voire à laquelle s'assimile) la conception linguistique de la modalité qui a fait l'objet de critiques, soit parce qu'elle est considérée peu précise, mal définie et trop large (voir par exemple Boye 2002, Douay 2003), soit parce que l'opposition entre un élément subjectif (le *modus*) et un élément objectif (le *dictum*) dans le langage est considérée comme inappropriée (voir Ducrot 1993). Ducrot (1993) en arrive jusqu'à mettre en question la raison d'être du terme de modalité, en maintenant que tout dans le langage relève de la subjectivité. Une notion linguistique qui se base sur l'opposition entre la subjectivité et l'objectivité dans le langage est donc infondée. Or, comme Vion (2004) le maintient, le point de vue de Ducrot, pour qui tout dans le langage remonte à la subjectivité, n'est pas tout à fait incompatible avec une conception de la modalité qui se fonde sur la dichotomie *modus – dictum* ; Vion (2004) partage le point de vue de Ducrot selon lequel le *dictum* ne constitue pas une entité objective qui relève de la pure description, mais qu'il est plutôt une représentation subjective de la réalité telle qu'elle est conçue par le locuteur. Or, même si le *dictum* relève, lui aussi, en fin de compte, de la subjectivité, on peut l'opposer à un *modus* qui traduit la réaction du locuteur

---

<sup>19</sup> Il convient de noter que la théorie du dédoublement énonciatif (Vion) est censée décrire uniquement la modalisation (dans le sens de Vion) et non les modalités, et Nølke affirme qu'il reste à voir si la théorie de la polyphonie peut s'appliquer aux modalités assertées (Nølke 1989 : 61).

<sup>20</sup> Comme nous l'avons vu, la modalité déontique se trouve ici dans une position intermédiaire, vu qu'il n'exprime pas toujours une attitude (cf. note 11, le présent chapitre).

par rapport à cette représentation. Plutôt que de constituer deux éléments dans le langage, l'un objectif et l'autre subjectif, le dictum et le modus constituent deux plans différents dans le discours : il s'agit d'une structure hiérarchique où le dictum constitue un « déjà-là discursif » sur la base duquel le locuteur énonce une réaction (Vion 2004 : 98).

Si la dichotomie modus – dictum est défendable, les conceptions traditionnelles de la modalité en linguistique n'échappent pas à d'autres reproches. Certains ont critiqué l'accent mis sur la modalité comme expression d'une attitude (voir Gosselin 2000 : 58 et Boye 2002 : 6-8) et d'autres encore (par exemple Kronning 1996, 2001a et b et Conte 1998) critiquent la bipartition du domaine modal autour de l'épistémique et du déontique<sup>21</sup>. Nous avons vu que l'insistance sur la modalité comme quelque chose qui exprime une *attitude* tend à mener à une telle bipartition. Selon Kronning (2001b : 98), cette « bipartition déontico-épistémique de l'espace conceptuel de la modalité linguistique » est réductrice et insuffisante. Nous partageons entièrement ce point de vue, car une telle réduction impliquerait qu'un verbe tel que *pouvoir* ait plusieurs valeurs non modales, par exemple quand il dénote une capacité ou une possibilité unilatérale sous-déterminée (cf. la section 3.1.2.2). Cela constituerait une solution bien insatisfaisante, puisque les emplois épistémique et déontique – qui selon cette optique seraient les seuls emplois modaux de cet auxiliaire – ne représentent qu'une partie restreinte de l'ensemble d'occurrences de *pouvoir* (voir aussi Cotte 2002 : 2). Le verbe est toutefois considéré comme un auxiliaire modal, quelle que soit sa valeur précise. La même problématique semble être valable pour *devoir*, bien que ce verbe n'ait pas autant de valeurs que *pouvoir*. En étudiant le verbe *devoir* et ses différentes nuances de sens, Kronning (1996, 2001a et b) est amené à conclure qu'afin de décrire d'une façon satisfaisante le comportement sémantique et syntaxique de ce verbe, il faut prendre en compte un troisième type de modalité, celle qu'il appelle la *modalité aléthique*.

Nous avons donc vu que les conceptions traditionnelles en linguistique, qui se basent sur les théories de l'énonciation et la pragmatique, ne peuvent pas décrire et expliquer tous les types de modalité traditionnellement reconnus dans la tradition philosophico-logique, ni tous les emplois des auxiliaires modaux par excellence, que sont *pouvoir* et *devoir*<sup>22</sup>. Un point de vue différent en linguistique est offert par les approches de Gosselin (2000) et de Kronning (1996, 2003), qui s'inspirent moins d'une tradition énonciative mais d'autant plus d'une

---

<sup>21</sup> Il nous semble que cette bipartition vaut principalement pour les travaux qui portent sur les verbes modaux. Dans les travaux qui portent sur la modalité en général, on trouve très souvent inclus d'autres types de modalité, notamment la modalité axiologique ou évaluative (cf. par exemple Breivega 2003 et Le Querler 1996).

<sup>22</sup> Ce qui ne veut évidemment pas dire que ces conceptions ne peuvent pas être fort fructueuses dans la description de la modalité épistémique et de la modalité axiologique.



tradition philosophico-logique ; il ne s'agit pas pour autant de la tradition aristotélicienne, qui est prédominante dans le domaine de la modalité, mais plutôt de la tradition kantienne. Ces approches tiennent compte du fait que l'absence de marqueurs spécifiques est aussi une manière de marquer l'énoncé (cf. note 1, le présent chapitre), et elles évitent aussi les problèmes liés aux concepts de subjectivité et d'attitude ; on parle moins du locuteur et d'autant plus de l'instance modalisante (Kronning 2003) ou l'instance de validation (Gosselin 2000), – ce qui permettra d'inclure plus facilement des exemples tels que *les élèves doivent se présenter en classe à l'heure* (cf. note 11, le présent chapitre). Le fait d'éviter la notion d'attitude permet aussi d'inclure dans la typologie des modalités des exemples relevant de catégories « objectives » telles que la modalité aléthique, exemplifiée dans Gosselin (2000), entre autres par l'énoncé *Pierre a eu la possibilité de marcher* (ibid. : 65). En tant que modèle global, ces approches semblent donc susceptibles de constituer une alternative fructueuse aux modèles traditionnels, qui n'ont pas, comme nous l'avons vu, la même portée globale.

Par ce qui précède, nous avons relevé certains aspects problématiques de la définition de la modalité qu'on adopte traditionnellement en linguistique. Pour la présente étude, c'est surtout la notion de modalité épistémique qu'il importe de définir de façon très précise ; comme nous l'avons vu, la modalité épistémique est peut-être la seule parmi les types de modalité traditionnellement recensés par la tradition philosophico-logique qui se laisse intégrer sans difficultés dans une définition selon laquelle la modalité exprime une attitude de la part du locuteur envers un contenu informatif. Si nous hésitons à adopter une telle définition de la modalité comme catégorie globale, nous l'utiliserons volontiers pour décrire la modalité épistémique. On pourrait aussi préciser de quel type d'attitude il s'agit (à savoir un jugement sur la valeur de vérité), et la modalité épistémique se définit donc comme l'expression d'un jugement du locuteur par rapport à la valeur de vérité d'un contenu informatif. En ce qui concerne la conception de la modalité en général, nous nous contenterons de dire que la base de la présente étude est une conception médiane qui s'inspire à la fois de la conception philosophico-logique – sans prétendre par là que les modalités logiques peuvent décrire de façon satisfaisante le langage naturel – et de la tradition énonciative. Plus précisément, la typologie adoptée (cf. section 3.1.2) s'inspire de la tradition philosophico-logique alors que l'analyse du rôle des modalisateurs épistémiques dans le discours (cf. chapitre 7) s'inscrit dans une perspective énonciative.

Ci-dessous seront donc présentés les cinq types de modalité recensés par le philosophe et logicien von Wright. Ce choix est motivé par un besoin de distinguer clairement et de façon intelligible la modalité épistémique d'autres types de modalité apparentés. Les multiples

autres modalités mentionnées par différents linguistes (*modalités bouliques, modalités appréciatives, modalités implicatives, etc.*) ne sont pas apparentées à la modalité épistémique de la même manière, et elles sont donc facilement distinguables de celle-ci. Les relations entre les modalités recensées par la tradition philosophico-logique se manifestent par le fait que les mêmes marqueurs linguistiques peuvent renvoyer aux différents types ; cela vaut notamment pour le verbe modal *pouvoir*, qui est susceptible de véhiculer les cinq types de modalité. Pour exprimer les modalités bouliques, appréciatives, implicatives etc., on doit avoir recours à d'autres marqueurs que ceux examinés dans cette étude et ces types n'ont donc pas posé de difficultés pour le codage des marqueurs (voir la section 6.4.1.3 pour plus d'information sur le codage des marqueurs).

### 3.1.2 Types de modalité

Nous présenterons dans cette section les cinq types de modalité évoqués par von Wright (1951). L'accent est mis sur la modalité épistémique, étant donné que c'est ce type de modalité qui est l'objet du présent travail.

#### 3.1.2.1 La modalité épistémique

La modalité épistémique concerne nos connaissances du monde, elle exprime les jugements du locuteur par rapport à la fiabilité de l'information transmise (voir Dendale 1994 : 25). La modalité épistémique qualifie donc la valeur de vérité d'une proposition. Lyons (1994 : 797) en donne la définition suivante :

Any utterance in which the speaker explicitly qualifies his commitment to the truth of the proposition expressed by the sentence he utters, whether this qualification is made explicit in the verbal component [...] or in the prosodic or paralinguistic component, is an epistemically modal, or modalized, utterance.

Pour lier cette modalité de connaissances aux deux opérateurs modaux de la logique modale, on pourrait dire que la possibilité épistémique est une possibilité envisageable, présumée ou probable, c'est-à-dire une possibilité de caractère hypothétique ou une *éventualité*. Ce type de possibilité est parfois appelé une *possibilité bilatérale*, ce qui signifie que la possibilité épistémique exprime « ce qui peut être, mais qui peut aussi ne pas être » (Le Querler 1996 : 37). Dans l'exemple (16), l'état de choses est tel que le sujet *il* peut être arrivé, mais il peut aussi ne pas être arrivé :

(16) Peut-être qu'il est arrivé maintenant

La nécessité épistémique, de son côté, concerne les déductions et les inférences du locuteur :

(17) Il doit être arrivé, puisque je l'ai vu partir d'ici déjà vers 13 heures.

Il ressort de la définition de Lyons que les moyens d'exprimer la modalité épistémique sont d'une grande hétérogénéité. La qualification épistémique peut se réaliser par des moyens extralinguistiques tels que les gestes (par exemple un haussement d'épaules), par la prosodie (par exemple par une intonation interrogative), ou par des moyens grammaticaux ou lexicaux. La présente étude étant une étude linguistique portant sur le discours écrit, il va sans dire que ce sont ces derniers types (les moyens grammaticaux et lexicaux) qui retiennent notre attention dans cette étude. Nous reviendrons dans la section 3.3 sur les différents moyens d'expression de la modalité épistémique.

Il ressort aussi de cette définition que la modalité épistémique est extra-prédicative ou *de dicto*, c'est-à-dire qu'elle porte sur une proposition ou un *dictum*. Rappelons que la modalité *de re* ou intra-prédicative porte sur la relation entre sujet et prédicat (cf. *supra*, section 3.1.1.4), alors que la modalité *de dicto* ou extra-prédicative porte sur une proposition entière. Ainsi, dans l'exemple 18 ci-dessous, le verbe *pouvoir* qualifie la proposition *il est malade*, et dans l'exemple 19 il qualifie la proposition *elle a trente ans*.

(18) Il peut être malade<sup>23</sup>

(19) Elle pouvait avoir trente ans<sup>24</sup>

La distinction *de re* / *de dicto* ou intra-prédicative / extra-prédicative n'est pas à confondre avec la distinction que fait Kronning (1996, 2003) entre modalités véridicibles et modalités montrables, bien que ces deux types d'oppositions se recouvrent dans une large mesure. Les modalités véridicibles font partie de la partie véridicible de l'énoncé, c'est-à-dire la partie qui peut être discutée en termes de vérité et de fausseté. Les modalités montrables n'appartiennent pas à cette partie ; ainsi, les modalisateurs épistémiques montrables qualifient la valeur de vérité d'une proposition mais ne sont pas eux-mêmes véridicibles. Dans l'exemple (16) c'est l'assertion *il est malade* qui est véridicible, tandis que le *peut* du locuteur exprime une possibilité envisagée par celui-ci et le jugement en lui-même ne peut pas être discuté en

---

<sup>23</sup> Exemple emprunté à Tasmowski et Dendale (1994 : 45).

<sup>24</sup> Exemple emprunté à Kronning (1996 : 32).

termes de vérité et de fausseté. Il en va de même pour l'exemple (17), où c'est l'assertion *elle a trente ans* qui est véridicible, alors que le jugement exprimé par *pouvait* ne l'est pas. Dans la terminologie de Kronning (1996), la quantification modale effectuée dans (16) et (17) est donc *montrable*, mais non *véridicible*.

La plupart des expressions épistémiques sont montrables, mais il existe pourtant des marqueurs épistémiques véridicibles<sup>25</sup> (Kronning 1996, 2003, à paraître ; Vet 1997), tels que *il est possible que* et *il est probable que*. Kronning (1996, 2003, à paraître) et Vet (1997) montrent par des tests syntaxiques que ces expressions font partie de la partie véridicible d'un énoncé. Ces expressions sont toutefois extra-prédicatives, dans la mesure où elles qualifient une proposition (par exemple, dans *il est possible qu'il soit malade*, l'expression modale *il est possible que* qualifie la proposition *il est malade*), mais elles contribuent en même temps à une nouvelle proposition, plus complexe, elle-même véridicible (voir aussi Vion 2004 : 98 et Gosselin 2000 : 67). D'un point de vue sémantique le type de modalité exprimée est le même (Vet 1997). La distinction entre modalisateurs épistémiques véridicibles et modalisateurs épistémiques montrables ne sera donc pas très pertinente pour le présent travail, mais le concept de véridicibilité peut nous aider à expliquer la différence entre *pouvoir* épistémique et *pouvoir* dit *aléthique* (*infra*, section 3.1.2.2).

Une troisième chose qui ressort de la définition de Lyons est que la modalité épistémique englobe non seulement les marqueurs exprimant l'*incertitude*, mais aussi ceux qui expriment la *certitude*. Cela est aussi en concordance avec l'optique de Le Querler, affirmant que « par la modalité épistémique, le locuteur exprime son degré de certitude sur ce qu'il affirme » et que « le degré de certitude du locuteur peut aller de la certitude absolue à l'incertitude totale [...] » (Le Querler 1996 : 71). Les expressions épistémiques sont souvent considérées comme situant le contenu propositionnel sur un axe certain – incertain dont les nuances sont multiples et dont les deux extrêmes sont constitués par une assertion simple et sa négation (cf. par exemple Le Querler 1996 : 71, Breivega 2003 : 156, Hyland 1998 : 2). Cependant, l'insistance sur une échelle qui implique un niveau décroissant de certitude ne tient pas suffisamment compte de marqueurs tels que *guère* ('*hardly*', '*neppure*'), *improbable* ('*improbable*', '*usannsynlig*') et *douteux* ('*unlikely*', '*tvilsomt*'), qui marquent un degré de certitude relativement haut, mais sur la forme négative de la proposition et non sur la version affirmative. Il est donc plus correct de dire que l'axe va de la certitude absolue *via* l'incertitude totale et de nouveau vers la certitude absolue. C'est ce que montre la figure 2 sur

---

<sup>25</sup> Parfois inclus sous le terme de *marqueurs aléthiques* chez Kronning.

la valeur de vérité de la proposition *Pierre est dans sa chambre*. Elle peut être qualifiée de diverses manières<sup>26</sup> :

Pierre est dans sa chambre

*Pierre est sans aucun doute dans sa chambre*

*Pierre est certainement dans sa chambre*

*Pierre doit être dans sa chambre<sup>27</sup>*

*Pierre est probablement dans sa chambre*

*Je crois que Pierre est dans sa chambre*

*Pierre est peut-être dans sa chambre*

*Pierre peut être dans sa chambre*

*Il est possible que Pierre soit dans sa chambre*

*Je ne crois pas que Pierre soit dans sa chambre*

*Il est douteux que Pierre soit dans sa chambre*

certitude absolue



incertitude totale



certitude absolue

Pierre n'est pas dans sa chambre

Figure 2 : L'axe épistémique

Comme déjà mentionné, le domaine épistémique est souvent conçu comme un domaine modal entre une assertion simple et sa négation. Ces deux extrêmes sont souvent considérées comme non-marquées ou modalement neutre (voir par exemple Halliday 1994 : 357). Dans cette optique, l'assertion *Pierre est dans sa chambre* et sa négation *Pierre n'est pas dans sa chambre* représenteraient deux pôles d'une échelle, entre lesquels on pourrait imaginer d'innombrables énoncés marqués épistémiquement. La conception de l'énoncé négatif comme neutre et non modalisé peut pourtant être contesté, dans la mesure où presque toute négation peut être perçue comme plus ou moins polémique (Fløttum 2005c), c'est-à-dire que l'énoncé négatif s'oppose à un énoncé implicite, à savoir l'énoncé assertif correspondant. En cela, il est marqué par une polyphonie implicite, et dans un modèle où la modalité est conçue comme un

<sup>26</sup> Il va de soi qu'il est difficile de ranger d'une façon exacte et dans un ordre décroissant les différents marqueurs sur cette échelle : s'il est facile de placer les extrêmes, il ne semble pas possible de distinguer entre, par exemple, *Pierre est peut-être dans sa chambre* et *Pierre peut être dans sa chambre*, en ce qui concerne le niveau de certitude exprimé.

<sup>27</sup> Cet énoncé se prête à deux lectures, l'une épistémique, l'autre déontique, selon la valeur du verbe modal : Dans un contexte où Pierre n'est pas descendu pour dîner avec sa famille, et où sa sœur demande *où est Pierre*, à quoi sa mère répond *mais il n'est pas descendu, il doit être dans sa chambre*, l'énoncé aurait une interprétation épistémique. Inversement, dans un contexte où Pierre a été consigné par ses parents, l'énoncé aurait reçu une lecture déontique.

dédoublé énonciatif (Vion 2001, 2004), il est aussi modalisé, dans le sens où il peut être perçu comme un commentaire sur une autre énonciation (à savoir l'énoncé assertif). Comme nous l'avons vu (cf. *supra*, section 3.1.1.1), même la conception de l'assertion simple comme modalement neutre peut être contestée. C'est en effet elle qui marque le plus haut degré de certitude (voir ci-dessous). Cependant, comme cette étude s'intéresse aux modalisateurs épistémiques en tant qu'atténuateurs, le statut modal qu'on attribue à ce type d'énoncés est une question secondaire.

Certains linguistes<sup>28</sup> tendent à restreindre la zone modale de l'épistémique aux marqueurs de *l'incertitude*. Il existe pourtant de bons arguments pour inclure le marquage de la certitude dans le domaine de la modalité épistémique. Premièrement, l'étymologie de l'adjectif *épistémique*, dérivé du nom grecque *episteme*, qui veut dire *savoir, connaissances*, signale qu'il s'agit de ce qu'on sait (ou prétend savoir) aussi bien que de ce qu'on ne sait pas. Deuxièmement, il y a un glissement graduel de la certitude vers l'incertitude, et il est difficile de couper en deux cet axe. Un énoncé non marqué (c'est-à-dire modalement « neutre ») est souvent perçu comme plus catégorique, plus « sûr », qu'un énoncé marqué par une expression modale de certitude. Ainsi, *Ce traitement est sans risque* peut être perçu comme plus sûr que *Il est certain que ce traitement est sans risque*. Le fait que ce sont souvent les assertions « neutres » qui sont perçues comme les plus sûres est dû à la polyphonie implicite dans les énoncés explicitement marqués comme certains. Par l'adjonction d'un modalisateur épistémique le locuteur signale qu'il envisage la possibilité qu'il existe un doute ou un autre point de vue, et il doit donc souligner qu'à son avis ou en ce moment, il n'y a pas de doute. Ainsi, toute expression épistémique, qu'elle marque la certitude ou le doute, tend en quelque sorte à mettre en question la vérité du contenu propositionnel par le fait qu'elle ajoute un élément de jugement humain, souvent le jugement personnel du locuteur (voir à ce sujet Dyvik 1980, Breivega 2003 : 156 et Lyons 1994 : 808-809). Dans un énoncé non-marqué, comme *Ce traitement est sans risque*, le doute n'est pas même envisagé. Il convient aussi de noter qu'il existe des expressions qui littéralement expriment une certitude, mais qui en réalité ne véhiculent pas du tout cet effet-là. Tel est le cas pour *sans doute* en français et *sikkert* en norvégien, qui littéralement expriment la certitude et l'absence de doute, mais dont le sens correspond à *probablement* et *vraisemblablement*, voire *peut-être*. Les expressions de ce type illustrent bien le glissement graduel entre la certitude et l'incertitude.

---

<sup>28</sup> Voir par exemple Bouscaren et Chuquet (1987 : 37), cités dans Gresset (2002 : 4).

Cependant, si les marqueurs de certitude peuvent signaler une certaine incertitude comparés à l’assertion simple, on peut difficilement dire qu’ils fonctionnent comme des atténuateurs<sup>29</sup>. Ce sont donc les expressions d’incertitude qui sont au centre d’intérêt pour la présente étude, étant donné qu’elles constituent un type important d’atténuateurs, qui à leur tour constituent un trait typique du discours scientifique (Hyland 1998). En général, dans la recherche portant sur le discours scientifique, les expressions d’incertitude semblent avoir reçu plus d’attention que les expressions de certitude, ce qui s’explique sans doute par le simple fait qu’elles sont plus fréquentes.

La liste à la fin de la section 3.3.2 donne une idée des différentes notions sémantiques comprises dans le domaine de la modalité épistémique : on voit que l’axe épistémique va du vérifié et de ce qui est considéré comme évident à ce qui est réfuté ou falsifié (considéré comme prouvé faux), en passant par des stades de spéculation, d’inférence et de déduction, de supposition et de conviction. Evidemment, les marqueurs linguistiques concernés ne véhiculent pas toujours un sens épistémique. Nous reviendrons à la question de la polysémie dans la section 6.4.1.3.

### 3.1.2.2 La modalité aléthique

La modalité aléthique n’est pas une notion dont le sens serait unifié et consensuel. Cette notion est souvent utilisée dans un sens strictement logique, référant au carré logique d’Aristote qui schématise les relations entre quatre types de modalités : le nécessaire, le possible, le contingent et l’impossible. Ainsi, Lyons dit de la modalité aléthique qu’elle concerne « the necessary or contingent truth of propositions » (Lyons 1994 : 791). Les conceptions de la nécessité aléthique en linguistique se basent souvent sur l’acceptation des logiciens, qui tendent à restreindre la nécessité aléthique à des propositions exprimant des vérités analytiques et tautologiques (voir Hughes et Cresswell 1968 et McCawley 1993). Kronning (1996, 2001a et b) opte pour un autre point de vue : il élargit la conception de la nécessité aléthique et y inclut la nécessité synthétique aussi bien que la nécessité logique/analytique – un point de vue partagé par Kolflaath (correspondance personnelle 01.10.2003). Étant donné que c’est la première conception qui semble être prédominante en linguistique, il va de soi que la nécessité aléthique a reçu peu d’intérêt de la part des linguistes ; et pour cause : les énoncés qui l’illustrent sont rares dans le langage naturel. La

---

<sup>29</sup> Cela ne vaut pas pour les marqueurs qui littéralement expriment une certitude, mais qui en réalité véhiculent un sentiment d’incertitude (*sans doute, sikkert*) – ces marqueurs peuvent fonctionner comme de véritables atténuateurs.

possibilité aléthique n'a pas non plus reçu beaucoup d'attention en linguistique, et il est rare de trouver des explications de ce qui est entendu par ce terme. Lyons (1994 : 791) dit que « propositions that are not necessarily false (i.e. propositions that are true in at least one logically possible world) may be described as alethically possible ». Or, cette définition n'aide pas vraiment à opposer le possible aléthique du possible épistémique. Selon Kolflaath (2003), la modalité aléthique s'oppose à la modalité épistémique en ce qu'elle concerne le monde tel qu'il est, alors que la modalité épistémique concerne nos savoirs. La modalité aléthique se situe donc, selon Kolflaath, sur un axe vrai – faux, opposé à l'axe certain – incertain de l'épistémique<sup>30</sup>. Comme l'humanité ne peut pas exister indépendamment de ses connaissances, la différence peut paraître subtile, mais il faut souligner qu'il s'agit de deux manières différentes de *présenter les choses* : les énoncés aléthiques ne décrivent pas nécessairement le monde tel qu'il est, mais ils sont censés *représenter* le monde tel qu'il est (voir aussi Vion 2001 : 212, Gosselin 2000). Cette conception de la modalité aléthique semble concorder avec celle de Gosselin (2000), pour qui la modalité aléthique se distingue de la modalité épistémique entre autre par le choix que le locuteur fait de ce qu'il appelle l'instance de validation : dans les énoncés aléthiques, l'instance de validation est « le réel lui-même » (ibid. : 59), dans les énoncés épistémiques, c'est un sujet qui est l'instance de validation. Une possibilité épistémique est donc une possibilité présentée comme hypothétique, alors qu'une possibilité aléthique est une possibilité présentée comme réelle ou attestée (selon Kolflaath (correspondance personnelle 01.10.2003), il s'agit souvent d'une possibilité *future*). L'opposition se manifeste dans le langage en ce que la modalité aléthique s'inscrit dans la proposition même, alors que la modalité épistémique en est détachée. La possibilité aléthique s'exemplifie ainsi :

(20) L'équipe française peut toujours gagner la coupe du monde.<sup>31</sup>

(21) Il peut faire un 6 (concernant un jet de dés)

Contrairement à *pouvoir* épistémique, *pouvoir* aléthique est intrapredicatif et véridicible. Ainsi, dans 20, la possibilité assertée est vérifiable, par exemple en regardant la liste et la position de l'équipe française sur celle-ci. Et il suffit de regarder le dé pour vérifier la validité de l'énoncé 21. Avec une lecture aléthique, 20 signifie quelque chose comme : il existe encore une possibilité pour l'équipe française de gagner (mais si elle perd le prochain match, elle

<sup>30</sup> Ou plutôt l'axe vérifié – réfuté/falsifié, cf. *supra*, section 3.1.2.1.

<sup>31</sup> Exemple inspiré par un exemple en norvégien, élaboré par Kolflaath (correspondance personnelle 01.10.03).



n'aura plus cette possibilité). Il faut pourtant souligner que, hors contexte, les phrases contenant *peut* sont très souvent polysémiques. Pour l'exemple 20 par exemple, une lecture dynamique vient très naturellement à l'esprit, c'est-à-dire qu'on interprète l'énoncé dans le sens de : *l'équipe française est encore capable de gagner la coupe du monde*, et on pourrait peut-être envisager encore d'autres lectures<sup>32</sup>.

Bien que certains linguistes tiennent compte de la valeur aléthique, la plupart ne le font pas. Selon Kronning (2001b : 103), l'indifférence vis-à-vis de la modalité aléthique s'explique ainsi : ou l'on a ignoré les exemples de la modalité aléthique (ce que l'on a eu tendance à faire dans la tradition linguistique française), ou l'on a reconnu leur existence, mais en insistant sur leur marginalité (ce que l'on a fait dans la tradition linguistique anglo-saxonne, voir par exemple Perkins 1983 et Palmer 1986), ou l'on a intégré l'aléthique dans le domaine de l'épistémique, par une sorte de subordination de l'aléthique à l'épistémique (procédé que suivent entre autres Dendale (1994), van der Auwera et Plungian 1998). Cette indifférence vis-à-vis de la modalité aléthique est possible lorsque le terme est pris dans un sens strictement logique. Dans une optique telle que celle esquissée ci-dessus, où l'aléthique englobe ce qui est présenté comme objectivement vrai (ou faux), il est plus difficile de négliger cette catégorie. Elle peut comprendre des énoncés du type 22, qui exprime la simple constatation d'un fait, c'est-à-dire qu'il y a « affirmation d'une possibilité » (Ducrot 1972 : 66, cité dans Le Querler 1996 : 45) :

(22) La venue de Pierre est possible<sup>33</sup>

Ducrot cherche à opposer un exemple comme 22 à un énoncé comme 23, où « le locuteur prend une certaine attitude, qui n'est ni affirmation ni refus, vis-à-vis de l'événement envisagé » (Ducrot 1972 : 66, cité dans Le Querler 1996 : 45).

(23) Peut-être Pierre va-t-il venir

Il critique la logique modale pour ne pas pouvoir rendre compte de cette distinction, mais cette accusation ne semble pas tout à fait justifiée. Il y a lieu de croire que le premier énoncé serait étiqueté *aléthique* et le second *épistémique*.

---

<sup>32</sup> Ces lectures ne sont pas nécessairement mutuellement exclusives, et elles relèvent plus souvent de l'indétermination que de l'ambiguïté.

<sup>33</sup> Les exemples 22 et 23 sont empruntés à Ducrot (1972 : 66), cité dans Le Querler (1996 : 45).

On peut donc distinguer les exemples comme 20-22, typiques de la possibilité aléthique en ce qu'ils marquent une possibilité future, des exemples épistémiques. Si cette distinction est justifiée, il est encore plus important pour nous d'établir une distinction entre des exemples tels que 24 (qualifié d'aléthique) et 25 (qualifié d'épistémique), car ce type d'exemples est plus fréquent dans notre corpus que les exemples typiques de la possibilité aléthique :

24. Ces formes peuvent théoriquement recevoir deux analyses différentes. D'une part, elles **pourraient** être analysées comme les exemples [5-6], c'est-à-dire comme *en* quantitatif accompagné d'un quantificateur qui serait ici de type nominal. [...] D'autre part, les exemples [1-4] **pourraient** être rapprochés plutôt de [7-8], où l'on trouve *en* adnominal, [...]. (frling06)
25. Ainsi, nous avons estimé que 18 % des bénéficiaires SR ont reçu un traitement correspondant au stade 2. Il reste qu'une certaine proportion de ces personnes sont traitées plutôt en fonction d'un stade 3, car 60 % d'entre elles ont également pris de la théophylline à action prolongée. Cela **pourrait** signifier qu'au lieu d'une augmentation du dosage du corticostéroïde en inhalation, le traitement ferait appel à des doses accrues de médicament au même dosage, ou ajouterait la théophylline à action prolongée plutôt que d'augmenter la dose de corticostéroïde. Cela **pourrait** également indiquer que les personnes ont déjà utilisé de la théophylline, à laquelle ont été ajoutés des corticostéroïdes, ou pour lesquels ils ont été substitués au cours de l'année. (frmed03)

La modalité exprimée dans ces deux exemples, qui à première vue peuvent paraître assez semblables, n'est pas du même type. Dans le premier exemple, l'auteur évoque deux alternatives, deux possibilités d'interprétation présentées comme réelles. Dans le second exemple, l'auteur exprime des hypothèses sur ce que peut signifier l'observation qu'il vient de faire. C'est uniquement dans ce dernier exemple que les occurrences de *pourrait* prennent un sens épistémique. Dans la présente étude, nous avons besoin de pouvoir distinguer ces deux types de possibilité, pour nous centrer sur la possibilité épistémique. Nous qualifions d'aléthiques les occurrences dans le premier exemple, mais la question de savoir si ce type d'exemples peut entrer dans la catégorie philosophico-logique de la possibilité aléthique reste non résolue. Ce n'est donc pas sans hésitation que nous adoptons le terme d'aléthique. Il semble que la possibilité aléthique dans l'optique esquissée ci-dessus soit à peu près assimilable à ce que Blanché (1970 : 67) appelle le *pur possible* ou ce que Granger (1976) appelle le *possible large*. Cependant, opter pour une de ces notions, qui – elles aussi – sont des notions logiques, reviendrait à l'utiliser dans un sens quelque peu différent que celui envisagé à l'origine. Pour opposer ce type de possibilité à la possibilité épistémique, qui se caractérise par sa bilatéralité, on pourrait l'appeler une *possibilité unilatérale* ou avoir recours

à la notion traditionnelle de *possibilité radicale* ('*root possibility*'). Ces notions seraient pourtant trop larges, car elles caractérisent la possibilité déontique et la possibilité dynamique aussi bien que le type de possibilité exprimé dans 22. On pourrait préciser en désignant ce type de possibilité par la notion de possibilité unilatérale (ou radicale) *sous-déterminée*, c'est-à-dire un type de possibilité unilatérale qui ne peut pas être précisé par la valeur de déontique ou de dynamique. Pour des raisons pratiques, nous utiliserons toutefois également le terme de possibilité aléthique, en soulignant qu'il n'est pas à prendre au sens strictement logique.

Pour récapituler, la possibilité unilatérale sous-déterminée exprime la simple constatation d'un fait, l'affirmation d'une possibilité, alors que la possibilité épistémique exprime, comme nous l'avons vu, une possibilité bilatérale, c'est-à-dire une éventualité (ce qui peut être, mais qui peut aussi ne pas être).

### 3.1.2.3 La modalité existentielle

Ce petit exposé sur la modalité existentielle, ou les modes d'existence, se restreindra à la *possibilité existentielle* (excluant la nécessité), car c'est elle qui est susceptible de se confondre avec la modalité épistémique. La possibilité existentielle revient à la *sporadicité*, qui réfère, comme le terme l'indique, à ce qui est sporadique, c'est-à-dire à ce qui se produit de temps en temps, d'une manière irrégulière. Le Querler (2001 : 19-20) remarque que la sporadicité peut être référentielle :

(26) Les Alsaciens peuvent être obèses  
(s'interprète comme *Certains Alsaciens sont obèses*)

ou temporelle :

(27) Jean peut être odieux  
(s'interprète comme *Jean est odieux à certains moments*)

Les modes d'existence de von Wright semblent correspondre au premier type, mais pas forcément au second.

La sporadicité semble être quelque chose d'intermédiaire entre l'aléthique et l'épistémique. Elle peut se rattacher à l'épistémique en ce qu'elle exprime une relation qui *peut être, mais qui peut aussi ne pas être*, c'est-à-dire qu'il s'agit d'une possibilité bilatérale. En outre, les marques de sporadicité peuvent servir à atténuer l'énoncé, le rendre

moins catégorique. Les adverbes de sporadicité (*souvent, parfois*) sont généralement comptés parmi les marqueurs d'atténuation (voir par exemple Hyland 1998). En ce sens, marquer la sporadicité sert à faire des réserves, à s'exprimer avec précaution, ce qui est un trait typique de la modalité épistémique telle qu'elle est employée dans le discours scientifique. Cependant, il y a une différence remarquable entre l'épistémique et le sporadique : celui-ci marque une possibilité qui se fait valoir dans certains cas au moins, alors que celui-là marque une possibilité qui est *susceptible* de se faire valoir, mais qui peut ne se faire valoir dans aucun cas. Comme la sporadicité marque une possibilité réelle et affirmée (réelle parce qu'elle est manifestement réalisée dans certains cas), elle a peut-être plus en commun avec l'aléthique qu'avec l'épistémique. Kronning (1996) la considère aussi comme aléthique, étant donné qu'elle est véridicible et pas montrée (correspondance personnelle 20.10.2003), c'est-à-dire que la modalisation fait partie de la partie véridicible de l'énoncé. Ainsi, 26' est une glose fidèle de 26

26' Il est vrai que < *Les Alsaciens peuvent être obèses* >

alors que 26'' ne l'est pas ; il ne rend pas correctement le sens de 26 :

26'' Il est possible que < *Les Alsaciens être obèses* >

Gosselin (2000 : 64) situe aussi la sporadicité dans le domaine de l'aléthique. Plus nombreux sont pourtant ceux qui rattachent la sporadicité à la modalité épistémique (par exemple Halliday 1994, Le Querler 1996 et Cotte 2002), ce qui s'explique par la parenté qui indéniablement existe entre ces deux catégories. En général, on peut dire que la possibilité épistémique se paraphrase par *peut-être*, et la possibilité sporadique par *parfois*. Mais il n'en reste pas moins que dans certains cotextes, *peut-être* est remplaçable par *parfois*, et en outre, il y a des cas où un énoncé se laisse paraphraser par les deux sans que le choix de l'un ou de l'autre amène à un changement remarquable de sens :

28. Il n'est pas sûr qu'on doive d'abord se donner un objet bien précis, une méthodologie définie et appliquer tout cela ensuite pour obtenir des résultats évaluables par conformité avec des hypothèses. **Il se peut aussi qu'une recherche corresponde davantage à une promenade où ce qu'on rencontre au cours de la promenade modifie la signification qu'on se proposait au début.** (frling12)  
(Paraphrases : Une recherche correspond parfois davantage à... / Une recherche correspond davantage peut-être à...)

Il est donc possible de rattacher la sporadicité au domaine de l'aléthique aussi bien qu'au domaine de l'épistémique. Que l'on choisisse l'une ou l'autre solution, il n'en reste pas moins que la sporadicité est en quelque sorte un hybride de l'aléthique et de l'épistémique, et inévitablement certains exemples dans notre corpus s'interprètent soit comme sporadiques, soit comme épistémiques. Nous reviendrons à ce problème dans le chapitre 6. En règle générale, les occurrences qui expriment la sporadicité ne sont pas incluses dans ce que nous appelons les occurrences épistémiques.

#### 3.1.2.4 La modalité déontique

Avec la modalité épistémique, la modalité déontique est le seul type de modalité qui est établi de manière consensuelle dans la tradition linguistique. Comme nous l'avons déjà vu, cela s'explique par le fait que les modalités déontiques aussi bien que les modalités épistémiques révèlent une attitude du locuteur par rapport à ce qu'il énonce. Tandis que la modalité épistémique révèle l'attitude du locuteur par rapport à la *valeur de vérité* de la proposition, la modalité déontique concerne l'attitude du locuteur par rapport à l'*exécution* de ce qu'il dit. Voilà pourquoi les modalités épistémiques sont traditionnellement appelées les modalités de l'*être* et les modalités déontiques sont appelées les modalités de *faire* (Pottier 1983, Kronning 1996, 2001b). L'instance réceptrice, c'est-à-dire l'interlocuteur, a un statut central dans l'analyse des modalités déontiques, car il s'agit de *faire faire* quelque chose à quelqu'un. C'est cet aspect-là qui lie étroitement les modalités déontiques aux actes directifs.

Si la possibilité épistémique équivaut à l'éventualité, la possibilité déontique équivaut à la permission :

(29) Tu peux rentrer (tu as la permission de rentrer)

Et si la nécessité épistémique revient à des déductions, la nécessité déontique revient à l'obligation :

(30) Pierre doit rester dans sa chambre (Pierre est obligé de rester dans sa chambre)

Les moyens linguistiques typiquement associés à ce type de modalité sont le mode impératif et les verbes modaux *devoir* et *pouvoir*. Cependant, l'impératif et l'emploi déontique de *devoir* ne véhiculent pas toujours une obligation, ils peuvent aussi exprimer des souhaits, des

demandes, des conseils etc. Par ailleurs, le domaine du déontique embrasse, comme il a déjà été mentionné, plus que la permission et l'obligation. Selon Lyons (1994 : 825) et Faarlund et al. (2002 : 580), la modalité déontique concerne aussi les intentions, les désirs et les volontés, alors que d'autres subsument de telles valeurs sous les appellations *modalités bouliques* (Gosselin 2000 : 59, 61) ou *volitives* (Laurendeau 2004).

### 3.1.2.5 La modalité dynamique

Nous finirons par une brève introduction à la modalité dynamique, sans nous y attarder, car ce type de modalité semble assez facilement distinguable de la modalité épistémique. Le terme de modalité dynamique est peu utilisé en français, mais le correspondant en anglais, *dynamic modality*, s'utilise de plus en plus (voir par exemple Palmer 2001), et Bentley (2003 : 64) le traduit en français par *modalité dynamique*. Ce type de modalité concerne, selon von Wright (1951), les modes de capacité, et la possibilité dynamique dénote donc des capacités, des aptitudes et des disponibilités (Bentley 2003 : 64). (31) est un exemple de la possibilité dynamique :

(31) Je peux soulever cette valise<sup>34</sup>

En français, *pouvoir* dynamique alterne avec *savoir* : *Jean peut / sait parler l'anglais*.

Traditionnellement, l'emploi dynamique de *pouvoir* a été considéré comme non modal (voir par exemple Faarlund et al. 2002 : 581), sans doute parce qu'on ne peut dire qu'il exprime une « attitude subjective » ou un « point de vue ». Cependant, comme il a déjà été mentionné, plusieurs linguistes commencent à attribuer à cet emploi une valeur modale.

### 3.1.3 Bilan

En résumé, nous avons vu qu'il existe de nombreuses conceptions de la modalité. Ces conceptions se repartissent globalement en deux groupes majeurs : l'un comprend des caractérisations de la modalité à partir des termes de possibilité et de nécessité, l'autre comprend des caractérisations en termes d'attitude ou de point de vue du locuteur. Ces deux conceptions ne semblent pas conciliables, étant donné qu'il y a plusieurs types de possibilité et de nécessité qui n'expriment pas l'attitude ou le point de vue du locuteur. Dans le domaine de la linguistique, il semble donc qu'on oscille entre deux positions divergentes, ce qui

---

<sup>34</sup> Exemple emprunté à Kronning (1996 : 32).

entrave l'unanimité par rapport à la conception de la modalité. Or, il importe dans la présente étude de définir la modalité épistémique d'une façon précise, et ce type de modalité se laisse facilement intégrer à une définition traditionnelle de la modalité en linguistique, selon laquelle la modalité exprime l'attitude du locuteur par rapport au contenu informatif de son énoncé. Cependant, afin de distinguer la possibilité épistémique d'autres types de possibilité, il a fallu parcourir les divers types de possibilité. Plus précisément, nous sommes partie des types de possibilités recensés dans la tradition philosophico-logique depuis von Wright (1951), afin de voir comment ils se distinguent de la possibilité épistémique. Ayant ainsi essayé d'identifier la valeur effectivement épistémique parmi la multiplicité de valeurs modales (ou, si l'on veut, de types de possibilité), nous pouvons maintenant nous tourner vers la catégorie d'épistémicité afin de distinguer les marqueurs de la modalité épistémique et les marqueurs strictement évidentiels.

## 3.2 L'épistémicité

### 3.2.1 Définir la catégorie de l'épistémicité

Par épistémicité – terme dérivé du nom grec *episteme*, qui signifie *connaissances* – nous entendrons une catégorie linguistique qui regroupe les moyens langagiers dont dispose le locuteur pour s'exprimer sur ses connaissances du monde. On tient compte de deux stratégies majeures pour s'exprimer sur ses connaissances. En premier lieu, le locuteur peut marquer son degré de certitude, c'est ce que nous avons appelé la modalité épistémique, qui est exemplifiée par les énoncés suivants :

(1) Je suis sûr que Pierre a joué au football hier soir

(2) Peut-être que Pierre a joué au football hier soir

En second lieu, le locuteur peut indiquer (explicitement par le langage) la source ou la nature de la source de l'information transmise par l'énoncé, et c'est ce que nous appellerons l'évidentialité (cf. Kronning 2003 : 131). En (3), par exemple, le locuteur indique qu'il tient l'information de quelqu'un d'autre et dans (4), il indique qu'il l'a acquise de ses propres yeux:

(3) Selon Marie, Pierre a joué au football hier soir

(4) Hier soir, j'ai vu Pierre jouer au football

L'épistémicité n'a pas reçu beaucoup d'attention en tant que catégorie unifiante, mais ces dernières années, les sous-domaines de la modalité épistémique et de l'évidentialité ont chacun de leur côté suscité un intérêt croissant de la part des linguistes. La modalité épistémique a été amplement étudiée dans la linguistique anglo-saxonne (voir par exemple Coates 1983 et Perkins 1983). Dans les recherches portant sur le discours scientifique, c'est surtout le phénomène d'atténuation qui a reçu beaucoup d'intérêt. Il existe un nombre considérable d'études portant sur l'atténuation, et la plupart d'entre elles se concentrent sur les modalisateurs épistémiques (voir par exemple Salager-Meyer 1994, 1997 ; Hyland 1996a et b, 1998 ; Vihla 2000 et Varttala 1999, 2001). En linguistique française, les études sur la modalité épistémique sont peut-être moins nombreuses, mais il existe plusieurs travaux importants qui traitent de la modalité épistémique sous différents angles (voir par exemple Le Querler 1996, Dendale 1994, 1999, Kronning 2001b, 2002, 2003). Ces études portent sur le langage général et ne se restreignent pas à un type spécifique de discours. Contrairement à



ce que l'on observe dans la tradition anglo-saxonne, il n'existe pas dans la linguistique française de tradition consistant à étudier l'emploi de la modalité épistémique dans le discours scientifique.

Le phénomène d'évidentialité a fait l'objet de multiples études ces dernières années : en 1994 a paru un numéro spécial de *Langue française* entièrement consacré à ce phénomène, et marquant, de bien des manières, le début des études portant sur l'évidentialité en linguistique française (Dendale et Tasmowski 1994 : 6-7). Sept ans plus tard a paru un numéro spécialisé de *Journal of pragmatics* (Dendale et Tasmowski 2001a), lui aussi consacré au phénomène d'évidentialité et rédigé par les mêmes chercheurs. Ce numéro est, autant que nous sachons, le deuxième numéro spécial en anglais qui traite de l'évidentialité, le premier étant la bien connue collection d'articles rédigée par Chafe et Nichols en 1986.

Bien que l'intérêt augmente pour la modalité épistémique ainsi que pour l'évidentialité, le terme d'épistémicité est peu utilisé comme terme générique. Boye (2002) utilise le calque danois *epistemicitet* et Dendale et Tasmowski (1994 : 4) proposent *marquage épistémique* comme hypéronyme des deux catégories, un terme repris par Kronning (2003). Dans ce qui suit, nous parlerons d'épistémicité et de marquage épistémique de façon interchangeable.

Dans l'optique adoptée ici, le domaine de l'épistémicité se scinde donc en deux : la modalité épistémique concerne le degré de certitude du locuteur et la fiabilité qu'il accorde à ce qui est dit, ce qui implique que la modalité épistémique exprime une *attitude* par rapport à la valeur de vérité d'une proposition. L'évidentialité concerne les sources du savoir : il ne s'agit pas d'exprimer une *attitude* par rapport à l'information transmise, mais d'explicitier d'où ou de qui on tient cette information. Comme plusieurs linguistes romanistes l'ont déjà remarqué (Guentchéva 1994 : 9, Dendale et Tasmowski 1994 : 3, Kronning 2003 : 132), le terme d'évidentialité peut sembler mal approprié, étant donné que l'on l'associe facilement à ce qui est évident, alors que le sens est plutôt le contraire : il s'agit d'expliquer comment on a appris ce qu'on dit. Le terme est un anglicisme, dérivé du nom anglais *evidence*, signifiant « preuve », et n'a donc rien à voir directement avec la notion d'évidence en français. Voilà pourquoi certains linguistes romanistes préfèrent utiliser le terme de *médiatif* plutôt que celui d'évidentialité (par exemple Guentchéva 1994).

Pour résumer, la modalité épistémique concerne la valeur de vérité de l'information transmise par l'énoncé, alors que l'évidentialité concerne les sources de cette information. Les deux catégories ont ceci en commun qu'elles apportent un commentaire sur nos connaissances du monde, et par conséquent, qu'elles appartiennent à l'espace épistémique.

### 3.2.2 Conceptions différentes

Comme nous l'avons déjà indiqué, ce modèle de classification ne fait pas l'unanimité. Pour le moment il n'existe pas de consensus quant aux relations qui existent entre l'évidentialité et la modalité épistémique. Le fait d'indiquer la source de l'information (évidentialité) implique parfois un marquage plus ou moins explicite du degré de fiabilité de cette information. Selon la nature et la crédibilité de la source, l'information transmise peut être perçue comme plus ou moins certaine. Ce constat est à l'origine d'une certaine confusion en ce qui concerne les rapports entre les deux catégories. Dendale et Tasmowski (2001b : 342) remarquent que certains linguistes préfèrent considérer l'évidentialité comme une sorte de modalité épistémique, alors que d'autres préfèrent inclure la modalité épistémique dans le domaine de l'évidentialité. Les adhérents du premier point de vue maintiennent – comme il vient d'être mentionné – que tout marquage de la source marque, de façon plus ou moins directe, une attitude par rapport à la valeur de vérité du message transmis par l'énoncé. Dans cette hypothèse, l'évidentialité n'est qu'un sous-type de la modalité épistémique, et telle était la classification proposée par Palmer dans la première édition de son *Mood and modality* (1986). Telle est aussi la classification de Hyland (1998), qui maintient que « epistemic modality clearly encompasses evidentiality » (ibid. : 47).

Vue sous cet angle, l'évidentialité sera directement reliée à la catégorie de la modalité. Un tel point de vue se heurtera, à notre avis, à de grandes difficultés quant à la délimitation des catégories en question. Car bien qu'il soit facile de voir l'effet modal de certains marqueurs évidentiels tels que *il paraît que*, qui sert à marquer la proposition qui suit comme une information incertaine (en même temps qu'il signale un lien de non-responsabilité entre le locuteur et l'information communiquée, cf. Nølke 1993), il n'est pas aussi facile de détecter un effet modal du marqueur évidentiel *selon x* dans l'énoncé (3) ci-dessus, repris ici pour des raisons de facilité de lecture :

(3) Selon Marie, Pierre a joué au football hier soir

Le marqueur *selon x* exprime aussi un lien de non-responsabilité entre le locuteur et l'information communiquée, mais à l'opposé de *il paraît que*, il n'implique pas nécessairement que cette information soit qualifiée d'incertaine. L'interprétation de (3) variera naturellement selon le cotexte et la situation d'énonciation, mais hors contexte, il n'y a rien dans cet énoncé qui qualifie directement la valeur de vérité du contenu propositionnel. Le

locuteur ne se prononce pas sur la valeur de vérité, mais il indique la source de l'information. Or, on peut facilement s'imaginer une situation où le marqueur évidentiel *selon x* influence le degré de certitude qu'on attribue à l'information exprimée. Dans une situation où le locuteur dit : *Selon Marie, Pierre a joué au football hier soir*, à quoi l'allocutaire répond : *Marie ! Elle ne fait que mentir*, l'information est perçue comme incertaine parce que la source n'est pas digne de confiance (selon l'allocutaire), mais il n'y a pas de marques *linguistiques* de cette incertitude. L'incertitude naît d'une connaissance du référent de *Marie*, et n'est pas exprimée par le langage. On pourrait également s'imaginer une situation où *selon x* sert à renforcer la fiabilité de l'information, par exemple quand le référent de *x* est reconnu comme une autorité dans le domaine dont on parle. De ces exemples on peut conclure que certains marqueurs évidentiels (comme *il paraît que*) marquent, par leur contenu sémantique même, le degré de certitude, alors que d'autres (comme *selon x*) ne le font pas. Le fait d'inclure l'ensemble de marqueurs évidentiels dans la catégorie de la modalité épistémique exigerait donc une définition élargie de cette catégorie, et par conséquent, une définition élargie de la catégorie de la modalité en général.

Le second point de vue suppose, selon Dendale et Tasmowski (2001b : 342-343, 1994 : 4), une conception large de l'évidentialité où la catégorie de l'évidentialité englobe non seulement les spécifications de la source du savoir, mais aussi les spécifications de la valeur de vérité. Dans cette optique, le terme d'évidentialité devient donc un quasi-synonyme de notre *épistémicité*. L'inconvénient de ce modèle est que ce que nous avons appelé l'évidentialité (sens restreint) ne peut être désigné par un seul terme : on doit utiliser des paraphrases comme 'spécification de la source de l'information' (voir Dendale et Tasmowski 2001b : 343). Les spécifications de la valeur de vérité, par contre, sont toujours désignées par le terme de 'modalité épistémique'.

Ces trois conceptions sont sans doute les plus fréquentes, mais Kronning liste six (!) différents types de relations entre les deux catégories qu'il a trouvés attestés dans la littérature linguistique (voir Kronning 2003 : 135). Malgré cette grande diversité dans les types de relations, il nous semble que la linguistique se dirige vers un consensus, selon lequel l'évidentialité et la modalité épistémique constituent deux sous-espaces en théorie disjoints qui relèvent de la catégorie de l'épistémicité, autrement dit une conception comme celle adoptée dans la présente étude. Dendale et Tasmowski (1994 : 4, 2001) argumentent en faveur d'une telle conception, et ce classement est repris par Kronning (2003), qui l'élabore et l'affine. Boye (2002) opte pour cette même solution dans ses études du danois, et Palmer (2001), affinant ses analyses de 1986, traite l'évidentialité et la modalité épistémique comme

deux catégories distinctes, contrairement à ce qu'il faisait dans la première édition (1986) où l'évidentialité était subordonnée à la modalité épistémique. De Haan (2001) et Squartini (2004) défendent aussi une distinction nette entre ces deux catégories, mais sans les regrouper sous une même catégorie générique.

Notre vue est donc en concordance avec celle de Kronning (2003), Boye (2002) et Dendale et Tasmowski 1994 : 4, 2001b), mais elle diffère un peu de celle de Palmer (2001), qui continue à réunir la modalité épistémique et l'évidentialité sous la catégorie principale de *modalité*, et non sous celle de l'épistémicité / du marquage épistémique.

Bien que l'évidentialité et la modalité épistémique soient deux catégories théoriquement disjointes, il existe, naturellement, de nombreux marqueurs qui relèvent des deux catégories à la fois. Nous avons déjà mentionné la locution *il paraît que*, qui exprime à la fois le oui-dire (évidentialité) et un certain degré d'incertitude (modalité épistémique). Kronning appelle ces marqueurs des *marqueurs mixtes* (2003) ou *bicatégoriels* (à paraître), c'est-à-dire des marqueurs qui expriment en même temps une modalité épistémique et l'évidentialité. Récemment on a proposé de considérer des marqueurs traditionnellement reconnus comme épistémico-modaux, tels que *devoir<sub>E</sub>*<sup>35</sup> et *pouvoir<sub>E</sub>* en français et *may<sub>E</sub>* et *must<sub>E</sub>* en anglais, comme des marqueurs évidentiels (voir Dendale 1994, Tasmowski et Dendale 1994, Drubig 2001 et Kronning 2003 : 131). Que ces marqueurs soient considérés comme des marqueurs évidentiels n'empêche pas qu'ils ne soient en même temps considérés comme des marqueurs modaux épistémiques. De ce point de vue, les catégories de l'évidentialité et de la modalité épistémique se chevauchent et s'entremêlent. Certains de ces chevauchements seront explicités dans la section suivante.

### 3.2.3 Evidentialité et marqueurs mixtes

L'évidentialité est définie par Dendale et Coltier (2003 : 106) comme « le phénomène linguistique qui consiste à indiquer par des moyens langagiers *comment* le locuteur sait ce qu'il dit », une définition qui, selon les auteurs, s'inspire de celle de Hardman (1986 : 115). Il s'agit donc de signaler explicitement, par des marqueurs morphologiques, lexicaux ou autres, la source, ou la nature de la source, de l'information transmise. Tous les marqueurs évidentiels ne sont pas pertinents pour cette étude. Ce sont les marqueurs mixtes, ou les marqueurs *évidentio-modaux* dans la terminologie de Dendale (1999), qui nous intéressent,

---

<sup>35</sup> Le <sub>E</sub> signifie qu'il s'agit de la valeur épistémique de ces verbes modaux.

puisque ces marqueurs, tout en exprimant une valeur évidentielle, véhiculent en même temps une valeur épistémico-modale. Quels sont alors les marqueurs mixtes ?

Prenons comme point de départ les trois grandes sources d'information traditionnellement reconnues : la perception, l'inférence et l'emprunt (cf. Dendale et Tasmowski (1994 : 4) et Dendale et Coltier (2003 : 107)) :

- *la perception*

Il s'agit notamment de la perception visuelle ou auditive, comme exemplifié par les énoncés suivants :

(5) J'ai vu Pierre jouer au football

(6) J'ai entendu Pierre chanter

Les marqueurs indiquant *la perception* ne sont pas censés qualifier explicitement la valeur de vérité, bien que le contenu informatif d'un énoncé introduit par *j'ai vu* soit, dans la grande majorité des cas, considéré comme vérifié. La valeur de vérité n'est pas explicitement qualifiée par l'expression *j'ai vu*, mais la perception visuelle est jugée être une source très fiable et par conséquent, l'information transmise sera considérée comme un fait certain (pourvu qu'on fasse confiance au locuteur). Par le marqueur *j'ai vu*, le locuteur indique explicitement la manière dont il s'est procuré l'information, mais son degré de certitude par rapport à l'information transmise n'est indiqué que d'une façon implicite.

- *l'inférence*

Par *inférence* on entend le procès consistant à tirer une conclusion en se basant sur des indices ou sur d'autres types de connaissances. Selon Dendale (1999 : 8), le terme d'inférence couvre «aussi bien la *déduction* et *l'induction* que le *calcul* ou *l'estimation* ». L'énoncé suivant exemplifie une déduction :

(7) Pierre doit être allé jouer au football (car ses chaussures de football ne sont pas là).

En français, les marqueurs évidentiels indiquant *l'inférence* expriment très souvent (mais pas toujours, voir Kronning 2003 : 143-144) une modalité épistémique. Il s'agit notamment des auxiliaires modaux *devoir* et *pouvoir*, qui dans leur acception épistémique fonctionnent, selon Tasmowski et Dendale (1994), aussi comme des marques d'une inférence : *devoir<sub>E</sub>* indique

que le locuteur ne voit qu'une seule conclusion, alors que *pouvoir<sub>E</sub>* indique que le locuteur reste ouvert à d'autres possibilités :

(8) Ça doit être Jean (seule conclusion possible)

(9) Ça peut être Jean (une parmi plusieurs conclusions possibles)

(cf. Tasmowski et Dendale 1994 : 54-55)

Un autre marqueur d'inférence est la particule *donc*. Cette conjonction peut marquer un haut degré de certitude. On pourrait donc dire qu'elle implique, dans bien des cas, une modalité épistémique, mais elle se situe plutôt du côté *certitude* de l'échelle épistémique :

(10) Ça sent bon et la table est mise. Marie a donc déjà préparé le dîner.

Il convient de rappeler que *donc* possède toute une série d'emplois, et l'emploi évidentio-modal d'inférence n'est probablement pas le plus fréquent (voir Didriksen 2004 pour une présentation de différents emplois de *donc*).

- *l'emprunt*

La catégorie de l'emprunt ne semble pas très bien circonscrite. Les exemples les plus typiques dans la littérature contiennent soit un conditionnel de citation soit une locution telle que *selon X* ou *il paraît que*. Cette dernière locution signale que l'information est empruntée à quelqu'un d'autre (non spécifié), plus précisément qu'il s'agit d'un oui-dire ou d'une rumeur :

(11) Il paraît que Pierre est malade

Les exemples types sont pourtant rarement accompagnés d'une définition précise de la catégorie de l'emprunt. Kronning (2003) fait figure d'exception, en argumentant en faveur d'une distinction nette entre l'emprunt et le discours rapporté. Selon lui, ces deux catégories se distinguent en ce que le discours rapporté rend compte d'un acte d'énonciation et le présente comme tel, alors que par l'emprunt, on transmet une information en l'attribuant à quelqu'un d'autre « *sans faire référence à cet acte d'énonciation* » (ibid. : 143). Dans cette optique, un exemple typique du discours rapporté tel que *Marie a dit que Pierre joue au football*, ne sera pas un exemple d'emprunt, étant donné qu'en l'énonçant, le locuteur fait explicitement référence à l'acte d'énonciation du locuteur source. Les énoncés suivants, par

contre, seront des exemples d'information empruntée, car l'acte d'énonciation n'est pas explicitement indiqué :

(12) Le président serait malade (exemple emprunté à Kronning 2003)

(13) Selon Marie, Pierre joue au football

D'autres linguistes (Kronning mentionne Li 1986) assimilent le discours rapporté à l'emprunt, et d'autres encore ne font aucune remarque sur cette problématique. La distinction entre le discours rapporté et l'emprunt nous semble nécessaire, pourtant il faut avouer que cette distinction n'est pas toujours évidente au niveau du sens. Car si *il paraît que* sera de ce point de vue considéré comme un marqueur évidentiel de l'emprunt, la paraphrase *on dit que* ne le sera pas, bien que le sens soit à peu près le même. La question n'est pas centrale dans cette étude, mais les relations entre l'information empruntée et le discours rapporté méritent d'être examinées plus profondément.

Les marqueurs évidentiels marquant *l'emprunt* ont ceci en commun qu'ils signalent que le locuteur ne prend pas en charge l'information transmise – il n'en est pas responsable. Les marqueurs d'emprunt impliquent souvent une modalité épistémique d'incertitude, comme avec le conditionnel de citation<sup>36</sup> (cf. l'exemple 12) et la locution *il paraît que* dans l'exemple ci-dessous :

(14) Il paraît que le président est malade

D'autres marqueurs de l'emprunt ne font qu'indiquer la source, sans pour autant marquer le degré de certitude. Il s'agit par exemple du marqueur *selon x* (voir plus haut). Les énoncés introduits par ce marqueur peuvent être perçus comme de simples rapports : le locuteur constate que quelqu'un a dit ceci ou cela, sans exprimer sa propre attitude. Comme il a déjà été noté, le message d'un tel énoncé peut être perçu comme une information incertaine, si par exemple la source est jugée indigne de confiance. Or, l'incertitude n'est pas marquée linguistiquement, comme elle l'est pour les marqueurs *il paraît que*, *devoir* épistémique etc.

Comme le montrent les exemples cités ci-dessus, les marqueurs évidentiels en français constituent un groupe assez hétérogène, de morphèmes grammaticaux (par exemple le tiroir -

---

<sup>36</sup> Dendale (1993, 1999) a montré que le conditionnel de citation n'assume pas toujours une valeur d'incertitude, il peut également signaler l'indifférence du locuteur par rapport à la valeur de vérité de l'information rapportée. Selon Dendale, cela s'explique par le fait que c'est la valeur de non prise en charge qui est la première. Cette valeur de base se double très souvent d'une valeur d'incertitude, qui reste pourtant toujours secondaire.

*rait* dans le conditionnel de citation) à des marqueurs lexicaux variés comme *j'ai vu, selon x* etc., en passant par des verbes modaux (par exemple *devoir* épistémique). D'autres langues disposent de systèmes morphologiques affinés pour indiquer la source de l'information transmise, et parfois le marquage de la source est obligatoire. L'exemple le plus cité est celui du tuyuca (initialement discuté par Barnes 1984), une langue amérindienne où le marquage évidentiel est à la fois très précis et obligatoire (voir par exemple Dendale et Tasmowski 1994 : 4, Kronning 2003, Palmer 1986). Dans cette langue, des morphèmes grammaticaux spécifiques sont attachés aux propositions afin d'indiquer quelle est la nature de la source. Le morphème *-wi* par exemple, indique que l'information provient de la perception visuelle.

Le phénomène d'évidentialité est donc parfaitement grammaticalisé dans certaines langues, alors que dans d'autres, comme le français, l'anglais et le norvégien, les marqueurs évidentiels constituent un groupe plus hétérogène qui est moins facile à cerner. Il n'y a pas lieu ici de fournir une liste des divers marqueurs susceptibles d'exprimer l'évidentialité. Mais, comme nous l'avons vu, plusieurs de ces marqueurs expriment en même temps une modalité épistémique. Le chapitre suivant fournira une liste des différents moyens linguistiques qui expriment la modalité épistémique, comprenant, mais pas exclusivement, les marqueurs évidentio-modaux.



### 3.3 La réalisation linguistique de la modalité épistémique

#### 3.3.1 Moyens grammaticaux et moyens lexicaux

Ayant précisé, dans les sections précédentes, le sens que nous accordons à la notion de *modalité épistémique*, il est maintenant temps de regarder de plus près les moyens linguistiques par lesquels cette catégorie modale se réalise dans les trois langues. Cet exposé se restreindra aux moyens *linguistiques* et *écrits* pour exprimer la modalité épistémique, c'est-à-dire que les moyens extralinguistiques tels que les gestes, ainsi que les moyens prosodiques (l'intonation, l'accentuation) qui ne sont pertinents que pour la langue parlée, ne seront pas considérés.

On a déjà vu que les moyens d'exprimer la modalité épistémique sont d'une grande diversité (cf. *infra* 3.1.2.1). Il est commun de diviser ces moyens en deux groupes principaux : les moyens grammaticaux d'un côté et les moyens lexicaux de l'autre (Lyons 1981 : 238, Andersen 2003 : 43sq, Palmer 1986 : 6). Les moyens grammaticaux comprennent typiquement des catégories telles que les modes, les désinences temporelles à valeur modale et les auxiliaires modaux. Les moyens lexicaux renvoient à une série de marqueurs différents, tel que verbes, adjectifs et adverbes épistémiques. Hyland (1998 : 45) souligne en outre l'importance de moyens syntaxiques tels que la passivation, l'interrogation et les subordonnées de condition. Il est pourtant souvent difficile de distinguer les moyens lexicaux et les moyens dits syntaxiques, puisque ce sont souvent les lexèmes utilisés et non les structures syntaxiques elles-mêmes qui véhiculent la modalisation. Le passif par exemple, n'est pas en soi une structure modalisante, mais des constructions au passif (telles que celle citée par Hyland (1998 : 46) : *the large cells [...] are suspected to be...*) peuvent avoir un effet modal et atténuant en raison des lexèmes choisis (en l'occurrence le verbe *suspect*). Dans cette section, nous nous contenterons de distinguer les moyens grammaticaux et les moyens lexicaux. Dans l'analyse exploratoire qui suit (voir section 6.3), nous cherchons cependant à être attentive à toutes sortes de moyens susceptibles d'exprimer une modalité épistémique, y compris les constructions syntaxiques évoquées par Hyland (1998). Soulignons aussi le besoin de tenir compte du rôle important que jouent les structures syntaxiques à l'égard de certains lexèmes spécifiques, par exemple à l'égard de l'adjectif *possible*. La valeur qu'on accorde à cet adjectif dépend en grand partie de la construction syntaxique dans laquelle il apparaît (voir *infra*, section 6.4.1.3)

Les termes de « moyens grammaticaux » et de « moyens lexicaux » ne sont pas toujours employés de façon uniforme. Il convient alors d'expliquer notre conception de ces

deux catégories : Par *moyens grammaticaux* nous entendrons, en bref, des éléments morphosyntaxiques. Plus précisément, il s'agit d'éléments sémantiquement pauvres qui sont *non autonomes* dans le sens qu'ils exigent la présence d'un autre mot afin de pouvoir former un énoncé sensé. Les moyens grammaticaux se divisent à leur tour en deux groupes : les éléments morphologiques ou synthétiques (modes, tiroirs verbaux à valeur modale) et les éléments syntaxiques ou analytiques (auxiliaires modaux). Les flexions verbales sont inséparablement liées au verbe qu'elles modifient et les auxiliaires dépendent d'un verbe principal. Ces moyens ont donc en commun le fait d'être non autonomes, mais ils signalent pourtant un continuum entre éléments strictement grammaticaux (morphologie verbale) et éléments qui tendent plus vers la lexicalité (auxiliaires).

Les *moyens lexicaux* constituent une catégorie hétérogène qui regroupe des lexèmes relevant de nombreuses parties du discours (voir aussi Hyland 1996a ou 1998 : 102-156). Par exemple, les trois langues disposent toutes de verbes<sup>37</sup>, d'adjectifs<sup>38</sup> et d'adverbes<sup>39</sup> épistémiques.

Il est à noter que, du fait des processus de grammaticalisation, la distinction entre moyens grammaticaux et moyens lexicaux n'est jamais nette et définitive. Par *grammaticalisation* nous entendons le processus par lequel certains éléments lexicaux se transforment en éléments grammaticaux servant des fonctions grammaticales (Hopper et Traugott 1993 : xv). Il est communément admis que les auxiliaires modaux de plusieurs langues se sont peu à peu transformés de verbes lexicaux en auxiliaires. Étant donné que la grammaticalisation est un processus et un développement par degrés, il est parfois difficile de déterminer si un moyen linguistique est grammatical ou lexical. Ces frontières floues se reflètent dans le fait que les auxiliaires modaux sont par certains chercheurs (Hyland 1998, Lyons 1981 : 238) considérés comme des moyens lexicaux plutôt que grammaticaux. Qui plus est, il existe une série de verbes dits *semi-auxiliaires*, qui ne fonctionnent ni exactement comme des auxiliaires, ni tout à fait comme des verbes lexicaux : *sembler, laisser, paraître* etc. Plutôt que de constituer deux groupes distincts, les éléments grammaticaux et les éléments lexicaux forment un continuum, tout comme il y a un continuum au sein des moyens grammaticaux, du synthétique à l'analytique. Une classification d'éléments modalisants en « moyens grammaticaux » et « moyens lexicaux » ne se laisse donc pas faire de façon catégorique, et pourrait toujours être mise en discussion. Néanmoins, dans la figure ci-dessous

---

<sup>37</sup> Par exemple *supposer* (français), *assume* (anglais), *anta* (norvégien)

<sup>38</sup> Par exemple *possible* (français et anglais), *mulig* (norvégien)

<sup>39</sup> Par exemple *probablement* (français), *probably* (anglais), *sannsynligvis* (norvégien)

nous proposons une classification pour l'analyse de la réalisation linguistique de la modalité épistémique :

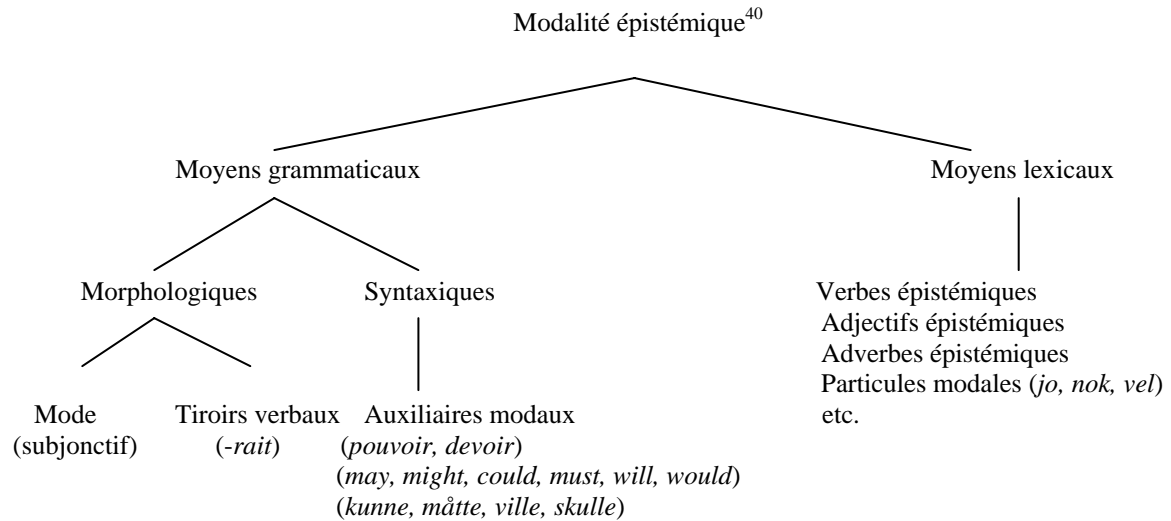


Figure 3 : La réalisation linguistique de la modalité épistémique

Partant de ce modèle, nous donnerons dans ce qui suit une brève présentation des moyens grammaticaux et lexicaux employés pour exprimer la modalité épistémique en français, anglais et norvégien. Cet aperçu de la réalisation linguistique de la modalité épistémique ne prétend point à l'exhaustivité, mais vise à illustrer la diversité qui caractérise la catégorie de marqueurs épistémico-modaux. Il sert de toile de fond pour l'analyse exploratoire dont les procédés seront décrits au chapitre 6 et qui a pour but de mener à une sélection bien-fondée de marqueurs à étudier.

Comme nous allons le voir, les trois langues disposent dans une large mesure des mêmes moyens, mais elles possèdent aussi certaines caractéristiques spécifiques dont nous tenterons de rendre compte dans ce qui suit.

<sup>40</sup> Le modèle n'intègre que les moyens disponibles à l'anglais, le français et le norvégien. Pour d'autres langues, il faudrait ajouter d'autres catégories.

### 3.3.2 La réalisation linguistique de la modalité épistémique en français, anglais et norvégien<sup>41</sup>

#### 3.3.2.1 Moyens grammaticaux

La modalité épistémique peut, comme il a déjà été mentionné, s'exprimer grammaticalement par les modes, les temps verbaux et les auxiliaires modaux<sup>42</sup>.

#### *Mode*

En général, c'est avant tout le mode du subjonctif qui est associé à la modalité épistémique. L'impératif est associé à la modalité déontique, et l'indicatif est traditionnellement considéré comme un mode neutre ou non marqué<sup>43</sup>. Il est vrai que le subjonctif du français est traditionnellement lié à l'hypothétique, l'irréel, l'incertain et le virtuel. Ces caractéristiques le lient à la modalité épistémique, mais comme nous l'avons vu (cf. section 3.1.1.3), il n'existe pas de relation obligatoire entre le mode du subjonctif et la modalité épistémique. Selon Riegel et al. (1998 : 288, 322) le subjonctif n'exprime pas en soi une modalité : c'est plutôt le contexte, par exemple le verbe principal, qui véhicule une valeur modale. Le subjonctif peut ainsi apparaître dans des contextes déontiques aussi bien que dans des contextes épistémiques. De plus, il n'exprime pas toujours la virtualité, qui est censée être sa valeur de base, mais peut fort bien exprimer des faits réels (voir *infra* 3.1.1.3). Même s'il ne fonctionne normalement pas comme un marqueur épistémico-modal à part entière, le subjonctif sert souvent à reprendre et renforcer la modalité du verbe régissant. Dans l'exemple ci-dessous, le subjonctif exprime, selon Huot (1986, citée dans Boysen 2003 : 46), une réserve plus grande que l'indicatif et il fonctionne alors comme un marqueur de la modalité épistémique, mais qui dépend – syntaxiquement aussi bien que sémantiquement – d'un autre modalisateur épistémique (à savoir la séquence *(ne pas) croire que*) :

- (1) a) Je ne crois pas que Jean soit un bon candidat
- b) Je ne crois pas que Jean est un bon candidat

---

<sup>41</sup> Cette présentation se base notamment sur les ouvrages suivants : Riegel et al. (1998) pour le français, Quirk et al. (1985) pour l'anglais et Faarlund et al. (2002) pour le norvégien.

<sup>42</sup> Certaines langues disposent d'autres types de modalité grammaticale, par exemple de systèmes complexes composés de diverses affixes modaux (voir Palmer 2001). Ce type de modalité grammaticale ne sera pas discuté ici vu qu'il n'est pas pertinent pour les langues étudiées.

<sup>43</sup> L'emploi des termes « neutre » et « non-marqué » peut être contesté, car même l'indicatif peut être considéré comme un mode marqué dans le sens où il marque la factivité.

Un autre exemple du subjonctif comme marqueur épistémico-modal est fourni par les constructions contenant la locution conjonctive *le fait que...* Cette locution est suivie tantôt de l'indicatif, tantôt du subjonctif, et dans les cas où elle est suivie du subjonctif, il ne s'agit peut-être pas après tout d'un *fait* au sens strict, puisque le subjonctif ajoute un certain doute par rapport à l'information exprimée. En tant que locution conjonctive, *le fait que* a un sens assez pauvre, et sa présence s'explique par des règles grammaticales. Le degré de certitude se reflète alors dans l'emploi des modes. Dans ces cas particuliers, on peut dire que le subjonctif fonctionne comme un marqueur épistémico-modal à part entière.

En norvégien, le subjonctif est une catégorie marginale (Faarlund et al. 2002 : 587, 590). Il ne s'emploie que dans quelques expressions figées et souvent archaïques du type *Leve kongen* et *Måtte han leve i hundre år*. Qui plus est, il exprime le plus souvent des souhaits, ce qui le lie plutôt à la modalité déontique ou à la modalité « boulique » qu'à la modalité épistémique.

En anglais, le subjonctif s'utilise un peu plus qu'en norvégien, mais pas en tant que modalisateur épistémique. Comme en norvégien, il apparaît plutôt dans des contextes exprimant une volonté que dans des contextes exprimant une incertitude (Quirk et al. 1985 : 156-158). Il s'emploie parfois dans des contextes hypothétiques (*if I were you...*), ce qui ne veut pas dire qu'il qualifie le degré de certitude du locuteur ; son rôle est tout simplement d'indiquer qu'il s'agit d'une situation imaginée.

Les modes en norvégien et en anglais ne sont pas pertinents pour notre analyse, étant donné qu'aucun d'eux ne réalise la modalité épistémique. La catégorie de mode semble donc être pertinente uniquement pour le français. Cependant, le rôle du subjonctif comme moyen d'exprimer la modalité épistémique est, à en juger d'après Riegel et al. (1998), assez restreint, étant donné qu'il apparaît rarement comme l'unique marque d'une telle modalité.

### *Tiroirs verbaux*

Temporalité et modalité sont deux notions liées. Nous avons vu qu'il n'y a pas de relation biunivoque entre les modes et les modalités. Il n'y a pas de relation biunivoque entre *temps grammatical* ('*tense*') et *temps notionnel* ('*time*') non plus (termes de Le Querler 1996 : 13). Les temps verbaux ne s'emploient pas toujours pour exprimer la temporalité ; ils s'emploient parfois dans un sens strictement modal. Cela est bien connu en ce qui concerne le conditionnel du français, qui connaît plusieurs emplois modaux outre son emploi temporel de futur du passé. Riegel et al. (1998) énumèrent ses valeurs modales à partir des exemples suivants :

- (2) S'il avait de l'argent, il achèterait une Mercedes (ibid. : 309) (marque le potentiel dans un contexte hypothétique)
- (3) Vous devriez approfondir cette question (« conseil poli », ibid. : 319)
- (4) On se croirait au Moyen-Âge (« opinion illusoire », ibid. : 319)
- (5) Elle cherche une théorie qui expliquerait l'univers (« éventualité », ibid. : 319)
- (6) Je serai Robin des Bois. Je dépouillerai les riches pour nourrir les pauvres... (« l'imaginaire », ibid. : 319)
- (7) Une navette spatiale partirait bientôt pour Mars (« information incertaine », ibid. : 320)

Bien que plusieurs de ces emplois expriment des faits hypothétiques et potentiels, c'est uniquement ce dernier, le conditionnel d'information incertaine, qui s'associe directement à la modalité épistémique telle qu'on l'a définie au chapitre 3. D'après nos définitions de l'évidentialité et de l'épistémicité, cet emploi du conditionnel fonctionne avant tout comme un marqueur évidentiel (voir aussi Dendale 1993), mais à la valeur de non-prise en charge vient s'ajouter, dans la plupart des cas, une valeur d'incertitude, dans la mesure où l'information exprimée a le statut d'information non confirmée. C'est donc notamment dans cet emploi que le tiroir du conditionnel fonctionne en tant que modalisateur épistémique<sup>44</sup>.

Le tiroir futur est aussi susceptible d'exprimer la modalité épistémique. Cela vaut notamment pour le futur antérieur, qui peut exprimer une hypothèse probable, plus précisément une déduction qu'a faite le locuteur à partir de certains indices (explicités ou non) :

- (8) Elle revient déjà : elle aura manqué son train (Riegel et al. 1998 : 315)

La valeur épistémico-modale du futur simple est moins évidente. A l'extrême, on pourrait dire que tout emploi du futur est modal, et plus précisément épistémiquement modal, car en se prononçant sur l'avenir, on se prononce sur un monde virtuel dont on ne peut rien savoir au juste. Toute prédiction, aussi certaine qu'elle soit, devient une présomption, et le tiroir du futur français devient conséquemment un marqueur épistémico-modal. Nous n'allons pas poursuivre une telle approche ici, qui nous semble un peu tirée par les cheveux. Le rôle du

---

<sup>44</sup> En norvégien, cet emploi s'exprime à travers l'auxiliaire *skulle* et il sera donc traité dans la section sur les auxiliaires modaux. L'anglais ne semble pas avoir grammaticalisé cet emploi spécifique (Kronning 2005).

tiroir futur est avant tout de situer, tout simplement, un procès dans l'avenir, et non de marquer une incertitude. Bien qu'on puisse maintenir, d'un point de vue épistémologique, que le fait de parler de l'avenir implique une part d'incertitude, cette incertitude n'est pas nécessairement marquée par le langage. Ce n'est pas une fonction prédominante ou explicite du tiroir du futur que d'exprimer l'incertitude.

Pour le norvégien et l'anglais, la question concernant le futur, le conditionnel et leurs emplois temporels et modaux est un peu différente. Grammaticalisés en français, ces temps ne se réalisent pas à travers des désinences en norvégien et en anglais, mais se composent à l'aide d'auxiliaires modaux tels que *skulle/shall* et *ville/will*. Ils seront par conséquent discutés dans la section sur les auxiliaires modaux.

Bien que ce soient en particulier le futur et le conditionnel qui se chargent de sens modaux, tous les temps verbaux sont en effet susceptibles de transmettre des valeurs modales. Dans l'exemple ci-dessous par exemple, le présent exprime un ordre, autrement dit un type de modalité déontique, et il n'a pas forcément de sens temporel :

(9) Tu te tais !

En norvégien :

(10) Nå tier du stille !

Le prétérit norvégien, l'imparfait français et le temps du passé de l'anglais sont aussi susceptibles de véhiculer des sens modaux du type *demandes polies*, *situations hypothétiques* etc :

(11) Var det mulig å forstyrre et øyeblikk? (Faarlund et al. 2002 : 630)

(On pourrait vous déranger une minute?)

(12) Je voulais vous demander un service (Riegel et al. 1998 : 310)

(13) I wondered if you could help us (Quirk et al. 1985 : 188)

(14) Om det brøt ut brann, dekket nok forsikringsselskapet skadene (Faarlund et al. 2002 : 629)

(S'il y avait un incendie, la compagnie d'assurance couvrirait sans doute les dégâts)

(15) S'il avait de l'argent, il achèterait une Mercedes (Riegel et al. 1998 : 309)

(16) If you really worked hard, you would soon get promoted (Quirk et al. 1985 : 188)

Selon Faarlund et al. (2002 : 628), le prétérit norvégien exprime, dans ces emplois, une modalité épistémique. Or, ces emplois ne rentrent pas forcément dans la catégorie de la modalité épistémique telle qu'on l'a définie ici. Les formes au prétérit dans ces exemples n'indiquent pas le degré de certitude du locuteur envers ce qu'il énonce, mais elles sont employées pour atténuer une demande (la rendre plus « polie », ex. 11-13) ou pour parler d'un état de choses imaginé (ex. 14 -16).

En conclusion, les tiroirs verbaux susceptibles d'exprimer la modalité épistémique se restreignent aux tiroirs français du conditionnel et du futur. Les tiroirs norvégien et anglais du passé peuvent marquer l'hypothétique, le potentiel et l'imaginaire, mais ces emplois ne sont pas compris dans notre définition de la modalité épistémique.

### *Auxiliaires modaux*

Dans la recherche portant sur la modalité, ce sont avant tout les auxiliaires modaux qui ont attiré l'attention des chercheurs. Les modaux anglais en particulier ont fait l'objet d'une littérature abondante. Certains les considèrent comme des moyens grammaticaux (Palmer 1986, 2001 ; Andersen 2003 : 43 sqq), alors que d'autres les subsument sous la catégorie des moyens lexicaux (Hyland 1998 ; Lyons 1981 : 238). Comme il a déjà été noté, nous les considérons comme des moyens grammaticaux. Ce choix n'est pourtant pas évident. Certes, les verbes modaux en anglais se distinguent clairement des verbes lexicaux par une série de critères grammaticaux aisément repérables. Par exemple, les modaux n'ont pas de formes non-finies, pas d'impératif et ils ne se conjuguent pas à la troisième personne du singulier (Palmer 1986 : 33-39). Ils se caractérisent aussi par le fait qu'ils ne peuvent pas être combinés entre eux, c'est-à-dire qu'il n'y a pas de *will can come or may shall be* (Palmer 1986 : 33). Mais ces critères ne valent pas pour les auxiliaires modaux en français, car tous possèdent des formes non-finies (infinitif, participe présent, participe passé) et *vouloir* se produit aussi à l'impératif. En outre, il n'y a pas de règles particulières propres à la conjugaison des auxiliaires modaux, et la combinaison de deux modaux est possible : *cela doit pouvoir aller*. Y a-t-il d'autres critères formels, spécifiques au français, qui peuvent distinguer les auxiliaires modaux d'autres verbes ? *Pouvoir* et *devoir* possèdent certaines particularités syntaxiques (incapacité d'être suivi par des complétives (Riegel et al. 1998 : 254), pas de forme impérative), mais *vouloir*, qui est normalement inclus parmi les auxiliaires modaux, ne partage pas ces traits. Selon Riegel et al. (1998 : 252-253), il est en fait difficile de distinguer les auxiliaires modaux d'autres verbes sans avoir recours à des critères sémantiques. La



question de savoir si les auxiliaires modaux en français constituent une véritable classe grammaticale demeure donc un point de discussion.

Les auxiliaires modaux en norvégien semblent se situer, en tant que classe grammaticale, quelque part entre les modaux de l'anglais et ceux du français. Ils possèdent des infinitifs et des participes passés, mais pas de participe présent et pas d'impératif. Bien qu'ils possèdent des participes passés, ils ne s'emploient pas au passif, à l'exception de *kunne* dans son sens dynamique (voir Faarlund et al. 2002 : 526). Comme les verbes en norvégien ne se conjuguent pas en personne, ce critère n'est pas applicable tel qu'il est formulé pour l'anglais. Néanmoins, les modaux prototypiques du norvégien se conjuguent selon un paradigme flexionnel qui diffère du paradigme d'autres verbes norvégiens. Ce paradigme est appelé *preterito-presentisk*<sup>45</sup> (Faarlund et al. 2002 : 491) et se caractérise entre autre par l'absence du *r* final que la grande majorité des verbes en norvégien reçoivent au présent. Cependant, ce paradigme flexionnel ne vaut pas pour tous les auxiliaires normalement considérés comme modaux. Une exception est le verbe *burde*, qui est indiscutablement modal, mais dont la conjugaison ne suit pas ce paradigme particulier (Faarlund et al. 2002 : 526). Et dans le groupe de verbes qui se conjuguent d'après le paradigme *preterito-presentisk*, on trouve le verbe *vite*, qui n'est pas un auxiliaire modal. Alors que le critère d'un paradigme de conjugaison spécifique s'applique aux auxiliaires modaux norvégiens (bien que ce ne soit pas de façon infaillible), le critère de non combinaison ne s'applique pas du tout. Les auxiliaires modaux en norvégien se combinent très fréquemment entre eux : *vil kunne komme, skal kunne gå osv.*

Une autre caractéristique des auxiliaires modaux est qu'ils se combinent uniquement avec l'infinitif *pur*, (c'est-à-dire sans les marques d'infinitif *de, à* pour le français, *å* pour le norvégien, et *to* pour l'anglais) : *il doit venir, det kan være slik*. Or, il existe en français toute une série d'autres verbes qui se construisent avec l'infinitif pur ; il s'agit toutefois souvent de verbes qui ont un contenu modal : *espérer, oser, savoir, penser*. Pour le norvégien, ce critère nous donnerait – outre *ville, skulle, burde* et *måtte – få* et *la* comme auxiliaires modaux. Mais ces derniers ne possèdent pas les autres caractéristiques des auxiliaires modaux : ils se conjuguent - comme le fait *burde* - avec des *r* finals au présent et les deux se mettent à l'impératif. Il existe aussi d'autres verbes qui peuvent se combiner avec l'infinitif pur et qui, dans ces emplois, ressemblent aux modaux : *hva tenker du gjøre ? ('qu'est-ce que tu penses*

---

<sup>45</sup> Le paradigme du 'prétérit-présent' (notre traduction). L'appellation renvoie au fait que les formes au présent de ces verbes se sont développées d'une forme ancienne au prétérit.

*faire ?*), *du trenger ikke gjøre det* ('tu n'a pas besoin de le faire') (voir Faarlund et al. 2002 : 530).

Pour résumer, nous avons vu que les auxiliaires modaux de l'anglais se laissent aisément définir par une série de critères formels. En ce qui concerne les modaux en français et en norvégien en revanche, les critères qui les distinguent d'autres verbes sont moins clairs et moins nombreux que ceux caractérisant les modaux anglais. Les modaux norvégiens se caractérisent, comme nous l'avons vu, par leur paradigme de conjugaison, qui est différent de celui de la grande majorité d'autres verbes, ainsi que par leur incapacité d'être mis au passif et par leur absence de formes impératives et de participes présents. Quant au français, on distingue difficilement les auxiliaires modaux d'autres verbes sans avoir recours à des critères sémantiques (cf. Riegel et al. 1998 : 252-253). Ils possèdent pourtant certaines particularités syntaxiques (construction avec l'infinitif pur), mais celles-ci ne sont pas exclusives à ce groupe de verbes.

Ce petit exposé nous montre à quel point il est difficile de ranger les éléments modalisants dans les catégories « moyens grammaticaux » et « moyens lexicaux ». Cette difficulté s'explique par le fait que la grammaticalisation est un développement progressif par degrés (Hopper et Traugott 1993). Alors que *pouvoir* et *devoir* se sont grammaticalisés dans une large mesure, *vouloir* ne s'est pas grammaticalisé au même point, ce qui fait que les auxiliaires modaux ne se laissent pas définir de façon globale par une série de critères. Faute de critères formels précis, peut-être serait-il plus approprié de considérer les auxiliaires modaux français comme des moyens lexicaux ? En tant que correspondants des auxiliaires modaux anglais et norvégiens, nous avons tout de même choisi de les traiter tous comme des moyens grammaticaux. Ce que cette problématique nous montre, c'est avant tout à quel point il est difficile, voire pas toujours fructueux, de prendre comme point de départ pour des études contrastives la modalité dans un sens strictement grammatical.

Compte tenu des difficultés qu'il y a à circonscrire une catégorie bien définie d'auxiliaires modaux, notre choix d'éléments à inclure dans cette petite présentation des moyens linguistiques susceptibles d'exprimer la modalité épistémique sera inévitablement un peu arbitraire. Nous nous contenterons ici de considérer les auxiliaires modaux traditionnellement reconnus comme tels, plus précisément *pouvoir*, *devoir* et (*vouloir*) pour le français, *kunne*, *måtte*, *burde*, *skulle* et *ville* pour le norvégien, et *may/might*, *can/could*, *must*, (*shall*)/*should* et *will/would* en anglais. *Vouloir* et *shall* ne possèdent pas de valeur épistémique et ils sont inclus uniquement pour des raisons de comparaison.

– *pouvoir, kunne, can/could, may/might*<sup>46</sup>

Ces verbes expriment tous diverses formes de possibilité. Ils ont en commun le fait qu'ils sont, chacun dans sa langue, l'auxiliaire modal employé pour dénoter la possibilité. On notera qu'on a deux verbes en anglais, à savoir *can* et *may* là où on n'a qu'un verbe en français (*pouvoir*) et un en norvégien (*kunne*). De ces deux auxiliaires anglais, c'est le verbe *may* qui est le plus important en ce qui concerne le marquage épistémique. L'emploi épistémique de *can* est une particularité des formes interrogative et négative ; *can* assertorique ne s'utilise pas dans le sens épistémique (Hyland 1998 : 109), mais est restreint aux valeurs radicales. Cette restriction ne vaut pourtant pas pour la forme au passé, *could*, qui s'emploie dans le sens épistémique aussi bien que dans les valeurs radicales. *May* de son côté est, comme *pouvoir* et *kunne*, susceptible de véhiculer tous les types de modalité répertoriés au chapitre 3, bien que *may* à la valeur dynamique semble rare.

Comme mentionné au chapitre 3, l'emploi épistémique de *pouvoir, kunne, could, may* et *might* dénote une possibilité éventuelle ; une hypothèse :

(17) Hun kan være syk

(18) Elle peut être malade

(19) She may be ill

En dépit de la polysémie inhérente de *may*, plusieurs grammaires reconnaissent la valeur épistémique comme la valeur principale de ce verbe (Quirk et al. 1985 : 223, Coates 1983 : 26). En ce qui concerne l'auxiliaire norvégien, *kunne*, la valeur épistémique est répertoriée dans les grammaires comme une des valeurs principales (voir par exemple Faarlund et al. 2002). Quant à leur correspondant en français, Le Querler (2001 : 22) note que l'emploi épistémique de *pouvoir* est relativement rare. Ce sens est néanmoins répertorié parmi les valeurs principales de ce verbe (Sueur 1979 ; Le Querler 1996).

En ce qui concerne l'anglais et le norvégien, l'emploi de la forme passée est censé mener à une atténuation plus forte du degré de certitude, c'est-à-dire que la possibilité est conçue comme moins probable (Andersen 2003 : 67, Quirk et al. 1985 : 223)

(20) Det kunne være at du har rett

(21) It might be that you are right

---

<sup>46</sup> Les formes *might* et *could* fonctionnent tantôt comme les formes passées de *may* et de *can* et tantôt comme des auxiliaires modaux à part entière (Coates 1983).

Ce point de vue est pourtant contesté par Coates (1983 : 146, 147, 149) qui ne voit pas de différence de degré de certitude entre *may* et *might* et maintient que ces deux formes sont interchangeable dans la plupart des cas.

– *devoir, måtte, skulle, burde, must, (shall)*

*Devoir, måtte* et *must* expriment différentes formes de nécessité dont la nécessité déontique (l'obligation) et la nécessité épistémique sont les deux formes principales. Dans la valeur épistémique, *devoir, måtte* et *must* expriment une déduction de la part du locuteur et un haut degré de probabilité (Dendale 1994, Faarlund et al. 2002 : 599, Quirk et al. 1985 : 225) :

(22) Du må være utslitt etter turen (Faarlund et al. 2002 : 600)

(Tu dois être épuisé après le voyage)

(23) You must be feeling tired (Quirk et al. 1985 : 224)

(24) Jean doit être fatigué parce qu'il a beaucoup travaillé (Dendale 1994 : 30)

*Devoir* en français possède en outre une valeur de futur. Dans ce sens, il correspond au verbe *skulle* en norvégien :

(25) Quand est-ce que Paul doit partir pour Paris ? (Kronning 2001a : 75)

Når skal Paul reise til Paris?

Dans ce dernier emploi, *devoir* et *skulle* n'ont pas de valeur épistémique, mais *skulle* peut prendre un sens épistémique dans d'autres cas. Ce verbe s'emploie par exemple comme marqueur évidentio-modal indiquant l'emprunt. Il correspond alors au conditionnel de citation du français (cf. section 3.2.) :

(26) Hun skal være svært flink (Faarlund et al. 2002 : 605)

(Elle serait très douée)

La forme au prétérit s'emploie en outre pour indiquer un haut degré de probabilité ou une sorte d'inférence (c'est-à-dire une valeur épistémico-modale) :

(27) De skulle være hjemme på denne tiden (Faarlund et al. 2002 : 610)

(Ils devraient<sub>E</sub> être à la maison à cette heure-ci)

Il en va de même pour le verbe anglais *should* (la forme au présent, *shall*, par contre, ne s'utilise pas avec une valeur épistémique) :

(28) The mountains should be visible from here (Quirk et al. 1985 : 227)

Dans le sens de 27, *skulle* est interchangeable avec *burde*, qui lui aussi est susceptible de prendre un sens épistémique, bien que le sens principal soit une valeur déontique de recommandation ou de conseil. Dans les énoncés épistémiques, *burde* garde souvent un sens déontique secondaire (Faarlund et al. 2002 : 610).

– *vouloir, ville, will/would*

Le norvégien et l'anglais disposent d'un autre auxiliaire susceptible de véhiculer un sens épistémico-modal, à savoir *ville* et *will /would*, respectivement :

(29) That will be the postman (Quirk et al. 1985 : 228)

(30) Mange vil nok ha sett henne på TV-skjermen (Faarlund et al. 2002 : 617)<sup>47</sup>

(Beaucoup l'auront (probablement) vue à la télé)

Le correspondant en français, *vouloir*, ne possède pas de valeur proprement épistémique. Pour un équivalent en français d'un énoncé tel que 29, on utiliserait plutôt d'autres constructions ou d'autres marqueurs, par exemple le verbe *devoir* : *Ça doit être le facteur*. Il convient de noter que la valeur épistémique de *ville* et de *will* n'est pas une valeur principale de ces verbes. Ils possèdent d'autres valeurs qui sont beaucoup plus fréquentes ; par exemple l'auxiliaire norvégien prend fréquemment un sens déontique (ou « boulique ») de volonté, auquel cas il correspond au verbe '*want*' en anglais. En outre, ils sont couramment employés avec l'infinitif d'un verbe lexical pour composer le temps du futur.

---

<sup>47</sup> La valeur épistémique est ici renforcée par la particule modale *nok* ('probablement'). Cette particule n'est pas indispensable pour garder une lecture épistémique de l'énoncé, mais il semble que *ville* épistémique est souvent accompagné d'autres marqueurs servant à renforcer le sens épistémique.

### 3.3.2.2 Moyens lexicaux

Selon notre définition de la modalité épistémique, tout lexème indiquant le degré de certitude du locuteur à l'égard de ce qu'il affirme est un marqueur épistémico-modal. Il peut s'agir d'éléments relevant de diverses parties de discours, et il semble y avoir un haut degré de correspondance entre les trois langues<sup>48</sup>. Les moyens lexicaux qui expriment la modalité épistémique incluent les sous-types suivants, sans s'y limiter :

– *Semi-auxiliaires*

exemples : français : *sembler, paraître* ; anglais : *seem, appear* ; norvégien : *synes, se ut (til/som), virke (som)*

– *Verbes épistémiques*

exemples : français : *croire, supposer* ; anglais : *believe, assume* ; norvégien : *tro, anta*

– *Adjectifs, adverbes et noms épistémiques*

exemples : français : *possible, vraisemblablement, possibilité* ; anglais : *possible, likely, possibility* ; norvégien : *mulig, trolig, mulighet*

En norvégien, il existe aussi un certain nombre d'adverbes de phrase monosyllabiques qui sont relativement fréquents, en particulier dans la langue parlée (Andersen 2003 : 45), mais qui n'ont pas forcément d'équivalents précis en français ou en anglais. Ces petits adverbes sont souvent appelés des « particules modales » (« modalpartikler ») et les exemples types sont : *nok, vel, jo* et *visst*. En indiquant que l'information transmise est ou devrait être déjà connue, *jo* se situe sur le côté certitude, et ce sont donc *nok, vel* et *visst* qui pourraient être pertinents pour la présente étude. En tant que particules modales<sup>49</sup>, ils atténuent l'énoncé et expriment une certaine réserve de la part du locuteur à propos de la vérité de l'information véhiculée (voir Fretheim 1981, 1988). Le marqueur *visst* indique que le locuteur tire une conclusion à partir de certains indices (Faarlund et al 2002 : 826), et il est donc un marqueur évidentio-modal :

---

<sup>48</sup> C'est-à-dire que les trois langues semblent à peu près disposer des mêmes unités sémantiques. Pour la plupart des lexèmes, il existe dans les autres langues des correspondants ou des quasi-équivalents relevant de la même partie du discours. Il reste à voir dans quelle mesure l'usage effectif des différents lexèmes coïncident.

<sup>49</sup> Il est à noter que ces mots sont, comme tant d'autres marqueurs susceptibles d'exprimer la modalité épistémique, largement polysémiques. Ainsi, *nok* fonctionne aussi comme un adverbe de quantité : *Han har nok penger (il a suffisamment d'argent)* et *vel* fonctionne entre autre comme un attribut du sujet dans le sens de *bien* : *hun føler seg ikke vel (elle ne se sent pas bien)*.

(31) Han er nok heime no (Faarlund et al. 2002 : 825)

(Il est probablement chez lui maintenant)

(32) Han er vel ikkje heime no (Faarlund et al. 2002 : 825)

(Il n'est pas chez lui maintenant, je suppose)

(33) Han har visst gått ut (Faarlund et al. 2002 : 826)

(Il paraît qu'il est sorti)

C'est peut-être par rapport aux particules modales que le niveau de correspondance entre les trois langues semble être le plus bas. Pour rendre le même contenu sémantique en français ou en anglais, on doit avoir recours à d'autres constructions, lesquelles varieront selon les cas et les cotextes.

Le domaine de la modalité épistémique englobe toute une série de notions sémantiques, par exemple la certitude, la probabilité, la possibilité, les présuppositions, les déductions, les inférences, le doute, les impressions etc. Comme il a déjà été mentionné, les marqueurs de certitude sont exclus de l'étude, et les autres notions se ramènent à divers types de possibilité. Néanmoins, ces différentes notions sémantiques se réalisent linguistiquement par toute une gamme de moyens lexicaux qui n'ont pas nécessairement de similitude avec des mots comme *possibilité* et *probabilité*. A titre d'illustration, la figure ci-dessous donne quelques exemples de moyens lexicaux français susceptibles d'exprimer la modalité épistémique, allant de la certitude définitive à l'égard d'une proposition *p* en passant par l'incertitude totale à la certitude de non *p*.

l'évident	il est évident que	évidemment	naturellement
	il va sans dire que	il va de soi que	
certitude	je suis sûr que	je suis certain que	on sait que cela prouve que
	c'est un fait que	sans aucun doute	sans conteste cela montre que
conviction	je suis convaincu /persuadé que		
probabilité	certainement		
	vraisemblablement	probablement	
	sans doute	cela indique que	il est probable que
	je crois que	il paraît que	
supposition	je suppose que		
impression	il (me) semble que	j'ai l'impression que	apparemment
possibilité	il est possible que	possiblement	peut-être que
	l'hypothèse que	éventuellement	on ne sait pas si
doute	je doute que	je ne crois pas que	il est peu probable que
	il est invraisemblable que		il est contestable que

Figure 4 : Moyens lexicaux français appartenant au domaine modal de l'épistémique

### 3.3.3 La combinaison de moyens grammaticaux et lexicaux

Plusieurs chercheurs ont observé que, dans le discours scientifique, les atténuateurs s'accumulent souvent en ce qu'on peut appeler, avec Salager-Meyer (1994 : 154-155) 'compound hedges' (atténuateurs composés) ou, avec Halliday (cité dans Palmer 2001 : 35) 'harmonic combinations' (combinaisons harmoniques). Deux ou plusieurs marqueurs relevant du même domaine modal concourent à la transmission d'un certain sens modal. Il s'agit souvent de cas où un verbe lexical épistémique se voit modifié par un auxiliaire de possibilité, comme dans *disse tallene kan tyde på...*<sup>51</sup> (voir Breivega 2003 : 153) ou *it may suggest that...* (Salager-Meyer 1994 : 154-155). Dans ces cas, la combinaison est grammaticalement facultative : l'un ou l'autre des marqueurs est suffisant pour exprimer la modalité envisagée, mais l'ajout d'un deuxième marqueur sert à renforcer la modalité

<sup>50</sup> Le choix des termes *vérifié* et *réfuté* peut être mis en discussion. Il convient de noter que ces concepts sont ici à prendre dans le sens de « montré ou présenté comme vérifié ou réfuté ». Il ne s'agit pas nécessairement d'énoncés exprimant des faits effectivement vérifiés ou réfutés, mais présentés comme tels par le locuteur.

<sup>51</sup> Traduction française : *ces chiffres peuvent suggérer que...*



exprimée, atténuant plus encore l’assertion. Le fait d’omettre le marqueur grammatical ou de remplacer le marqueur lexical par un marqueur plus « neutre » ou exprimant un degré plus haut de certitude changerait l’interprétation de l’énoncé : cela lui donnerait une lecture plus catégorique. Les tournures *this suggests that...* et *this may show that...* sont ainsi toutes les deux plus fortes que *this may suggest that...*

En français, ces combinaisons harmoniques s’inscrivent parfois dans les règles grammaticales mêmes. Il s’agit de constructions (relativement restreintes en nombre) où le degré de certitude exprimé par un marqueur lexical se reflète dans l’emploi des modes. Ainsi, dans *il est fort probable qu’il viendra*, le marqueur lexical *fort probable* marque un haut degré de certitude à l’égard de la vérité de la proposition (*p*) *il vient*. Le mode de l’indicatif reflète cette certitude à l’égard de la vérité de *p*. Dans *il est peu probable qu’il vienne*, le marqueur lexical *peu probable* marque un doute fort par rapport à la vérité de *p*, un doute qui se reflète dans le mode du subjonctif. L’emploi de l’un ou l’autre mode n’est pas optionnel ici, mais gouverné par des règles grammaticales<sup>52</sup>.

En ce qui concerne le discours scientifique anglais, ce phénomène d’accumulation d’atténuateurs a été observé plusieurs fois (voir par exemple Salager-Meyer 1994 et Hyland 1998). Breivega (2003 : 153) a fait la même observation pour le discours scientifique norvégien. Au chapitre 7, nous examinerons dans quelle mesure ce phénomène se fait valoir dans nos données norvégiennes et françaises.

### 3.3.4 Bilan comparatif

Qu’est-ce qu’on peut retenir de ce bref survol des différents moyens d’exprimer la modalité épistémique en français, norvégien et anglais? Comme nous l’avons vu, les trois langues disposent dans une large mesure des mêmes moyens linguistiques. On verra dans les chapitres à suivre ce qu’il en est de l’usage effectif. Certains traits spécifiques se dessinent cependant : la catégorie morphologique de *mode* ne semble être pertinente que pour l’analyse du français (mais même en français son rôle paraît limité). De plus, c’est uniquement en français que la modalité épistémique peut être exprimée à travers des tiroirs dits temporels. On peut en tirer la conclusion que le français utilise des moyens morphologiques pour exprimer la modalité épistémique dans une plus large mesure que ne le font l’anglais et le norvégien. En revanche, il semble que le norvégien et l’anglais font un usage plus étendu d’auxiliaires modaux, étant

---

<sup>52</sup> Si les grammaires recommandent l’usage de l’indicatif après *il est probable que* (voir par exemple Pedersen et al. 1992), on constate toutefois que l’emploi du subjonctif se rencontre relativement souvent dans le français moderne.

donné que les auxiliaires susceptibles de prendre une valeur épistémico-modale sont plus nombreux en norvégien et en anglais. En outre, la valeur épistémique de *pouvoir* serait, selon Le Querler (2001 : 22), relativement rare. En ce qui concerne les moyens lexicaux, ceux-ci sont nombreux et variés dans les trois langues. Le norvégien dispose d'une catégorie de modalisateurs dont il est difficile de trouver des équivalents directs en français et en anglais, à savoir les particules modales du type *vel, nok, visst*. Ces particules sont supposées typiques de la langue parlée (Andersen 2003 : 45), et il reste à voir si elles sont courantes dans le discours scientifique écrit.

Dans ce chapitre, nous avons tenté de ranger les divers marqueurs épistémico-modaux en groupes selon leurs caractéristiques formelles. Or, on notera que ces caractéristiques formelles vont souvent de pair avec des traits sémantiques spécifiques. Nous reviendrons sur cette question à la section 6.4.2.2.

## 4 RECHERCHES ANTÉRIEURES SUR LE PHÉNOMÈNE D'ATTÉNUATION

### 4.1 Le concept d'*atténuation*<sup>1</sup> ('*hedging*')

Le terme *atténuation* a souffert, et souffre toujours dans une certaine mesure, de la même absence de consensus et de clarification conceptuelle que la notion de modalité. Cela implique qu'il faut faire preuve de prudence en rapportant et interprétant des résultats antérieurs ; ceux-ci ne sont pas nécessairement comparables, ni entre eux, ni avec les nôtres, puisque on ne parle pas toujours de la même chose. Au cours de cette section, qui donne un bref aperçu du développement du terme '*hedging*' et de certaines des conceptions le plus souvent rencontrées dans la littérature<sup>2</sup>, nous verrons plusieurs exemples de cette confusion définitionnelle. Comme il existe un vaste ensemble de recherches qui abordent le phénomène de '*hedging*', il n'est pas possible ici de retenir davantage qu'une sélection des travaux les plus importants.

Dans l'optique adoptée ici, les modalisateurs épistémiques d'incertitude constituent un type d'*atténuateurs* (cf. *supra* 1.2), qui à leur tour se définissent comme des marqueurs servant à amoindrir la force de ce qui est dit (Foullioux et Tejedor de Felipe 2004 : 114). Si le phénomène d'atténuation a attiré relativement peu d'attention de la part de chercheurs situés dans la tradition linguistique française<sup>3</sup>, la situation est toute autre dans la tradition anglo-saxonne. Dans la linguistique anglo-saxonne, on s'intéresse à ce phénomène depuis au moins trois décennies ; la littérature est abondante sur le sujet et elle continue de croître. C'est un article de Lakoff, publié en 1972, qui éveille initialement l'intérêt pour le phénomène d'atténuation et qui marque à bien des égards le début de ce champ de recherche en linguistique. Pour Lakoff (1972), les atténuateurs ('*hedges*') sont des marqueurs qui modifient l'appartenance d'un prédicat ou d'un syntagme nominal à une catégorie particulière (voir

---

<sup>1</sup> La notion d'*atténuation* est ici à prendre comme le correspondant du terme anglais '*hedging*'. Bien qu'il soit difficile de trouver en français des équivalents parfaits des termes anglais '*hedges*' et '*hedging*', notions si bien établies dans la linguistique anglo-américaine, les termes *atténuateurs* et *atténuation* semblent être parmi les meilleurs candidats. Ce sont les termes proposés par le glossaire français-anglais de terminologie linguistique de SIL (<http://www.sil.org/linguistics/glossary%5Ffe>) et les termes utilisés dans *Langue française* 142, un numéro spécial consacré aux études de ce phénomène. Le glossaire en ligne du SIL propose aussi le terme *adoucisateurs* pour traduire '*hedges*'. Dans la présente thèse seront employés de façon interchangeable les termes *atténuateurs* et *adoucisateurs*.

<sup>2</sup> Pour un compte rendu plus complet du terme '*hedging*' et de ses emplois, voir Varttala 2001.

<sup>3</sup> C'est-à-dire l'atténuation comme phénomène en général. Il existe, bien entendu, beaucoup d'études portant sur des phénomènes rattachables au domaine de l'atténuation ; nous pensons par exemple aux études de Nølke portant sur des adverbes comme *peut-être* (Nølke 1993 : 145-180) et à diverses études concernant les verbes modaux (voir par exemple Dendale et Auwera 2001). Or, l'atténuation en tant que catégorie principale a reçu peu d'intérêt ; une exception est le numéro spécial mentionné ci-dessus de la revue *Langue française*, paru en 2004 et consacré, dans sa totalité, au phénomène d'atténuation.

Markkanen et Schröder 1997 : 5). Par exemple, dans l'énoncé *A penguin is sort of a bird*, le marqueur *sort of* modifie l'appartenance de *penguin* à la catégorie d'oiseaux, en indiquant que cette appartenance n'est pas indiscutable (Markkanen et Schröder 1997 : 3). Chez Lakoff, il s'agit donc d'expressions comme *sort of, rather, in a manner of speaking* etc., autrement dit d'expressions qui marquent que l'appartenance d'un certain objet à une certaine catégorie n'est pas une appartenance à part entière. Le sens originel de la notion d'atténuation, telle qu'elle a été définie par Lakoff, a changé, et le champ de recherche souffre depuis d'une absence de définition précise unanimement acceptée.

Il est clair que la notion d'atténuation s'est élargie depuis son introduction par Lakoff en 1972<sup>4</sup>. Cet élargissement a mené à une dichotomie des adoucisseurs en ce qu'on appelle en anglais *approximators* ('approximateurs') et *shields* (littéralement 'boucliers') (voir Markkanen et Schröder 1997 : 5). Le terme *approximateurs* renvoie à des expressions du type de celles dont parlait Lakoff (*sort of, rather*), mais aussi à certains adverbes de quantité tels que *roughly, about*, alors que le terme *shields* réfère à des expressions qui reflètent le degré de certitude du locuteur par rapport au contenu de son message. Cette dernière catégorie correspond donc à peu près à ce que nous appelons les modalisateurs épistémiques d'incertitude, c'est-à-dire des unités lexicales ou grammaticales qui qualifient explicitement la valeur de vérité d'un contenu informatif, en mettant en doute la vérité de ce contenu. Selon Prince, Frader et Bosk (1982), les *approximateurs* font partie de la proposition et concernent les *conditions de vérité* de celle-ci, alors que les *shields* qualifient une proposition entière. Ce sont uniquement les expressions de ce deuxième groupe qui rentrent dans la catégorie de la modalité épistémique telle que nous l'avons définie dans la présente étude.

Le concept d'atténuation n'a pas seulement été élargi jusqu'à inclure la catégorie des *shields*, il inclut également, chez certains, diverses structures syntaxiques telles que le passif sans agent (Clyne 1991 ; Salager-Meyer et Defives 1998 ; Markkanen et Schröder 1997 : 6) et des structures impersonnelles telles que *it is remarkable/obvious that...* (Clyne 1991 : 58) ainsi que les expressions de jugements de valeurs (*interestingly, surprisingly*) (Skelton 1988 ; Salager-Meyer 1994 ; Salager-Meyer et Defives 1998) et l'emploi (voire le non-emploi) de certains pronoms personnels (voir Markkanen et Schröder 1997 : 6).

Le nombre et le type de moyens inclus dans la catégorie d'atténuation varient donc d'un chercheur à l'autre. Salager-Meyer (1994) par exemple, inclut dans sa taxonomie

---

<sup>4</sup> Il est, bien entendu, difficile de déterminer qui est l'inventeur du terme – Toulmin parlait de '*hedges*' déjà en 1958 (Toulmin 1958) – mais on s'accorde généralement à attribuer l'introduction du terme dans la linguistique à Lakoff.

d'adoucisseurs les cinq catégories suivantes : les *shields* (exemples : *seem, probably*) – dans notre terminologie il s'agirait des marqueurs épistémico-modaux d'incertitude – , les approximateurs (exemples : *roughly, often*), les expressions exprimant le doute personnel et l'engagement direct de l'auteur (exemples : *it is our view that..., to our knowledge...*), les expressions de jugements de valeurs, qui sont des expressions évaluatives ou des jugements affectifs qui font porter les émotions ou les réactions de l'auteur, et qui sont appelées *emotionally charged intensifiers* dans la terminologie de Salager-Meyer (exemples : *surprisingly, extremely interesting*) et, finalement, les adoucisseurs composés ('*compound hedges*'), c'est-à-dire la juxtaposition de deux ou plusieurs adoucisseurs (exemples : *it seems reasonable to assume..., it may suggest that...*)<sup>5</sup>.

Dans une étude ultérieure, Salager-Meyer et Defives (1998) proposent une taxonomie légèrement différente, comprenant les catégories des *shields*, des approximateurs, des jugements de valeur et des constructions passives sans agent. Salager-Meyer et Defives (1998 : 135) définissent l'atténuation comme « the handling of potential fact », mais on voit difficilement comment les jugements de valeur rentrent dans une telle définition<sup>6</sup>. Salager-Meyer et Defives les ont peut-être inclus en suivant Skelton (1988), qui propose de substituer la notion de *commentative language* à celle de *hedging*. Il argumente en faveur d'une distinction entre *commentaire* et *proposition* où la proposition dénote ce qui est dit et le commentaire reflète ce que le locuteur pense de ce qu'il dit<sup>7</sup>. Les commentaires peuvent être du type jugement de vérité ou du type jugement de valeur<sup>8</sup>. Le désavantage d'une telle solution est que le terme *hedging* serait très large et non concordant avec la façon dont il est normalement employé. C'est peut-être pour cela que Skelton, dans un article plus récent (1997 : 125), restreint le terme *hedge* à ce qu'il appelle « deintensified truth-judgements » ('jugements amoindrissant de vérité').

Les jugements affectifs pourraient également entrer dans la catégorie d'atténuation si l'on accepte la définition proposée par Markkanen et Schröder (1997 : 5), dans la mesure où ils modifient l'importance de l'information donnée et/ou l'attitude du locuteur à l'égard de cette information :

---

<sup>5</sup> Tous les exemples sont ceux de Salager-Meyer (1994).

<sup>6</sup> Il n'est pas évident non plus de déterminer en quoi les constructions passives rentrent dans la catégorie, mais leur inclusion s'explique par la non prise-en-charge que le choix du passif peut impliquer : le locuteur se distancie en quelque sorte de l'information rapportée, il n'en prend pas la responsabilité (cf. Salager-Meyer et Defives 1998 : 137-138).

<sup>7</sup> On reconnaît ici la distinction *modus – dictum* (cf. *infra*, la section 3.1).

<sup>8</sup> Autrement dit le « discours commentatif » inclurait les modalités épistémiques et les modalités axiologiques.

[Hedges are seen] as modifiers of the writer's responsibility for the truth value of the propositions expressed or as modifiers of the weightiness of the information given, or the attitude of the writer to the information. (Markkanen et Schröder 1997 : 5).

On voit que Markkanen et Schröder adoptent un point de départ purement fonctionnel, qui permet d'inclure n'importe quelle expression langagière dans la catégorie d'adoucisseurs, pourvu que l'interprétation en contexte le justifie. Markkanen et Schröder admettent qu'une telle définition fonctionnelle mène à une conception extrêmement large (ibid. : 6). A notre avis, ce n'est pas le simple fait d'opter pour un point de départ purement fonctionnel qui mène à un tel élargissement, mais plutôt le fait de regrouper au moins trois fonctions sous le même concept : les atténuateurs sont en premier lieu des modificateurs de la prise en charge par le locuteur de la valeur de vérité de ce qu'il dit, en second lieu ils sont des modificateurs de l'importance de l'information donnée, et en troisième lieu, ils sont des modificateurs de l'attitude du locuteur à l'égard de cette information. Si ces fonctions s'entremêlent et se chevauchent dans bien des cas, elles restent quand même trois types distincts. Le problème avec une telle conception est qu'il est difficile de voir quelles sont exactement les caractéristiques qui relient et réunissent les différentes instances d'atténuation ; quel est le trait commun qui justifie de les subsumer sous la même catégorie ?

Alors que *hedging* devient, chez certains, une sorte d'expression passe-partout, d'autres cherchent à délimiter le concept, souvent en l'assimilant à celui de modalité épistémique. Dans un article critique, Crompton (1997) déplore le fait que les conceptions soient à ce point divergentes et il critique les recherches antérieures pour ne pas avoir été capables de fournir de description précise de ce que c'est qu'un *hedge*. Selon Crompton, les définitions proposées sont peu précises et l'extension de la notion n'est pas très bien déterminée. Les remarques critiques de Crompton sont justifiées, mais il faut noter que l'état de choses est plus clair aujourd'hui qu'au temps où Crompton a rédigé son article. Ce progrès est largement dû à la parution en 1998 de la monographie de Hyland, *Hedging in Scientific Research Articles*, qui est rapidement devenu un ouvrage de référence dans le domaine (voir ci-dessous). Dans son article, Crompton propose une nouvelle taxonomie d'adoucisseurs ainsi qu'une définition précise, mais restreinte (inspirée par la définition originelle que donne Lyons de la modalité épistémique), selon laquelle « a hedge is an item of language which a speaker uses to explicitly qualify his/her lack of commitment to the truth of a proposition he/she utters » (Crompton 1997 : 281). Cette définition de la notion d'atténuation peut paraître prometteuse, mais vu qu'elle revient à celle de la modalité épistémique, on n'aura pas

de terme générique susceptible de réunir les modalisateurs épistémiques et les autres types de marqueurs employés pour faire des réserves.

Hyland (1998) lie aussi le concept d'atténuation à celui de modalité épistémique, mais sa conception est tout de même beaucoup plus large que celle de Crompton<sup>9</sup>. Hyland inclut dans sa taxonomie les auxiliaires modaux à valeur épistémique, les verbes lexicaux épistémiques, les adjectifs, adverbes et noms épistémiques, les approximateurs et certaines stratégies discursives telles que les références à des connaissances restreintes, à des limites méthodologiques etc. En outre, il mentionne (bien que brièvement) le rôle de certaines structures syntaxiques telles que le passif et l'interrogation (ibid. : 5, 45, 77). A la différence de Salager-Meyer, il exclut explicitement les jugements de valeur (ibid. : 5). Comme mentionné plus haut, l'ouvrage de Hyland (1998) est vite devenu un ouvrage de référence pour ceux qui travaillent sur l'atténuation, et grâce à cette publication, il existe probablement un plus grand terrain d'entente aujourd'hui qu'il y a dix ans.

Comme déjà évoqué, pour Hyland (ibid. : 3, 5), l'atténuation est un type de modalité épistémique. Plus précisément, l'atténuation équivaut pour lui à ce qui relève de l'incertitude sur l'échelle épistémico-modale. Cependant, il inclut de nombreux éléments linguistiques qui n'entrent pas aisément dans sa définition de la modalité épistémique<sup>10</sup>, par exemple certains adverbes comme *about* et *around* :

(1) ... which accounted for about 4 % of the total bases present (Hyland 1998 : 140)<sup>11</sup>

et certains de ses exemples de stratégies discursives d'atténuation :

(2) We did not succeed in obtaining the complete transcript (ibid. : 147)

A notre avis, la relation entre les deux catégories est plutôt inverse : les modalisateurs épistémiques (côté incertitude) constituent un type d'atténuation, mais la notion d'atténuation englobe en outre plusieurs autres catégories, qui ne se laissent pas ramener à la modalité épistémique : il pourrait par exemple s'agir d'approximateurs (comme '*about*' dans le premier exemple tiré de Hyland 1998), de manifestations linguistiques des limites de la recherche (comme dans le deuxième exemple tiré de Hyland 1998) ainsi que de marqueurs de

---

<sup>9</sup> Ce qui suggère que ces deux chercheurs ont des conceptions différentes de ce qu'est la modalité épistémique.

<sup>10</sup> « Epistemic modality expresses the speaker's opinion or belief concerning the truth of what is said » (Hyland 1998 : 44).

<sup>11</sup> Le soulignement dans cet exemple et dans les exemples à suivre est de Hyland.

sporadicité tels que *rarement*<sup>12</sup> et de certains adverbes modificateurs (*assez, relativement, généralement*)<sup>13</sup>

Il n'y a pas lieu ici de discuter et d'évaluer l'ensemble des emplois du terme *hedging* qui ont été avérés au cours des années – certains sont plus ou moins absolètes et leurs auteurs semblent dans certains cas avoir changé de point de vue (nous pensons par exemple à Skelton, qui définit *hedge* de deux manières différentes dans ses articles de 1988 et de 1997). Néanmoins, nous sommes d'accord avec Crompton sur le fait qu'il faut aspirer à une définition qui soit aussi claire que possible. Et bien qu'il existe probablement un plus grand terrain d'entente aujourd'hui qu'il y a dix ans, on trouve toujours certaines incohérences dans les travaux portant sur l'atténuation. Par exemple, d'aucuns semblent parfois inclure sous le terme d'atténuation des items et des exemples qui, à notre avis du moins, se laissent difficilement catégoriser ainsi. Par exemple, nous avons du mal à voir comment certains des exemples du passif de Hyland pourraient jouer un rôle atténuateur :

(3) No detectable signs could be observed... (Hyland 1998 : 77)

Et il inclut encore d'autres exemples dont la valeur atténuante ne nous paraît pas claire, comme en (4) :

(4) However, these procedures are consistent with the large body of literature (ibid.: 203)

Pareillement, dans son travail portant sur l'atténuation dans le discours scientifique<sup>14</sup>, Varttala inclut dans son analyse des adverbes comme *drastically, greatly* et *significantly* (Varttala 2001 : 130-132, 301) et certaines occurrences de noms comme *idea* et *notion* (Varttala 1999 : 183, Varttala 2001 : 142) :

(5) This implies several simple but powerful *ideas* for making vertical systems work (Varttala 2001 : 142)<sup>15</sup>

Or, il nous paraît difficile de voir comment ces mots pourraient avoir l'effet d'un *hedge*.

---

<sup>12</sup> En ce qui concerne la sporadicité, nous avons vu au chapitre 3 que cette catégorie est souvent incluse sous le terme de modalité épistémique (voir par exemple Halliday 1994 : 357 et McEnery et Kifle 2002)

<sup>13</sup> Nous reviendrons à cette question au chapitre 6.

<sup>14</sup> *Hedging in Scientifically Oriented Discourse* (Varttala 2001)

<sup>15</sup> Les italiques sont celles de Varttala.



Enfin, dans leur étude sur les critiques de livres, Salager-Meyer, Ariza et Pabon (2007) visent à examiner l'emploi de modalisateurs épistémiques en tant qu'atténuateurs. Initialement, les auteurs semblent restreindre la catégorie de la modalité épistémique aux auxiliaires modaux à valeur épistémique, aux semi-auxiliaires comme *sembler* et *paraître*, et aux noms, adjectifs et adverbes épistémiques (ibid. : 114). Or, les exemples donnés ne respectent pas cette délimitation, ce qui laisse le lecteur dans un état de confusion en ce qui concerne les types de marqueurs inclus et exclus. Parmi les marqueurs commentés, on trouve le conditionnel (hors de l'emploi spécifique appelé le conditionnel épistémique), le passif et des expressions comme *un peu*. Sont inclus des exemples comme ceux-ci :

(6) Il *aurait été utile* que l'éditeur... (Salager-Meyer et al 2007 : 117)<sup>16</sup>

(7) On *peut/pourrait* (toutefois/cependant) *regretter* que les auteurs n'aient pas consacré un chapitre aux syndromes myasthéniques. (ibid. : 117)

(8) *The chapter on oncogenes* *should have been placed* earlier in the book. (ibid. : 118)

Alors qu'on pourrait peut-être considérer ces énoncés comme des exemples d'atténuation (du moins les deux premiers), leur valeur épistémico-modale est contestable. Le locuteur ne prononce pas de jugement sur la valeur de vérité de l'information ; il énonce simplement ses opinions, et dans le troisième exemple, il le fait même d'une manière très directe.

Un clin d'œil sur les différents exemples et items inclus par différents chercheurs révèle donc que la catégorie d'atténuation reste dans une certaine mesure une sorte d'assemblage composite d'éléments linguistiques que les chercheurs interprètent de diverses manières (cf. Crompton 1997). Par ailleurs, les notions d'atténuation et de modalité épistémique sont parfois confondues et utilisées de façon interchangeable (voir Markkanen et Schröder 1997 : 6-7). Cependant, comme nous l'avons déjà mentionné, la solution proposée par Crompton ne semble pas résoudre le problème : bien que son approche semble à première vue prometteuse, sa définition de ce qu'est un atténuateur correspond à la modalité épistémique et il n'y aura donc pas de terme générique pour les autres stratégies employées pour faire des réserves<sup>17</sup>.

---

<sup>16</sup> Nous avons gardé les italiques et le soulignement de Salager-Meyer et al. Les marqueurs considérés comme des atténuateurs sont soulignés (Salager-Meyer et al. 2007 : 117).

<sup>17</sup> Une autre chose est que Crompton paraît lui-même inclure des exemples discutables. Il est difficile de voir comment *rather* dans *I'm rather hungry* qualifie la valeur de vérité de la proposition *I'm hungry*, tel que Crompton le maintient. La proposition est tout aussi présentée comme vraie que l'expression *rather* soit là ou non, *rather* ne fait que spécifier le niveau de faim (voir aussi Salager-Meyer 1998 : 299).

En résumé, les conceptions du phénomène d'atténuation varient encore considérablement. Chez certains (Crompton 1997), les adoucisseurs semblent se restreindre aux marqueurs épistémico-modaux d'incertitude. Chez d'autres (Markkanen et Schröder 1997), la liste d'expressions langagières susceptibles de remplir le rôle d'un adoucisseur ne connaît aucune limite.

En ce qui concerne le rôle de l'atténuation dans le texte, il semble y avoir une plus grande unanimité. On s'accorde généralement à voir l'atténuation comme une stratégie rhétorique, employée par l'auteur dans le but de négocier avec ses lecteurs et, en fin de compte, de les convaincre. Certains mettent cet aspect davantage en avant que d'autres : Foullioux et Tejedor de Felipe (2004 : 114) avancent à cet égard: « ...il s'agit bien, en effet, de faire **paraître** « moins forte » une expression linguistique, mais non de la **rendre** « moins forte » » et « ...l'atténuation est un artifice, une feinte, en définitive une stratégie du détour. » Dans le même ordre d'idées, Beaufrère-Bertheux (1997), en parlant de l'atténuation dans les textes scientifiques médicaux, maintient que l'auteur d'un texte scientifique recourt à l'atténuation afin de rétablir avec ses lecteurs une connivence en train de disparaître : « il lui faut donc abuser de précautions pour se rapprocher de celui dont il s'est éloigné » (Beaufrère-Bertheux 1997 : 235), et elle avance même l'hypothèse que « le *hedging* est proportionnel à la certitude » (ibid. : 235). Alors que certains insistent beaucoup sur le rôle rhétorique et persuasif de l'atténuation, d'autres prennent soin de rappeler que les atténuateurs sont parfois employés simplement pour exprimer une incertitude réelle, c'est-à-dire pour être précis et pour transmettre ses observations avec exactitude (e.g. Salager-Meyer 1997).

Les marqueurs épistémico-modaux d'incertitude, au centre d'intérêt de cette étude, sont reconnus comme la classe dominante d'adoucisseurs. Bien que cette thèse ne traite pas du phénomène d'atténuation en général, c'est dans les études portant sur l'atténuation que nous trouvons les hypothèses et les résultats les plus intéressants pour la présente étude. Une grande partie des recherches empiriques menées dans ce domaine porte sur les marqueurs modaux épistémiques, uniquement (Vihla 2000) ou partiellement (Hyland 1998 ; Varttala 1999, 2001 ; Salager-Meyer 1994).

## 4.2 Quelques résultats antérieurs

Il existe donc de nombreux travaux portant sur le phénomène d'atténuation, mais comme nous l'avons vu, les chercheurs semblent en avoir des conceptions assez différentes, ce qui se reflète dans les résultats parfois divergents des diverses études quantitatives. Dans ce qui suit,

nous présenterons certains des résultats quantitatifs provenant d'études antérieures, en mettant l'accent sur les études qui portent sur des articles de recherche et qui intéressent donc directement le présent travail. La présentation ne prétend évidemment pas à l'exhaustivité ; nous avons choisi de signaler certaines études importantes qui pourront servir de base de comparaison à nos propres résultats.

Comme déjà évoqué, la grande majorité des études portant sur l'atténuation dans les articles scientifiques concernent des articles rédigés en anglais. En outre, la plupart de ces études se restreignent à une discipline, la discipline de médecine étant la plus amplement étudiée. Nous allons d'abord présenter quelques résultats de ces études anglophones et unidisciplinaires, et nous passerons ensuite aux études comparatives.

Comme mentionné plus haut, une des études les plus influentes dans le domaine de l'atténuation est celle de Hyland (1998). Dans cette étude, Hyland a examiné en détail l'emploi d'atténuateurs dans un corpus de 26 articles scientifiques en biologie. Il a repéré dans le corpus une fréquence relative d'atténuateurs de 20,9 pour mille du nombre total des mots (Hyland 1998 : 104). Les catégories grammaticales les plus fréquemment employées étaient (en ordre décroissant) les verbes lexicaux, les adverbes, les adjectifs et les auxiliaires modaux (ibid. : 104). La forme la plus récurrente était le verbe *indicate*<sup>18</sup>, suivi des auxiliaires *would* et *may* (ibid. : 149). En ce qui concerne la distribution sur les différentes parties rhétoriques, Hyland a observé que la section Discussion était la plus riche en adoucisseurs avec une fréquence relative de 36,4 pour mille<sup>19</sup> (ibid. : 153).

Cette dernière observation corrobore ceux de Salager-Meyer (1994), qui a étudié la distribution d'adoucisseurs dans les différentes sections à partir d'un corpus d'articles de recherche en médecine<sup>20</sup>. Les résultats ont montré que la section Discussion est de loin la plus riche en adoucisseurs, au point que 13% des mots dans les parties de discussion ont été classifiés comme des adoucisseurs (Salager-Meyer 1994 : 155). Ce taux extrêmement élevé s'explique probablement par la définition large de l'atténuation qu'adopte Salager-Meyer dans cette étude (cf. *supra*, 4.1). En ce qui concerne les différents types d'adoucisseurs, Salager-Meyer a observé que ce sont notamment les verbes modaux qui jouent le rôle d'adoucisseurs (ibid. : 155).

---

<sup>18</sup> Hyland ne semble pas avoir mis en discussion la polyvalence du verbe *indicate* (sur laquelle nous reviendrons au chapitre 6) et dans le cas où toutes les occurrences de *indicate* ont été comptées comme des atténuateurs, cela pourrait expliquer la haute fréquence de ce verbe.

<sup>19</sup> En comparaison, les fréquences correspondantes pour les autres sections sont : 20,0 pour les sections Introduction et Résultats et 3,4 pour la section Méthode.

<sup>20</sup> Outre les articles scientifiques, le corpus de Salager-Meyer (1994) comprend aussi des études de cas. Les résultats de cette partie du corpus ne seront pas commentés ici.

Dans son travail sur l'emploi d'atténuateurs dans des textes médicaux, Varttala (1999) tire la même conclusion : ce sont les verbes modaux (auxiliaires et verbes lexicaux) qui constituent la classe grammaticale la plus importante pour exprimer l'atténuation. Varttala a identifié 80 expressions de possibilité épistémique, y compris des prototypes tels que *may*, *possibly*, *perhaps*, ainsi que des expressions dont l'effet adoucissant peut paraître moins clair (*idea*, *notion*, *propose*, *claim*). De même que les études de Hyland (1998) et de Salager-Meyer (1994), celle de Varttala (1999) montre que c'est la section Discussion qui contient le plus grand nombre d'adoucisseurs. Les expressions étudiées constituaient en moyenne 16,42 pour mille des mots dans les sections Discussion, 9,09 pour mille dans les introductions et 3,02 pour mille dans les sections Résultats et Méthode (Varttala 1999).

Ces recherches antérieures constatent donc au moins deux choses : 1) les verbes épistémiques (auxiliaires et verbes lexicaux) constituent une stratégie d'atténuation particulièrement importante et 2) le nombre d'adoucisseurs est particulièrement élevé dans la section Discussion et assez limité dans les sections Matériels/Méthode et Résultats. Il convient de noter que toutes ces études portent sur des articles tirés de disciplines des sciences naturelles : l'étude de Hyland traite d'articles en biologie et les études de Salager-Meyer et de Varttala portent sur des articles en médecine. La question de la distribution des atténuateurs dans les différentes parties d'un article n'est pas aussi pertinente pour les sciences humaines, qui ne sont pas soumises aux mêmes contraintes concernant la composition rhétorique d'un article. Etant donné que la distribution d'atténuateurs dans les différentes sections d'un article scientifique a été si bien renseignée, cette question ne sera pas examinée plus profondément dans la présente étude. Nos résultats quantitatifs peuvent cependant révéler si ce sont également les verbes qui constituent la classe d'atténuateurs la plus importante dans les articles rédigés en français et en norvégien.

Bien que les tendances attestées dans ces études soient les mêmes, on remarque que les chiffres obtenus sont très différents. Dans le corpus de Varttala, 16,42 mots pour mille dans les sections Discussion sont classifiés comme des adoucisseurs, alors qu'il s'agit de 13 *pour cent* des mots dans le corpus de Salager-Meyer. Cela peut naturellement refléter une vraie différence entre les deux corpus, due par exemple aux sujets traités ou à une différence dans les pratiques et les préférences des revues. Mais avant tout, cela montre dans quelle mesure la sélection de marqueurs à inclure sous le concept d'atténuation influence les résultats. Cela rend difficile la comparaison des résultats quantitatifs des différentes études.

L'étude qui est, peut-être, la plus proche de la nôtre, méthodologiquement et conceptuellement, est celle de Vihla (2000). Cette étude ne porte pas sur l'atténuation en

général, mais se restreint à certaines expressions de possibilité, en mettant l'accent sur la possibilité épistémique. Le corpus examiné est composé de différents types de textes médicaux (articles de recherche, manuels, éditorial, etc.), et les marqueurs étudiés sont : *may*, *might*, *maybe*, *perhaps*, *possibly*, *(it is) possible that* et *possibility that*. La sélection de marqueurs reflète le fait que l'auteur a voulu étudier les prototypes de marqueurs de possibilité épistémique, se concentrant sur des items communément admis comme tels (Vihla 2000 : 221). L'analyse a révélé que l'auxiliaire *may* dépasse de loin les autres marqueurs en fréquence. *May* est l'expression de possibilité la plus récurrente, avec une fréquence relative (en pour mille) de 1,9 dans les articles de recherche<sup>21</sup>. Cet auxiliaire modal est aussi l'expression la plus récurrente de la possibilité épistémique, *may* épistémique ayant une fréquence relative de 1,8 pour mille – ce qui implique que la grande majorité des occurrences de *may* transmettent une valeur épistémique. Le deuxième modalisateur épistémique est *might*, avec une fréquence relative de 0,4 dans les articles de recherche. Les autres expressions sont rares. L'adjectif *possible* par exemple, a une fréquence relative de 0,4 (toutes occurrences incluses), mais les occurrences épistémiques de cet adjectif ne constituent que 0,1 pour mille. Vihla n'a repéré aucune occurrence de l'adverbe *maybe* dans les articles de recherche, en revanche elle a repéré quelques occurrences de *perhaps* (fréquence relative de 0.1), et *maybe* est parfois utilisé dans les articles de vulgarisation scientifique. Au total, les sept marqueurs épistémico-modaux ont une fréquence relative de 2.5 pour mille des mots (Vihla 2000 : 219)<sup>22</sup>. Il sera intéressant de comparer ces chiffres, qui se basent sur un nombre de 33 articles de recherche, avec les nôtres, et de voir en outre dans quelle mesure la fréquence de ces expressions épistémiques prototypiques de l'anglais correspond à celle des correspondants ou quasi-équivalents en norvégien et en français.

Nous avons vu que c'est surtout la discipline de médecine qui a été étudié quand il s'agit de l'emploi d'adoucisseurs. Une exception est Hunston (1994) qui a étudié l'emploi d'adoucisseurs dans un corpus de 10 articles en linguistique appliquée. Ces articles sont tous expérimentaux et ils sont donc en quelque sorte apparentés aux articles en médecine, qui relèvent très souvent du sous-genre *article expérimental*. Hunston met l'accent sur les fonctions pragmatiques de divers types d'atténuateurs et l'approche est qualitative plutôt que

---

<sup>21</sup> Nous ne rapportons ici que les chiffres qui concernent les articles de recherche.

<sup>22</sup> La fréquence relative est plus élevée dans les autres genres examinés, l'article de recherche étant le genre le moins riche en adoucisseurs (Vihla 2000 : 212). Varttala (2001 : 200) a fait une observation similaire en comparant articles scientifiques et articles de vulgarisation scientifique : ces derniers contenaient beaucoup plus d'atténuateurs que les premiers. Ces observations suggèrent que l'atténuation n'est pas une caractéristique réservée au genre de l'article scientifique, mais qu'elle peut être également (ou plus) importante dans d'autres genres (du moins en ce qui concerne le discours médical).

quantitative. Les résultats de l'étude de Hunston seront plus amplement présentés dans la section 7.1, qui traite de la question des fonctions pragmatiques.

Le manque d'études comparatives a été signalé plusieurs fois et par différents chercheurs. Hyland (1998 : 10) rappelle que nous savons peu de choses sur d'éventuelles différences disciplinaires. Markkanen et Schröder constatent qu'il faut plus de recherches sur le phénomène d'atténuation, et qu'il y a notamment besoin d'études interculturelles et interdisciplinaires (Markkanen et Schröder 1997 : 14). Varttala (1999 : 195) clôt son article sur l'emploi d'adoucisseurs dans le discours médical anglais en proposant d'étudier l'emploi de tels marqueurs dans d'autres disciplines et dans d'autres langues. Jusqu'à ces toutes dernières années, il n'existait guère d'études comparant plusieurs disciplines ou plusieurs langues, mais ces derniers temps quelques études comparatives, notamment des études interdisciplinaires, ont vu le jour. Hyland (1999b) a comparé l'emploi d'atténuateurs dans huit disciplines<sup>23</sup> de différents domaines scientifiques (sciences humaines, sociales et naturelles), et il a observé un usage plus fréquent d'atténuateurs dans les disciplines les moins « dures », y compris la linguistique appliquée (ibid. : 106-107). La fréquence relative d'atténuateurs varie entre 8,2 pour mille en ingénierie électrique et 20,0 en marketing. Les items les plus récurrents sont *may* et *would* (cf. aussi les résultats de Hyland 1998 : 149), mais le verbe *indicate* figure maintenant beaucoup plus bas dans la liste (Hyland 1999b : 108), ce qui suggère que l'emploi de ce verbe pourrait être une caractéristique des articles en biologie. Hyland (1999b) conclut qu'il existe des différences considérables entre les disciplines des sciences naturelles et celles des sciences humaines en ce qui concerne la fréquence d'atténuateurs. Ses résultats corroborent ceux de Ventola (1997), qui a comparé 9 articles des sciences naturelles à 8 articles des sciences humaines et sociales et trouvé un emploi plus extensif d'adoucisseurs dans ce dernier groupe. Ces observations sont aussi en concordance avec l'hypothèse initiale de Markkanen et Schröder (1997 : 10) selon laquelle l'emploi d'adoucisseurs serait plus élevé dans des disciplines telles que la linguistique et la philosophie que dans des disciplines des sciences naturelles, en raison des différentes stratégies d'argumentation dont font usage ces disciplines. Dans les disciplines des sciences naturelles, l'argumentation se fonderait sur des données expérimentales et sur des preuves concrètes alors que dans les sciences humaines, l'expression linguistique est censée être plus centrale pour établir une argumentation convaincante. Cependant, après avoir fait référence à un petit

---

<sup>23</sup> Les disciplines examinées sont la microbiologie, la physique, le marketing, la linguistique appliquée, la philosophie, la sociologie, l'ingénierie mécanique et l'ingénierie électrique. Sept articles de chacune des disciplines ont été étudiés.

nombre d'études montrant que les adoucisseurs sont fréquents et jouent un rôle important y compris dans les textes des sciences naturelles (Butler 1990, Darian 1995), Markkanen et Schröder (1997 : 10) finissent par abandonner l'hypothèse qu'ils ont posée au début et suggèrent que le domaine de recherche joue un rôle moins important que ce qu'on a cru antérieurement.

Dans son ouvrage de 2004 [2000], Hyland poursuit ses études interdisciplinaires, et consacre un chapitre à l'emploi d'atténuateurs et d'amplificateurs. Cette fois, les données sur lesquelles se basent les résultats sont des lettres scientifiques et non des articles de recherche. Il n'y pas de grandes différences entre les trois disciplines investiguées (la biologie, la physique et la chimie) – ce qui ne doit pas surprendre compte tenu de leur parenté – mais Hyland conclut que les atténuateurs sont aussi importants dans le genre de la lettre scientifique que dans l'article de recherche. Cela suggère, encore une fois (cf. note 22, le présent chapitre), que l'article de recherche n'est pas le seul genre scientifique où l'atténuation se manifeste comme un trait caractéristique.

Varttala (2001) a étudié l'emploi d'atténuateurs dans un corpus composé d'un total de 30 articles scientifiques tirés des disciplines de médecine, d'économie et de technologie. Il observe que le nombre d'atténuateurs est plus élevé en économie (fréquence relative d'environ 30 pour mille) qu'en médecine et technologie (fréquences relatives d'environ 22 pour mille) (ibid. : 100). *May* épistémique est parmi les items les plus récurrents avec une fréquence relative de 2,97 pour mille en médecine, de 2,22 en économie et de 1,04 en technologie (ibid.: 118), alors que l'emploi épistémique de *would* s'avère très rare. Cette dernière observation s'oppose à celle de Hyland, mais encore une fois, la divergence semble résider dans des conceptions différentes de ce qu'est l'atténuation plutôt que dans de véritables différences de fréquence<sup>24</sup>.

Pour récapituler, plusieurs études comparatives suggèrent qu'il y a des différences considérables entre les domaines scientifiques concernant l'emploi d'atténuateurs, en ce sens que les disciplines des sciences humaines font un usage plus étendu d'atténuateurs que les disciplines plus « dures ». Cependant, d'autres études montrent que l'atténuation joue un rôle important dans les sciences naturelles aussi. En outre, les études comparatives restent rares, et les corpus sur lesquels elles se basent sont souvent relativement limités en taille. Toutefois, à en juger par les recherches antérieures, il semble raisonnable de supposer qu'on utilise généralement plus d'atténuateurs dans les sciences humaines que dans les sciences naturelles.

---

<sup>24</sup> Pour une discussion des différentes valeurs de *would* et leur statut épistémique, voir Varttala 2001 : 110-112.

Néanmoins, il faut avouer que la frontière qu'on trace traditionnellement entre sciences naturelles et sciences humaines paraît parfois quelque peu artificielle et trop catégorique. Plusieurs disciplines et sous-disciplines se trouvent quelque part entre les deux domaines, s'inspirant des deux champs, et le type de processus de recherche importe aussi ; par exemple, il pourrait y avoir une différence entre articles expérimentaux et articles plus théoriques ou articles de synthèse qui se manifesterait indépendamment des domaines scientifiques dont ils sont issus.

Les différences entre langues et cultures sont encore moins étudiées que les différences entre disciplines. Néanmoins, le peu d'études qui existent suggèrent que la fréquence d'adoucisseurs varie là aussi considérablement. Clyne (1991) observe que les chercheurs allemands emploient plus d'adoucisseurs que les chercheurs anglais, qu'ils écrivent dans leur langue maternelle ou en anglais<sup>25</sup>. Vassileva (1997), de son côté, compare des articles linguistiques rédigés en anglais et en bulgare et montre que les linguistes anglais utilisent plus d'atténuation que les bulgares. De même, une étude de Crismore, Markkanen et Steffensen (1993), comparant la fréquence d'adoucisseurs chez des étudiants finnois et américains, indique que les finnois employaient plus d'adoucisseurs que les américains.

Il peut également y avoir des différences au niveau des moyens linguistiques utilisés pour exprimer la modalité épistémique, même entre des langues proches qui disposent à peu près des mêmes moyens formels : Løken (1997 : 56, 2004 : 34) observe que les textes anglais de son corpus contiennent un plus grand nombre d'auxiliaires modaux à valeur épistémique que les textes norvégiens. En revanche, les collocations épistémiques sont plus fréquentes dans les textes norvégiens (Løken 2004 : 35). Il convient de noter que les résultats de Løken se basent sur des textes de fiction et ils ne sont donc pas nécessairement généralisables aux textes scientifiques.

Autant que nous sachons, aucune étude ne s'est proposée de comparer l'emploi d'atténuateurs dans le discours scientifique anglais avec leur emploi dans le discours scientifique français ou norvégien. Or, comme nous l'avons vu au chapitre 2, d'autres aspects de ces discours ont été comparés, suggérant d'un côté que le discours scientifique norvégien est relativement proche de celui de l'anglais, alors que le discours scientifique français s'en distingue plus nettement (Fløttum et al. 2006), et de l'autre côté que le discours scientifique français accepte et utilise dans une plus grande mesure des remarques critiques directes et non

---

<sup>25</sup> Rappelons que Clyne (1991) adopte une taxonomie relativement large, incluant le passif et différentes constructions impersonnelles (cf. *infra*, 4.1). Mais ce sont les auxiliaires modaux qui sont les moyens d'atténuation les plus souvent utilisés (Clyne 1991: 63).



atténuées par rapport à ce qui se fait dans le discours scientifique anglais. Il s'ensuit qu'on pourrait s'attendre à des différences importantes entre le corpus anglais et le corpus français, alors que le corpus anglais et le corpus norvégien seraient plus proches l'un de l'autre.

Dans l'étude de Crismore, Markkanen et Steffensen (1993), il s'avère par ailleurs que les hommes emploient plus fréquemment des adoucisseurs que les femmes. D'autres études indiquent le contraire ; Preisler (1986) observe que les auteurs féminins emploient plus d'adoucisseurs que les auteurs masculins, suggérant que les chercheurs féminins éprouvent un sentiment d'infériorité dans la communauté scientifique. D'autres encore maintiennent qu'un emploi extensif d'adoucisseurs ne reflète pas un sentiment d'impuissance, mais une attitude polie envers les lecteurs (Holmes 1990 : 202, citée dans Markkanen et Schröder 1997 : 9). Vu les résultats divergents et le nombre limité d'études effectuées dans ce domaine, la question de savoir si le sexe de l'auteur joue un rôle n'est donc pas encore résolue.

Malgré le grand nombre d'études qui portent sur l'atténuation en anglais, et notamment sur le discours médical en anglais, il reste, comme nous l'avons vu, beaucoup de questions non résolues dans ce domaine de recherche. En particulier, se manifeste le besoin d'études comparatives et d'études portant sur d'autres langues et sur d'autres disciplines. En mettant l'accent sur les discours scientifiques français et norvégien, la présente étude a pour but de pallier ce manque d'études portant sur d'autres langues. L'anglais a été inclus dans l'analyse quantitative afin d'établir une base de comparaison, mais il jouera un rôle secondaire par rapport aux analyses des deux autres langues. En ce qui concerne les disciplines, la médecine fait aussi l'objet d'étude du présent travail, avec la linguistique. Si la médecine est surreprésentée dans les recherches antérieures, cela vaut uniquement pour les études portant sur l'anglais, et le discours médical en norvégien et en français a été peu étudié. Outre le rôle des langues et des disciplines, la présente étude évoquera aussi, comme il a été mentionné dans l'introduction, la question du rôle du sexe de l'auteur, mais uniquement en ce qui concerne l'aspect quantitatif. Nous verrons par la suite dans quelle mesure nos observations sur la variation langagière et disciplinaire et sur les différences de sexe convergent et divergent avec celles qui ont été faites dans des études antérieures et qui ont été rapportées ici. La seconde partie de la thèse présente et discute les résultats obtenus (chapitres 6-7), après une présentation du corpus et de l'approche méthodologique adoptée (chapitre 5).



## SECONDE PARTIE

### ANALYSES



## 5 MATÉRIAUX ET MÉTHODE

Les chapitres précédents nous ont permis de préciser les fondements théoriques sur lesquels s'appuient nos analyses. Le chapitre 2 s'est centré sur le genre de l'article de recherche, ses caractéristiques et sa variabilité interne, et sur des recherches antérieures portant sur ce genre, notamment celles qui adoptent une approche comparative. Ont été signalés aussi quelques défis liés à la perspective doublement comparative adoptée dans la présente étude et plus largement dans l'ensemble des études du projet KIAP. Le chapitre 3 a passé en revue différentes conceptions de la catégorie de la modalité, ce qui nous a permis, finalement, de retenir notre propre conception de la modalité épistémique, conception qui à son tour guidera à la fois la sélection de marqueurs à étudier et le codage d'occurrences, dont nous rendrons compte dans le chapitre 6. Le chapitre 4 a traité de la notion de 'hedging', souvent assimilée à celle de modalité épistémique, et des recherches antérieures dans ce domaine. Ces recherches antérieures, avec celles présentées au chapitre 2, permettent de faire certaines hypothèses concernant la variation entre disciplines et langues dans l'emploi d'atténuateurs ; nous y reviendrons à la section 6.1.

Nous passons maintenant aux analyses du corpus. Ce premier chapitre rendra compte de la composition des matériaux (5.1) et des grandes étapes de la démarche méthodologique suivie lors du travail (5.2).

### 5.1 Présentation des matériaux

Comme mentionné au début de ce travail, le corpus de textes sur lequel se base la présente étude est extrait du corpus KIAP ; il s'agit d'une large base électronique où sont rassemblés 450 articles de recherche tirés des disciplines de médecine, de linguistique et d'économie politique et rédigés en trois langues : l'anglais, le français et le norvégien. Les avantages d'utiliser un corpus déjà constitué et balisé sont multiples et évidents ; avant tout cela nous a épargné beaucoup de travail et de temps. Mais cela implique aussi que nous n'avons pas eu d'influence sur la sélection de textes et sur la composition du corpus, ce qui n'est pourtant pas jugé problématique pour la présente thèse, car notre objectif principal, celui d'examiner les variations entre disciplines et langues dans le discours scientifique, correspond à celui du projet KIAP et le corpus KIAP a été constitué précisément pour servir cet objectif-là. Afin d'assurer une bonne représentativité du corpus ainsi que de créer un corpus approprié à l'objectif visé, la sélection d'articles a été réalisée suivant certains critères spécifiques. Ces

critères, dont certains ont été mentionnés à la section 2.2.2, ne seront pas présentés de façon détaillée ici ; pour une telle présentation nous renvoyons à Fløttum et al. (2006 : 7-16) et à la page web <http://www.uib.no/kiap/mdcorpusdescr.htm>. Il convient cependant de reprendre brièvement ici les principes les plus importants :

- Afin de garantir un certain niveau de scientificité, les articles sont tous tirés de revues reconnues qui utilisent un système d'évaluation par rapporteurs anonymes (des revues à comité de lecture).
- Les articles sont dans la grande majorité des cas écrits par des locuteurs natifs. Cependant, en ce qui concerne les articles en médecine, ceux-ci sont à quelques exceptions près rédigés par plusieurs auteurs en collaboration. Ces auteurs sont souvent de différentes nationalités et tous les auteurs ne sont donc pas nécessairement des locuteurs natifs de la langue en question.
- En vue d'assurer une certaine homogénéité dans les différents sous-corpus, on a cherché à éviter les domaines les plus « durs » de la linguistique et les plus « mous » de la médecine pour faire en sorte que le corpus médical et le corpus linguistique représentent adéquatement les sciences naturelles et les sciences humaines, respectivement, même si la médecine et la linguistique ne sont peut-être pas les représentants typiques de leurs domaines scientifiques respectifs. Cependant, on a dû assouplir quelque peu ce critère, pour des raisons pratiques.

L'établissement du corpus KIAP s'est fait par étapes et la présente étude se base sur la première partie du corpus, qui a été terminée en 2002 et qui rassemble 180 articles, équitablement répartis sur les trois disciplines et les trois langues mentionnées. Comme la présente étude ne porte pas sur la discipline d'économie politique, il reste un ensemble de 120 textes à étudier, soit un total de 494 111 mots. Pour nommer les différents sous-corpus, nous conservons les étiquettes utilisées par le projet KIAP : le corpus d'articles linguistiques rédigés en anglais sera désigné par l'étiquette Engling, le corpus d'articles linguistiques rédigés en français sera appelé Frling, et celui d'articles linguistiques rédigés en norvégien sera désigné par l'étiquette Noling. Les corpus d'articles médicaux seront désignés par les étiquettes Engmed, Frmed et Nomed, selon la langue dans laquelle les articles sont rédigés. Il convient de noter que le comptage de mots vaut pour le « corps » des articles, c'est-à-dire le texte complet à l'exception des résumés, notes, références bibliographiques, citations, exemples linguistiques, tableaux, figures et appendices. Les requêtes sur ordinateur, sur

lesquelles nous reviendrons au chapitre 6, portent aussi sur le corps des textes, les autres parties étant exclues.

Les périodiques dont sont tirés les 120 articles sont répertoriés dans le tableau 1. Le tableau 2 montre la répartition du nombre total de mots dans les différents sous-corpus. Une liste complète des articles individuels avec titre, auteur(s), journal, année de publication et nombre de mots est fournie en appendice 1. Les articles sont tous parus entre 1995 et 2002.

**Tableau 1 : Composition du corpus**

Langue	Discipline	Mots	Revues	Nombre d'articles
Anglais	Linguistique	170981	Journal of Linguistics	10
			English for Specific Purposes	5
			Language	4
	Médecine	59410	Linguistic Inquiry	1
			Journal of the American Medical Association	10
			British Medical Journal	5
			The International Journal of Cancer	5
Français	Linguistique	68727	Travaux de linguistique	18
	Médecine	61180	Marges linguistiques	2
			Annales de médecine interne	15
			Maladies chroniques au Canada	5
Norvégien	Linguistique	90579	Norsk lingvistisk tidsskrift	20
	Médecine	43234	Tidsskrift for Den Norske Lægeforening	20
<b>Total</b>		<b>494111</b>		<b>120</b>

**Tableau 2 : Répartition du nombre de mots**

	Anglais	Français	Norvégien	Total
Linguistique	170981	68727	90579	330287
Médecine	59410	61180	43234	163824
<b>Total</b>	<b>230391</b>	<b>129907</b>	<b>133813</b>	<b>494111</b>

Il ressort de ces tableaux au moins deux aspects qui méritent un commentaire. Premièrement, on remarquera que le nombre total de mots se répartit sur les six sous-groupes de façon très inégale. Généralement, les articles en médecine sont plus courts que ceux en linguistique, tendance qui peut s'expliquer par des contraintes éditoriales et par l'idéal en recherche médicale d'écrire des articles aussi brefs et concis que possible (Gilhus 2003 : 22, cf. *supra* 2.1.2). On n'a pas la même tradition en linguistique, où la longueur des articles peut beaucoup varier, même si la plupart des revues linguistiques indiquent, elles aussi, un nombre maximum

de mots dans leurs conseils aux auteurs. La différence de longueur entre les articles linguistiques et les articles médicaux ne doit donc pas surprendre. Ce qui est peut-être plus frappant, c'est la longueur des articles linguistiques en anglais comparée à celle des articles linguistiques en français et en norvégien. Les causes précises de cet écart assez notable entre la longueur des articles anglais d'un côté et celle des articles français et norvégiens de l'autre sont mal connues, mais il est raisonnable de supposer que l'orientation éditoriale des différentes revues joue un rôle important à cet égard. Vu la distribution inégale du nombre de mots, il sera particulièrement important de mettre en évidence les fréquences relatives plutôt que les chiffres absolus<sup>1</sup>.

Deuxièmement, on notera que le nombre de sources n'est pas le même pour les trois langues. Cela s'explique par le fait que le nombre de revues accessibles varie selon les langues (cf. Fløttum et al. 2006 : 8). Pour garantir la qualité scientifique des textes, les articles sont tous tirés de revues reconnues qui publient des articles évalués par les pairs. Alors qu'il existe un bon nombre de revues de ce type publiant des articles en français, et un nombre abondant pour les articles en anglais, le nombre de telles revues publiant des articles en norvégien reste très limité. Voilà pourquoi il n'y a que deux périodiques norvégiens sur la liste, un pour chaque discipline.

## 5.2 Méthode d'analyse

Dans cette section nous ne reportons que les grandes lignes méthodologiques suivies lors de la sélection et l'analyse des données. Des descriptions détaillées des principes et des démarches méthodologiques suivies lors des différentes étapes de l'analyse seront offertes en temps utile (plus précisément dans les sections 6.2, 6.3, 6.4.1, et dans le chapitre 7).

Comme mentionné au chapitre 1, une combinaison de méthodes quantitatives et qualitatives a été utilisée. L'analyse quantitative (chapitre 6) comprend les articles français, norvégiens et anglais, alors que l'analyse qualitative (chapitre 7) sera restreinte aux textes français et norvégiens, en mettant l'accent sur un nombre limité de marqueurs sélectionnés. Notons que la dichotomie entre analyses quantitatives et analyses qualitatives n'est pas nette est stricte : l'analyse quantitative n'est évidemment pas indépendante de toute approche qualitative et l'analyse qualitative n'est pas dépourvue de chiffres. La distinction reflète

---

<sup>1</sup> Cela ne va pas entièrement sans problèmes non plus. Les fréquences relatives se basent sur le nombre de mots ; pourtant la longueur des articles et par conséquent la fréquence relative d'un phénomène linguistique, peuvent être influencées par des différences systématiques entre les langues ; par exemple, au nom norvégien *teorien* correspondent deux mots en anglais : *the theory* et deux en français : *la théorie* (cf. Fløttum et al. 2006). Cependant, les différences de ce type n'auraient qu'une influence infime sur les chiffres observés.



simplement deux objectifs différents : dans le chapitre 6, l'accent est mis sur la fréquence des marqueurs (mais aussi sur la fréquence de différents *types* de marqueurs, dont le classement entraîne, inmanquablement, une dimension qualitative), alors que dans le chapitre 7, ce sont avant tout les buts communicatifs et les effets pragmatiques qui retiendront notre attention. Or, dans les études approfondies des marqueurs sélectionnés, la question de la fréquence réapparaît quand nous comparons l'emploi d'un marqueur à travers les disciplines ou avec son/ses correspondant(s) dans l'autre langue. La division en analyses quantitatives et analyses qualitatives correspond ainsi au fait que l'aspect quantitatif est plus saillant dans les premières, tandis que l'aspect qualitatif est plus saillant dans les secondes. Cependant, elle ne doit pas être conçue comme une frontière absolue entre les deux types d'approche.

Avant de procéder aux différentes étapes de l'analyse, nous voudrions signaler deux défis d'ordre méthodologique qui ont été centraux pour la présente étude. Premièrement, le procès de sélection même de marqueurs à étudier constitue un défi à cause de la nature de la catégorie linguistique concernée. Notre but étant d'étudier la réalisation linguistique d'une catégorie sémantique (à savoir la modalité épistémique), c'est-à-dire une catégorie qui ne se laisse pas définir extensionnellement, nos données ne sont pas « données » d'avance comme elles peuvent l'être pour celui qui, dès le départ, consacre son étude à une ou plusieurs expressions spécifiques, ou à une catégorie grammaticale bien définie. Il faut faire une sélection de marqueurs censés représenter la catégorie en question. Dans la présente étude, cette sélection se base sur une analyse exploratoire, elle-même guidée par la conception théorique de la modalité épistémique élaborée au chapitre 3 et par certains choix méthodologiques tel que celui d'exclure les marqueurs de certitude. Ces principes sont affinés et expliqués plus en détail dans la section 6.2.

Deuxièmement, la perspective interlinguistique constitue un défi en soi : comment savoir si l'on compare la même chose dans les trois langues ? Nous avons évoqué cette question au chapitre 2, où nous avons posé que l'analyse exploratoire constitue un moyen de rendre justice à chacune des trois langues (et des deux disciplines) en ce qu'elle mobilise des parties égales des différents sous-corpus. Ainsi, la sélection de marqueurs d'une langue s'est faite indépendamment de la sélection de marqueurs dans les autres langues. Elle ne se base pas sur des équivalents présumés, mais sur l'usage effectif de modalisateurs épistémiques dans les langues respectives.

L'analyse du corpus s'est donc faite en trois temps : d'abord un corpus exploratoire composé de 30 articles (5 articles de chacun des sous-corpus, c'est-à-dire 5 Engling, 5 Engmed, 5 Frling, 5 Frmed, 5 Noling et 5 Nomed) a été examiné dans le but d'identifier les

modalisateurs épistémiques les plus importants et afin d'avoir une idée préalable de l'emploi effectif des modalisateurs épistémiques potentiels. Cette phase a donné comme résultat une sélection de marqueurs qui se fonde sur l'usage effectif de modalisateurs épistémiques dans les trois langues, ce qui semble être un point de départ plus favorable que de fonder son analyse sur quelques marqueurs prototypiques ou sur des marqueurs antérieurement étudiés et leurs équivalents supposés dans les autres langues. En plus de fournir une liste concrète de marqueurs à étudier, l'analyse exploratoire nous a aussi permis de faire quelques observations préliminaires qui ont pu servir de base pour les études ultérieures. Les résultats de l'analyse exploratoire ainsi que les démarches méthodologiques précises qui ont été mises en œuvre lors de cette phase sont plus amplement décrits à la section 6.3.

Dans un deuxième temps, une analyse quantitative des marqueurs sélectionnés a été effectuée à l'aide de recherches automatiques. L'analyse quantitative porte sur l'ensemble des 120 articles et a pour but principal de déterminer la fréquence des modalisateurs épistémiques dans les différents sous-corpus. Les listes d'occurrences qui résultent d'une requête automatique ont été examinées manuellement afin d'éliminer les occurrences non pertinentes. Les principes de classement adoptés à cette fin sont plus amplement décrits à la section 6.4.1.

Une fois les fréquences relatives obtenues, les résultats ont été soumis à des tests statistiques afin de déterminer la significativité des différences observées. En ce qui concerne la variable « langue », nous avons d'abord effectué le test de Kruskal-Wallis sur les trois jeux de données (Engling/med, Frling/med, Noling/med), afin de déterminer si le paramètre de langue a un effet significatif sur la fréquence des modalisateurs épistémiques employés dans les textes. La réponse étant positive, nous avons procédé à la comparaison de paires de langues, utilisant le test bilatéral de Mann-Whitney<sup>2</sup>. La comparaison de paires de langues révèle laquelle ou lesquelles des paires attestent d'une différence significative entre les groupes. Quant à la variable « discipline », nous avons pu procéder directement à la comparaison entre paires, puisque les articles du corpus relèvent uniquement de deux disciplines différentes. Les résultats de l'analyse statistique seront exposés dans 6.4.2.1.

L'analyse quantitative n'a pas seulement révélé des différences de fréquence en général, elle a aussi mené à certaines observations intéressantes quant à la distribution de différents types de marqueurs. Nous reviendrons à cette question à la section 6.4.2.2.

---

<sup>2</sup> Le choix de tests statistiques est celui de l'équipe KIAP et il est justifié et expliqué dans Fløttum et al. 2006 : 42-46. Nous tenons à remercier chaleureusement Torodd Kinn, responsable de l'analyse statistique des données du projet KIAP, pour nous avoir aidée à effectuer des analyses statistiques sur nos propres données.

Dans un troisième temps, des analyses qualitatives de certains des modalisateurs les plus centraux ont été effectuées (voir les sections 7.2, 7.3, 7.4 et 7.5). Après une description des propriétés syntaxiques et sémantiques des marqueurs en question intervient un examen en détail de l'emploi de ces marqueurs. La première partie de cet examen est centrée sur le cotexte immédiat et elle a pour but d'examiner s'il y a des différences entre disciplines et entre langues en ce qui concerne les structures et les cotextes dans lesquels ces marqueurs sont employés. La deuxième partie traite du domaine fonctionnel de la modalisation épistémique et est centrée sur les fonctions pragmatiques remplies par les marqueurs. L'analyse des fonctions pragmatiques se fonde sur une lecture attentive de chaque occurrence dans son cotexte, c'est-à-dire le paragraphe où elle apparaît et dans la plupart des cas, les paragraphes voisins.

Pour finir, la partie « qualitative » inclut quelques études de cas d'articles individuels. Ces études ont été réalisées afin de mieux illustrer comment les modalisateurs épistémiques peuvent être utilisés au long d'un texte scientifique pour remplir des fonctions rhétoriques et pour contribuer à l'image d'un chercheur sérieux et honnête. Plus précisément, les études de cas permettront de mieux montrer comment les différents modalisateurs agissent ensemble pour produire l'effet recherché et comment ils interagissent avec d'autres éléments textuels, tels que les marqueurs de certitude et les structures impersonnelles. Les études de cas font l'objet de la section 7.6.



## 6 ANALYSE QUANTITATIVE

Comme signalé au chapitre 1, la présente étude a pour but de contribuer aux connaissances sur les variations disciplinaires et langagières du discours scientifique. Plus précisément, elle vise à fournir une réponse à la question de savoir dans quelle mesure l'emploi de marqueurs épistémico-modaux à fonction atténuante est déterminé par l'appartenance de l'auteur à une communauté langagière ou disciplinaire particulière. Dans le présent chapitre, il s'agira, comme le titre l'indique, des aspects quantitatifs de l'emploi de ces marqueurs, alors que le chapitre suivant traitera des aspects qualitatifs. La première section (6.1) présente les hypothèses et les questions de recherche qui ont guidé cette partie de l'analyse. Avant de passer à l'analyse exploratoire (section 6.3), nous présentons certains principes qui ont gouvernés la sélection de marqueurs à étudier (section 6.2). Ensuite, dans la section 6.4, nous continuons avec l'analyse quantitative elle-même, au niveau de la méthode (6.4.1) et des résultats (6.4.2).

### 6.1 Hypothèses et problématique de la recherche

La principale problématique de ce chapitre se focalise autour de la question de savoir dans quelle mesure les facteurs d'appartenance disciplinaire et langagière de l'auteur influencent la fréquence des modalisateurs épistémiques utilisés dans un texte. Comme déjà mentionné, nous voudrions aussi examiner l'effet potentiel du sexe de l'auteur sur la fréquence de ces marqueurs. A partir de ce qui est déjà connu sur la variation dans le discours scientifique en général (cf. *supra* chapitre 2), et dans le domaine de l'atténuation en particulier (cf. *supra* chapitre 4), nous avons pu formuler quelques hypothèses concernant la fréquence des modalisateurs épistémiques à fonction atténuante dans nos différents sous-corpus. Ces hypothèses sont présentées dans ce qui suit :

- *le rôle des disciplines.* On pourrait s'attendre à une fréquence plus élevée de modalisateurs épistémiques d'incertitude dans les articles de linguistique que dans les articles de médecine, car les recherches antérieures suggèrent que les atténuateurs sont plus fréquents dans les disciplines des sciences humaines et sociales que dans celles des sciences naturelles (Hyland 1999 a et b, Varttala 2001). Cependant, selon Varttala (1999 : 178), la discipline de la médecine a traditionnellement une fréquence d'atténuateurs relativement élevée, et la médecine et la linguistique ne sont peut-être

pas les représentants typiques de leur domaines de science respectifs (voir Fløttum et al. 2006 et Breivega 2003). Cela affaiblit quelque peu cette hypothèse de départ.

Cela étant, comme la rédaction médicale est assez standardisée, on peut s'attendre à une variation individuelle moins importante parmi les médecins-chercheurs que parmi les linguistes. Il est donc raisonnable de supposer que la fréquence des modalisateurs épistémiques variera plus entre les différents textes du corpus linguistique qu'entre ceux du corpus médical.

- *le rôle des langues.* Comme mentionné au chapitre 4, autant que nous sachions, il n'existe pas d'études comparant l'emploi d'atténuateurs dans les langues investiguées ici. Cependant, les résultats émanant d'études portant sur des phénomènes apparentés (révélant, entre autre, qu'il y a un aspect interactionnel plus saillant dans le discours scientifique anglais et une rhétorique de critique plus directe dans le discours scientifique français, cf. *supra* section 2.1) nous ont permis de postuler qu'il y aura un usage plus fréquent de modalisateurs épistémiques dans les articles anglais que dans les articles français.

Par ailleurs, nous avons postulé (cf. section 2.1) qu'il y aura un plus grand degré de similitude entre l'anglais et le norvégien qu'entre l'anglais et le français ou entre le français et le norvégien, étant donné la parenté typologique entre les langues norvégienne et anglaise, la grande influence que la culture anglo-américaine exerce sur la culture norvégienne, et le fait que les discours scientifiques anglais et norvégiens se sont avérés assez similaires à bien des égards (cf. Fløttum et al. 2006).

- *le rôle du sexe de l'auteur.* Selon la conception traditionnelle, mais peut-être obsolète, l'emploi d'expressions d'incertitude serait une caractéristique du discours féminin (cf. les chapitres 1 et 4). Or, les études empiriques divergent dans leurs résultats. Dans le projet KIAP, on a montré que le sexe de l'auteur n'a qu'un effet négligeable sur l'emploi des phénomènes linguistiques investigués (Fløttum et al. 2006), ce qui nous amène à postuler que le sexe de l'auteur n'aura pas beaucoup d'influence sur la fréquence des modalisateurs épistémiques employés dans les articles de recherche.

Nous allons voir par la suite dans quelle mesure ces hypothèses peuvent être étayées. Avant tout, il convient d'apporter certaines précisions méthodologiques par rapport à la sélection des marqueurs à retenir pour l'analyse.

## 6.2 La sélection des marqueurs : qu'est-ce qu'un marqueur épistémico-modal ?

Comme il a été mentionné au chapitre précédent, notre objet d'étude est une *catégorie sémantique*, qui ne se laisse pas définir extensionnellement, c'est-à-dire en énumérant toutes les expressions qui peuvent entrer dans la catégorie. Nous nous heurtons donc inévitablement à un *problème de sélection* : quels éléments faut-il inclure dans l'étude et lesquels faut-il exclure ? Afin de répondre à cette question, une définition précise et opératoire de la catégorie en question nous est indispensable.

Au chapitre 3, nous avons défini la catégorie de la modalité épistémique comme *l'expression des jugements du locuteur par rapport à la fiabilité de l'information transmise* (voir aussi Dendale 1994 : 25), ou, selon les termes de Lyons :

Any utterance in which the speaker explicitly qualifies his commitment to the truth of the proposition expressed by the sentence he utters, whether this qualification is made explicit in the verbal component [...] or in the prosodic or paralinguistic component, is an epistemically modal, or modalized, utterance. (Lyons 1977 : 797).

Comme nous l'avons vu au chapitre 3, ces définitions recouvrent en principe à la fois des marqueurs linguistiques et des marqueurs extralinguistiques, mais étant donné que nous nous intéressons à la catégorie de la modalité épistémique en tant que catégorie *linguistique*, et que nous étudions des *textes écrits* et non des discours oraux, certains types de marqueurs, tels que les gestes et la prosodie dans la chaîne parlée ont été exclus. Or, une question importante demeure : *dans le texte*, qu'est-ce qui est un marqueur épistémico-modal et qu'est-ce qui ne l'est pas ? À titre de réponse, nous proposons une définition opératoire de ce qu'est un marqueur épistémico-modal (ou bien un modalisateur épistémique), selon laquelle un marqueur épistémico-modal est une unité grammaticale ou lexicale qui qualifie explicitement la valeur de vérité d'un contenu informatif. En expliquant et justifiant l'exclusion et l'inclusion de différents types de marqueurs, les sous-sections suivantes visent à expliciter ce qui est entendu par « qualification explicite de la valeur de vérité d'un contenu informatif ».

### 6.2.1 Modalité épistémique et atténuation ('hedging')

Au chapitre 4, nous avons vu que, selon une conception habituelle (voir par exemple Hyland 1998 : 3, 44 et 2000 : 87), toutes sortes d'atténuateurs seraient des sous-types de la modalité épistémique. Dans une telle optique, qui n'a rien de contestable en soi, atténuation et modalité épistémique deviennent quasi-synonymes, ou plus précisément, l'atténuation correspond au

côté incertitude sur l'échelle épistémico-modale. Or, la catégorie de l'atténuation, telle qu'elle est définie par Hyland et telle qu'elle est généralement conçue, constitue une vaste catégorie relativement hétérogène. Outre les différentes catégories typiques de modalisateurs épistémiques, l'atténuation comprend plusieurs autres types de marqueurs qui ne se ramènent pas, à notre avis, à la modalité épistémique, du moins pas quand cette dernière notion est définie comme l'expression explicite de l'évaluation de la valeur de vérité de l'information transmise. Dans ce qui suit, nous présenterons brièvement certains groupes de marqueurs traditionnellement inclus sous le concept d'atténuation et nous expliquerons pourquoi ils ont été exclus de la présente étude.

#### 6.2.1.1 Modalité épistémique et adverbes « modificateurs »

Les adverbes modificateurs ('*downtoners*', voir Hyland 1998 : 139) constituent un sous-ensemble d'atténuateurs qui diminuent l'effet de ce qui est dit sans pour autant exprimer une évaluation de la valeur de vérité de la proposition. Dans l'exemple 1 ci-dessous, la question n'est pas de savoir si la place du circonstant est *fixe* ou non, mais dans quelle mesure elle est fixe. Et parallèlement, dans l'exemple 2, la question n'est pas de savoir si *le fait d'établir une hiérarchie de difficulté* (entre différents types de propositions) *dépend de la langue en question*, mais dans quelle mesure cela dépend de la langue :

1. Ces propriétés sont utilisées comme tests du caractère intra-prédicatif du circonstant. Sa place est **relativement** fixe (frling02)
2. Det å lage et slikt vanskelighetshierarki vil **stort sett** være et språkspesifikt anliggende. (noling11)  
'Le fait d'établir une telle hiérarchie de difficulté sera **en gros** une tâche qui dépend de la langue en question.'

Les expressions de ce type peuvent pourtant contribuer à une évaluation de la valeur de vérité de ce qui est dit, dans la mesure où elles modifient un modalisateur épistémique (de certitude ou d'incertitude) déjà présent :

3. Det er nemlig **ganske klart** at den ekspansjon *gi*-suffikset har fått i nordisk, nettopp forutsetter en rent negerende betydning. (noling08)  
'Il est en effet **assez clair** que l'expansion du suffixe *gi* dans les langues nordiques suppose justement un sens purement négatif.'
4. [...] l'hypothèse d'une agammaglobulinémie de Bruton est **peu probable** du fait de la survenue tardive des premières complications infectieuses [...] (frmed07)



Dans ces cas, les adverbes modificateurs (*ganske* ('assez') en 3 et *peu* en 4) font partie de la modalisation et contribuent à la force de celle-ci. Dans un exemple comme 3, où l'adverbe modificateur porte sur un marqueur épistémico-modal de certitude, cet adverbe a pour effet de rendre la proposition moins sûre. Dans ce cas particulier, la modalisation est néanmoins considérée comme se situant du côté de la certitude, à cause de l'adjectif *klart* ('clair', 'évident'), qui constitue la tête du syntagme. L'exemple 4 par contre, est inclus dans l'analyse et constitue une des 8 occurrences de *probable* repérées dans le corpus Frmed (voir le tableau 3 ci-dessous). Cela peut paraître discutable, étant donné que ce qui est jugé comme *peu probable* est loin d'être jugé comme *probable* : *peu probable* équivaut presque à *improbable*. Cependant, le fait que l'adjectif *probable* est modifié par un tel adverbe ne change pas sa valeur épistémico-modale ; il s'agit toujours d'exprimer une incertitude par rapport à ce qui est dit. Les exemples de ce type, où les expressions épistémico-modales retenues révèlent que le locuteur est assez sûr de la *fausseté* de l'information véhiculée, ne sont pas fréquents dans notre corpus. Certains d'entre eux seront commentés au chapitre 7.

#### 6.2.1.2 Modalité épistémique et approximateurs

Comme mentionné au chapitre 4, les approximateurs (*une espèce de, une sorte de*) ne sont pas inclus, car ils servent à exprimer un jugement sur l'appartenance de l'entité désignée à une catégorie lexicale plutôt qu'un jugement sur la valeur de vérité d'une proposition. On peut dire qu'ils expriment une certitude mitigée, mais celle-ci porte sur la catégorisation opérée et non sur la vérité de la proposition.

Les adverbes approximatifs comme *à peu près, approximativement* sont également exclus. Ces adverbes sont des expressions d'imprécision et ils appartiennent donc à ce qu'on appelle *le langage vague* ou *flou* ('*vague language*', cf. par exemple Varttala 2001 : 55). Or, on ne peut pas dire là non plus qu'ils transmettent une évaluation de la valeur de vérité de la proposition qu'ils accompagnent.

#### 6.2.1.3 Modalité épistémique et adverbes de fréquence (*sporadicité*)

Comme il a été mentionné au chapitre 3, on juxtapose souvent à l'échelle épistémique de probabilité une échelle épistémique de fréquence. Dans un tel dispositif, les adverbes de fréquence tels que *parfois* et *rarement* sont considérés comme des modalisateurs épistémiques. Nous préférons pourtant traiter cette échelle comme une catégorie à part, car il ne s'agit pas, dans l'emploi de ces expressions, d'une question de vérité ou de fausseté, mais d'une indication de la fréquence avec laquelle un certain état de chose se (re)produit.

#### 6.2.1.4 Modalité épistémique et compléments adverbiaux d'opinion

On distinguera par ailleurs les marqueurs épistémico-modaux et les marqueurs révélant une opinion subjective, tels que *à notre avis, selon moi, in our opinion, I think<sup>1</sup>, vi synes, etter mitt skjønn*. Ces expressions sont sans nul doute des marqueurs de subjectivité, mais ils ne situent pas forcément la proposition qu'ils accompagnent sur l'axe épistémico-modal. S'ils expriment un jugement de la valeur de vérité, cela est fait d'une manière plutôt implicite ; le marquage explicite se contente d'indiquer qu'il s'agit d'un point de vue personnel. Plusieurs chercheurs ont tout de même traité ce type d'expressions comme des marqueurs épistémico-modaux, mais en les situant sur le côté certitude de l'échelle épistémique (McEnery et Kifle 2002, Borillo 2004). De notre point de vue, en revanche, le rôle principal de ces expressions semble être de signaler que ce qui va venir ne peut pas être ontologiquement vrai ou faux mais peut être sujet à différentes opinions, ce qui expliquerait pourquoi on n'aura pas ces marqueurs en association avec des propositions dont le statut ontologique doit être soit vrai soit faux. Par exemple, on n'aura pas, ou difficilement, *à mon avis, Jean est mort<sup>2</sup>*. Or, on aura bien *à ma connaissance ou autant que je sache, Jean est mort*. À la différence des compléments adverbiaux d'opinion, ces dernières expressions sont de véritables modalisateurs épistémiques, ayant pour rôle principal de signaler le degré de certitude que le locuteur accorde au contenu de son message.

#### 6.2.2 Marqueurs de certitude, marqueurs « neutres »<sup>3</sup> et marqueurs d'incertitude

Au chapitre 3 a été souligné qu'il n'y pas de frontière nette entre les marqueurs de certitude et les marqueurs d'incertitude. Comparé à un énoncé catégorique ou assertif, un énoncé modalisé à l'aide d'un marqueur de certitude paraît souvent – mais pas toujours – moins « sûr ». Ainsi *je suis sûr que Jean est mort* paraît beaucoup moins sûr que *Jean est mort*. Pour ces raisons, Breivega (2003 : 153-157) maintient qu'il ne paraît pas utile de regrouper les marqueurs de certitude et ceux d'incertitude en deux catégories distinctes, mais qu'il vaudrait mieux les traiter ensemble sous le terme de marquage épistémique. Une telle solution prendrait en compte la potentielle part d'incertitude d'un marqueur dit de certitude, mais déclencherait en même temps un nouveau problème, à savoir la difficulté de distinguer les

---

<sup>1</sup> Il faut souligner que *think* est polysémique et peut signifier *penser* (verbe d'opinion) et *croire* (verbe épistémique).

<sup>2</sup> Il n'y a rien d'agrammatical dans cette phrase, mais il nous semble qu'elle demanderait un contexte assez particulier pour être acceptable d'un point de vue sémantique et pragmatique.

<sup>3</sup> « Neutre » est ici à prendre dans le sens de « non modal », selon une optique aristotélicienne de la modalité où les énoncés assertoriques sont conçus comme non modalisés (cf. chapitre 3).

marqueurs de certitude et les marqueurs « neutres ». Si l'on peut s'accorder sur le statut du verbe norvégien *bekreft* ('confirmer') comme marqueur de certitude dans un énoncé tel que 5, qu'en est-il du verbe *vise* ('montrer') dans un énoncé tel que 6? A-t-on affaire à un marqueur de certitude ou à un marqueur « neutre » dans un énoncé catégorique ?

5. Dataene **bekrefter** at blandingsmisbruk er svært utbredt blant injiserende heroinmisbrukere. (nomed08)  
'Les données **confirment** que l'usage de stupéfiants divers est très répandu parmi les toxicomanes qui s'injectent de l'héroïne.'
6. Gjennomgang av litteraturen **har vist** at man med PCR kan påvise HPV-DNA i vel 35 % av øre-nese-hals-karsinomer. (nomed19)  
'L'examen de la littérature **a montré** qu'avec PCR on peut démontrer la présence de HPV-DNA dans plus de 35 % des carcinomes nez-gorge-oreilles.'

Nous n'allons pas poursuivre cette problématique ici, étant donné que c'est l'aspect atténuateur de la modalité épistémique qui nous intéresse dans cette étude. Cela ne veut pas dire que nous n'avons pas éprouvé de difficultés liées à la distinction entre marqueurs de certitude et marqueurs d'incertitude. Cette distinction est loin d'être évidente, ce qu'illustreront les exemples ci-dessous :

7. [...] **il apparaît que** cette solution n'est pas sans problèmes (frling06)
8. Cependant, en considérant la solidarité du syntagme prépositionnel avec le noyau verbal, il **apparaît** possible de proposer une graduation, "d'un plus à un moins de solidarité", du centre (complément nucléaire) à la périphérie (frling02)
9. Og siden vi ikke har noen mulighet til å tolke s-passiver som adjektiviske passiver, **må** vi kunne konkludere med at norske object-experier-verb tillater passivering (noling11)  
'Et comme nous ne pouvons pas interpréter les passifs en -s comme des passifs adjectivaux, nous **devons** pouvoir conclure que les verbes norvégiens à expérimentateur objet permettent la passivation'
10. Ifølge Grimshaw er det den aspektuelle dimensjonen som bestemmer hvordan argumentene realiseres på D-struktur. Theme-argumentet til object-experier-verb **må** derfor være D-struktur-subjektet, og vi finner støtte for analysen som vi har brukt her av denne typen verb. (noling11)  
'Selon Grimshaw, c'est la dimension aspectuelle qui détermine comment les arguments se réalisent sur la structure D. L'argument *thème* du verbe à expérimentateur objet **doit** par conséquent être le sujet de la structure D, [...]

Dans la locution *il apparaît que* (cf. l'exemple 7), le verbe *apparaître* contribue à marquer le contenu de la proposition qui suit comme vérifié plutôt que comme quelque chose d'incertain : sa signification semble être de marquer que quelque chose est devenu clair, s'est

manifesté. En revanche, *apparaître* suivi d'un attribut (cf. l'exemple 8) peut marquer un degré d'incertitude (quoique relativement faible, semble-t-il), notamment si l'adjectif qui suit oriente l'interprétation vers l'*incertitude*. En l'occurrence, *apparaître* qualifie la proposition *il est possible de proposer une graduation...* en s'insérant entre le sujet formel et l'attribut à la place de *être*.

Les auxiliaires modaux *måtte* / *devoir* / *must* peuvent aussi servir à discuter la distinction entre certitude et incertitude. Dendale (1994 : 25) fait remarquer que « les grammaires françaises [...] n'associent généralement que la valeur de probabilité au verbe *devoir*<sub>E</sub><sup>4</sup> et pas celle de certitude. » Selon lui, la valeur modale de *devoir*<sub>E</sub> va pourtant « de la probabilité à la nécessité factuelle en passant par la (quasi-)certitude » (ibid. : 26). Il nous semble qu'il en va de même pour l'emploi épistémique des équivalents en norvégien et en anglais, *måtte* et *must*. L'emploi épistémique de *devoir* et ses correspondants marque toujours une inférence, mais le degré de certitude qu'on attache à la conclusion tirée varie selon les contextes. Les deux occurrences de *måtte*<sub>E</sub> repérées dans le corpus échantillon (cf. les exemples 8 et 9) ont été mises sur le côté *certitude*, non sans hésitation, car même si *måtte* épistémique et ses correspondants expriment en principe une certitude, il y a toujours un élément d'incertitude présent, ce qui explique la présence et la mise entre parenthèses de l'adverbe *quasi* dans la citation de Dendale ci-dessus.

Les difficultés liées à la distinction entre marqueurs de certitude et marqueurs d'incertitude résident donc d'une part dans la plasticité qui caractérise certains marqueurs et qui leur permet d'exprimer de la certitude aussi bien que de l'incertitude, et d'autre part dans l'indétermination de certains autres marqueurs, tels que *måtte* épistémique, qui exprime une certitude apparente tout en transmettant en même temps une impression d'incertitude.

Cet assez long développement relatif à différents groupes de marqueurs qui sont en fin de compte exclus de l'étude peut paraître quelque peu superflu. Cependant, sa principale raison d'être réside dans le fait que les notions d'atténuation ('hedging') et de modalité épistémique sont souvent confondues et mal circonscrites, comme nous l'avons vu au chapitre 4. A notre avis, une distinction plus claire entre ces catégories est utile, même nécessaire, pour éviter un élargissement trop vaste du concept de modalité épistémique. Par ce qui précède, nous espérons avoir justifié notre point de vue en mettant en relief les différences entre les divers types de marqueurs qui sont souvent subsumés sous la même catégorie. Le fait que ces types

---

<sup>4</sup> Rappelons que le <sub>E</sub> signifie qu'il s'agit de *devoir* épistémique.

de marqueurs ont des propriétés sémantiques bien différentes n'empêche pas qu'ils puissent transmettre des effets pragmatiques très similaires et être motivés par les mêmes facteurs ; par exemple, l'emploi d'un approximateur et l'emploi d'un modalisateur épistémique peuvent tous deux être motivés par un désir de la part de l'auteur de ne pas s'exprimer d'une façon trop catégorique. Nous trouvons donc tout à fait justifié de les traiter ensemble sous la catégorie de l'atténuation, qui est un phénomène pragmatique, mais nous préférons restreindre la catégorie de la modalité épistémique aux marqueurs qui expriment un jugement sur la valeur de vérité d'une proposition.

En résumé, pour qu'une expression soit incluse dans la présente analyse, il faut qu'elle exprime, de façon explicite, une évaluation de la valeur de vérité d'un contenu informatif, en ce sens que cette information acquiert le statut de « non vérifiée ».

Avant d'entreprendre l'analyse exploratoire, quelques précisions s'imposent par rapport au statut du locuteur comme *instance modalisante* ou *support modal*. C'est là le sujet de la section suivante.

### **6.2.3 Modalisation rapportée ou rapport modalisé : la question de savoir qui modalise**

Nous avons à plusieurs reprises défini la modalité épistémique comme l'expression du degré de certitude du *locuteur* par rapport à ce que celui-ci énonce. Or, il n'est pas toujours évident de déterminer qui est l'instance modalisante. Cette étude n'a pas posé comme critère que la modalisation soit explicitement attribuée à l'auteur ou à ses propres études. Ont été incluses aussi les formes au passif :

11. **It has been suggested** that differences in prognosis may relate to differences in host response reaction (Friedell et al., [1991]; Sato and Suchi, [1991]). (engmed09)

ainsi que les modalisations qui apparaissent dans un cotexte de citation, où l'auteur – d'un point de vue – rapporte les modalisations de quelqu'un d'autre dans son propre article, ou – d'un autre point de vue – effectue une modalisation sur une proposition qu'il reprend de quelqu'un d'autre. Ces modalisations rapportées ou rapports modalisés (selon le point de vue, voir discussion ci-dessous) peuvent servir au moins deux buts. D'une part, ils peuvent servir à la réfutation de théories proposées par d'autres chercheurs, comme dans l'exemple suivant :

12. Lorsque l'on rassemble des exemples pris aux différentes, on se convainc assez facilement que ces tournures existent sous leurs trois formes depuis le XV<sup>ème</sup> siècle, qu'elles ont été toutes trois exploitées à l'époque classique (cf. Haase 1914, N. Fournier 1998), qu'elles existent encore aujourd'hui et qu'on ne peut pas les placer l'une derrière l'autre dans une

perspective d'évolution historique, comme **l'ont supposé** de nombreux grammairiens. (frling01)

D'autre part, ils peuvent servir à l'appui des résultats ou des points de vue de l'auteur :

13. Comme cela **est suggéré** dans d'autres études [4],[6],[7],[8], il n'y a pas de stricte corrélation entre l'évolution virologique et immunologique chez les malades traités par IP : d'assez nombreux malades ont, en effet, une réponse immunologique prolongée alors que la CV reste détectable. (frmed10)

Dans l'exemple 12, *supposer* modalise la proposition *on peut les placer l'une derrière l'autre dans une perspective d'évolution historique*, mais ce n'est pas l'auteur qui accomplit cette supposition, ce sont d'autres grammairiens (ou plus précisément, selon le jugement de l'auteur, il s'agit d'une supposition – on ne sait pas ce qu'en ont pensé les autres grammairiens). Dans la terminologie des théories polyphoniques (voir Nølke et al. 2004), il y a là un exemple d'une *polyphonie externe* (quelqu'un d'autre que le locuteur est responsable du point de vue exprimé), et, comme l'auteur réfute la théorie de ces autres grammairiens en disant que leur supposition était fausse, il y a un *lien de réfutation* entre le locuteur et le point de vue exprimé (cf. *infra*, section 7.2.3). Dans l'exemple 13, en revanche, l'auteur renvoie aux résultats antérieurs pour étayer ses propres conclusions. C'est ce dernier type, où l'auteur accepte le point de vue rapporté, qui est le plus fréquent.

Ces types d'exemples ont traditionnellement été inclus dans des études sur l'atténuation ou la modalité épistémique (voir par exemple Hyland 1998, Varttala 1999). Cette pratique a pourtant été contestée (voir Crompton 1997) et il semble inapproprié de les inclure sans commentaire explicatif. Comme Crompton le maintient, il est dans certains cas impossible de déterminer qui est responsable de la modalisation, et il est bien possible que les auteurs cités eux-mêmes ont modalisé leur assertion et que l'auteur citant ne fait que copier la formulation originale. Dans ce cas, il s'agirait d'une modalisation rapportée. Or, le choix du verbe de citation demeure la responsabilité de l'auteur citant, et ce choix signale le degré de certitude que celui-ci accorde à ce qui est rapporté (voir Hyland 1998 : 124, 2004 : 38 ; Thompson 1996 : 522). En choisissant un verbe tel que *supposer* ou *suggérer*, l'auteur citant signale que, d'après lui en tout cas, il s'agit d'une spéculation et non pas d'un fait prouvé. Afin de savoir si les auteurs cités eux-mêmes ont modalisé leur assertion (et éventuellement de quelle manière), il faudrait se reporter au texte source, c'est-à-dire à l'article auquel l'auteur citant a fait référence. Le choix du verbe de citation, qui rend parfois les idées des autres d'une façon quelque peu infidèle, peut aussi produire l'effet inverse : alors que l'auteur

cité pourrait avoir modalisé son assertion, l'auteur citant pourrait être tellement convaincu par l'argumentation de l'auteur cité qu'il fait rapporter les idées de ce premier d'une manière plus assertive, par exemple en disant *x a montré* (ou même *x a brillamment montré*) tandis que l'auteur cité pourrait avoir écrit quelque chose comme *cette analyse semble suggérer que....*. Le verbe de citation peut donc fort bien refléter l'évaluation qu'a faite l'auteur citant plutôt que celle faite par l'auteur cité (voir aussi Varttala 1999 : 185). Dans ce cas, il s'agirait d'un rapport modalisé.

Si l'on choisit d'exclure des occurrences comme celle dans 14 ci-dessous,

14. Turning to the question of how children arrive at the adult grammar of adjunct PRO, Wexler (1992) **suggests** that the addition of this structure to the repertoire of PRO constructions in the language may be triggered by the maturation of the ability to represent a temporal operator present in the adult grammar of temporal adjuncts (Larson 1987), an operator that requires clausal structure. (engling19)

il suit logiquement qu'il faut exclure les exemples du type 11 ci-dessus aussi, où les références bibliographiques entre parenthèses renvoient à ceux qui ont suggéré (d'après l'auteur citant). Pourtant, l'exclusion de tels exemples semble quelque peu contre-intuitive. En outre, en ce qui concerne la discipline de médecine, il est raisonnable de supposer que cette exclusion mènera à un taux artificiellement bas de modalisateurs épistémiques, étant donné que les médecins-chercheurs utilisent beaucoup le passif. Nous avons donc choisi d'inclure dans l'analyse les exemples du type de ceux discutés ci-dessus.

### 6.3 Analyse exploratoire et premières observations

L'emploi d'une *méthode quantitative et automatique* exige une sélection préalable minutieuse d'expressions à étudier, car avec une analyse automatique, on ne trouvera, en principe, que ce que l'on cherche. Le processus de sélection des expressions pertinentes est donc une étape importante. Dans la présente étude, la sélection d'expressions à étudier quantitativement s'est effectuée à la suite d'une analyse manuelle scrupuleuse d'un échantillon du matériau. Comme nous l'avons mentionné au chapitre 5, ce procédé présente l'avantage de se fonder sur les expressions effectivement employées dans le corpus plutôt que sur des marqueurs considérés comme prototypiques ou sur des marqueurs antérieurement étudiés et donc retenus par tradition. Cela est particulièrement important en ce qui concerne les corpus norvégien et français, étant donné qu'il existe peu ou pas d'études empiriques sur l'emploi des modalisateurs épistémiques dans le discours scientifique de ces langues, et il n'est donc pas possible de fonder sa sélection sur des résultats quantitatifs antérieurs, comme il est possible de le faire – du moins dans une certaine mesure – pour l'anglais.

En conséquence, avant de réaliser l'analyse automatique propre, nous avons effectué une analyse exploratoire au cours de laquelle ont été examinés 30 articles de recherche, uniformément répartis sur les six sous-groupes. Dans le but de garantir sa représentativité (dans la mesure du possible), le corpus exploratoire reflète dans sa composition non seulement la distribution du corpus principal pour les différentes langues et disciplines, mais aussi pour les différentes revues. Cela pour minimiser l'influence éventuelle de contraintes éditoriales<sup>5</sup>. Les textes qui composent le corpus exploratoire sont les suivants :

Frling 01, 02, 03, 04, 05  
Frmed 02, 06, 07, 08, 09  
Noling 11, 12, 13, 14, 15  
Nomed 16, 17, 18, 19, 20  
Engling 01, 02, 03, 11, 19  
Engmed 01, 06, 11, 12, 13

Dans l'ensemble, le corpus échantillon consiste en 106 403 mots, qui se répartissent comme indiqué dans le tableau 1 :

---

<sup>5</sup> On sait par exemple que certaines revues remplacent systématiquement l'emploi d'atténuateurs « composés » tels que *may suggest* par un atténuateur simple (*suggest*) (voir Johns 2001 : 57).



**Tableau 1** : Répartition des mots dans le corpus exploratoire

	français	norvégien	anglais	total
linguistique	11 356	21 997	34 179	67 532
médecine	12 308	12 283	14 280	38 871
total	23 664	34 280	48 459	106 403

En suivant les principes de sélection décrits dans la section précédente, nous avons extrait manuellement tous les marqueurs épistémico-modaux que nous avons pu repérer dans les 30 articles. L'objectif principal était d'identifier les marqueurs épistémico-modaux les plus fréquents pour ensuite les soumettre à une analyse quantitative portant sur l'ensemble des articles.

Dans le tableau 2, les marqueurs que nous avons repérés au cours de l'analyse exploratoire sont affichés par ordre décroissant selon leur nombre total d'occurrences. Les marqueurs lexicaux (qui sont largement majoritaires) sont écrits en minuscules, alors que les moyens grammaticaux et les constructions syntaxiques à valeur épistémico-modale sont écrits en majuscules. Comme il ressort de la liste, deux moyens grammaticaux ont été repérés – il s'agit de certaines occurrences du subjonctif et du tiroir du conditionnel en français – et une construction syntaxique, à savoir certaines subordonnées conditionnelles, désignées dans la liste par OM en norvégien et IF et WHETHER pour l'anglais.

Deux choses sont à noter avant de se reporter à la liste : Premièrement, les marqueurs sont donnés dans leur *forme de base*, c'est-à-dire que les formes répertoriées dans le tableau 2 (ainsi que dans les tableaux à suivre) sont des lemmes illustrant les lexèmes qu'ils représentent. Ainsi, la forme *probable* par exemple, recouvre les occurrences de la forme fléchie *probables* aussi bien que les occurrences de la forme non fléchie. En ce qui concerne les verbes, l'infinitif sert de forme de base, et toutes les formes d'un verbe sont incluses, qu'il s'agisse des formes du présent, du passé, du futur etc. Une exception a été faite pour les auxiliaires modaux : le modal français *pouvoir* figure avec deux formes de base : *peut* couvre toutes les occurrences épistémiques des formes du présent, et *pourrait* représente les formes du conditionnel (il n'y avait pas d'occurrences épistémiques des autres temps de ce verbe dans le corpus-échantillon). En ce qui concerne le modal norvégien *kunne*, la forme du présent a été choisie comme forme de base (il n'y avait pas d'occurrences épistémiques des autres temps dans le corpus-échantillon). Quant à l'anglais, le choix des formes de base *may* et *might* s'impose en vertu des propriétés morphologiques de ces verbes.

Deuxièmement, il convient de noter que les chiffres donnés dans le tableau 2 reflètent le nombre d'occurrences *épistémico-modales*. Bien que certains marqueurs soient intrinsèquement épistémico-modaux (auquel cas le nombre d'occurrences épistémico-modales équivaudra au nombre total d'occurrences), il y en a beaucoup qui possèdent plusieurs valeurs potentielles. En général, le cotexte favorisera une des lectures et permettra ainsi d'étiqueter le marqueur comme épistémico-modal ou non-épistémico-modal. Les principes qui ont guidé ce processus d'interprétation seront décrits plus en détail dans la section 6.4.1.3.

**Tableau 2 :** Marqueurs repérés dans le corpus échantillon avec leur nombre d'occurrences

Corpus français	total	ling	med	Corpus norv.	total	ling	med	Corpus angl.	total	ling	med
sembler	22	12	10	kan	22	10	12	may	41	20	21
pourrait	10	0	10	synes	17	10	7	suggest	25	17	8
paraître	8	7	1	sannsynligvis	11	5	6	appear	22	22	0
probable	5	0	5	mulig	9	5	4	assume	22	22	0
probablement	5	0	5	se ut	9	8	1	might	20	17	3
sans doute	5	4	1	anta	8	5	3	seem	16	16	0
peut	4	1	3	kanskje	7	7	0	perhaps	7	6	1
suggérer	4	1	3	trolig	7	6	1	indicate	6	2	4
SUBJONCTIF	4	3	1	tyde på	7	2	5	could	5	3	2
apparaître	3	3	0	antagelse	6	5	1	expect	4	0	4
estimer	3	0	3	muligens	5	1	4	believe	3	3	0
hypothèse	3	1	2	OM	4	0	4	plausibly	3	2	1
-RAIT	3	2	1	sies å	4	4	0	possible	2	0	2
possible	2	0	2	vite (NEG)	4	2	2	estimate	2	0	2
supposer	2	2	0	neppe	3	2	1	hypothesize	2	1	1
à notre connaissance	1	0	1	nok	3	2	1	IF	2	2	0
apparemment	1	1	0	regne med	3	3	0	likely	3	1	2
croire	1	0	1	foreslå	2	2	0	probably	2	1	1
évident (NEG) <sup>6</sup>	1	1	0	hypotese	2	2	0	proposal	2	2	0
éventuellement	1	1	0	kan tenkes	2	0	2	think	2	2	0
indiquer	1	0	1	peke i retn. av	2	2	0	unlikely	2	0	2
peut-être	1	1	0	sannsynlig	2	0	2	WHETHER	2	2	0
supposément	1	1	0	ansees	1	0	1	would	2	2	0
suspecter	1	0	1	antagelig	1	1	0	apparently	1	1	0
vraisemblablement	1	0	1	indikerer	1	0	1	belief	1	0	1
<b>Total</b>	<b>93</b>	<b>41</b>	<b>52</b>	mistenker	1	1	0	hypothesis	1	1	0
<b>Fréquence rel.</b>	<b>3,9</b>	<b>3,6</b>	<b>4,2</b>	mulighet	1	1	0	interpret	1	0	1
				skulle kunne	1	1	0	plausible	1	1	0

<sup>6</sup> « NEG » signifie qu'il s'agit de formes négatives : *il n'est pas tout à fait évident* (voir l'exemple17)

tror	1	1	0
usikkerhet	1	0	1
virke som	1	1	0
<b>Total</b>	<b>148</b>	<b>89</b>	<b>59</b>
<b>Fréq. rel.</b>	<b>4,3</b>	<b>4,0</b>	<b>4,8</b>

possibility	1	0	1
presumably	1	0	1
should	1	1	0
speculate	1	1	0
speculation	1	0	1
submit	1	1	0
suggestion	1	1	0
will	1	1	0
<b>Total</b>	<b>210</b>	<b>151</b>	<b>59</b>
<b>Fréq. rel.</b>	<b>4,3</b>	<b>4,4</b>	<b>4,1</b>

Considérons d'abord les résultats pour le corpus français. On remarquera que, bien qu'une série de marqueurs soient repérés, la plupart ne sont utilisés qu'une ou deux fois et il y a peu de marqueurs véritablement récurrents. En fait, seul *sembler* paraît récurrent à proprement parler : ce verbe apparaît beaucoup plus souvent que les autres marqueurs et on le trouve dans presque tous les articles<sup>7</sup>. Il est manifestement le marqueur épistémico-modal le plus fréquent dans le corpus échantillon. Viennent ensuite *pourrait* et *paraître* qui semblent être relativement récurrents, mais qui apparaissent dans un nombre restreint d'articles et uniquement dans une des disciplines : sept sur huit occurrences de *paraître* apparaissent dans les articles de linguistique Frling01 et Frling04 et les occurrences de *pourrait* relèvent toutes des articles de médecine Frmed02 et Frmed07. *Probable*, *probablement* et *sans doute* semblent aussi spécifiques à l'une des disciplines, mais les chiffres sont si peu élevés qu'il n'est pas possible de savoir dans quelle mesure cette première observation reflète l'usage effectif.

La désinence du conditionnel *-rait* s'est avérée peu fréquente en tant que modalisateur épistémique, ce qui n'est peut-être pas étonnant, puisque cette valeur du conditionnel est typique du discours journalistique mais n'est pas spécifiquement associée au discours scientifique. Il faut cependant noter que les occurrences de *-rait* comme désinence d'un verbe épistémique n'ont pas été comptées comme des marqueurs épistémico-modaux en elles-mêmes, bien qu'elles contribuent sans doute à renforcer l'aspect d'incertitude véhiculé par le verbe qu'elles modifient. L'investigation du corpus exploratoire suggère que *-rait* est employé comme marqueur épistémico-modal notamment avec le verbe *pouvoir*, et rarement avec d'autres verbes. Pour cette raison, nous avons choisi de retenir pour l'analyse quantitative le conditionnel de *pouvoir* plutôt que la désinence *-rait* en général.

A la différence du corpus français, l'exploration du corpus norvégien révèle que plusieurs modalisateurs épistémiques sont relativement récurrents dans les textes examinés (par exemple *kan*, *synes*, *sannsynligvis*, *mulig* et *se ut*). Bien que le marqueur *kan* soit plus fréquent que les autres, il n'est pas un marqueur aussi largement dominant que l'est *sembler* dans le corpus français. Il apparaît aussi que les différences entre disciplines sont moins importantes dans le corpus norvégien que dans le corpus français : les marqueurs les plus fréquents sont relativement récurrents dans les deux disciplines et leur emploi ne paraît pas se limiter à une discipline spécifique (à l'exception de *se ut*, qui est quasi-absent dans le corpus-échantillon médical). Cependant, les variations individuelles sont considérables, du moins en

---

<sup>7</sup> Les tableaux dans l'appendice 2 montrent la distribution des différents marqueurs entre les articles.

ce qui concerne l'emploi du verbe *synes* : les 10 occurrences de ce verbe qui ont été repérées dans le corpus de linguistique relèvent toutes d'un même article, à savoir de Noling13, et six sur sept occurrences dans le corpus médical relèvent de l'article Nomed18.

Pour ce qui est du corpus anglais, la situation est quelque peu différente. Comme pour le norvégien, une série de marqueurs se révèlent relativement récurrents (*may, suggest, appear, assume, might, seem*). Cependant, plusieurs de ces marqueurs paraissent spécifiques à la discipline de linguistique. Il semble donc qu'il y ait une différence sensible entre les deux disciplines, non pas en ce qui concerne la fréquence relative des modalisateurs épistémiques, mais en ce qui concerne le choix de marqueurs.

En règle générale, les marqueurs qui apparaissent avec une valeur épistémico-modale cinq fois ou plus ont été retenus pour une analyse quantitative de l'ensemble du corpus, ce qui nous donne *sembler, pourrait, paraître, probable, probablement*, et *sans doute* pour le français, *kan* ('peut'), *synes* ('sembler'), *sannsynligvis* ('probablement'), *mulig* ('possible'), *se ut (til/som)* ('paraître'), *anta* ('supposer'), *kanskje* ('peut-être'), *trolig* ('vraisemblablement'), *tyde på* ('suggérer'/'indiquer'), *antagelse* ('supposition') et *muligens* ('peut-être'/'éventuellement') pour le norvégien et *may, suggest, appear, assume, might, seem, perhaps, indicate* et *could* pour l'anglais.

A part les marqueurs relativement fréquents, les listes dans le tableau 2 contiennent toute une gamme d'expressions qui ne se reproduisent qu'une ou deux fois. Pour des raisons de lisibilité du tableau, elles apparaissent dans la liste isolées et détachées de leur cotexte, et il peut être difficile d'en voir la valeur épistémique. À titre d'illustration, nous donnons donc ci-dessous quelques exemples d'occurrences qui ont été classées comme épistémico-modales. Que ces exemples aient été étiquetés épistémico-modaux ne veut évidemment pas dire que nous avons accordé la même signification à toutes les occurrences du marqueur en question ; le codage des occurrences dépend largement du cotexte.

16. Det har blitt **foreslått**, blant annet av Belletti & Rizzi (1988), at psyk-verb av object-experiencer-typen ikke utdeler noen ekstern rolle. [...] Gitt at Belletti og Rizzis teori stemmer, vil aktivsetninger med object-experiencer-verb ha et spor (noling11)

'Il a été **suggéré**, entre autres par Belletti & Rizzi (1988), que les verbes psychologiques à expérientiateur objet ne donnent pas de rôle externe [...] Pourvu que la théorie de Belletti et Rizzi soit correcte, les phrases à l'actif avec verbes à expérientiateur objet porteront une trace'

17. Certes, il **n'est pas tout à fait évident** que ces contraintes **soient** valables pour tous les mots communément considérés comme des prépositions. C'est là une question qui mérite des recherches plus approfondies (frling04)

18. **Whether** the precise interaction of multiple factors in performance can be further predicted and constrained **remains to be seen.** (engling03)

Malgré ses avantages, cette méthode de sélection, fondée sur l'usage effectif, comporte quelques limites. Comme l'analyse exploratoire n'englobe qu'un quart des articles, il est possible que certaines expressions courantes nous aient échappé, et il va de soi qu'avec un petit échantillon de cinq articles de chacun des sous-corpus, cette méthode est relativement vulnérable aux variations internes qui existent dans les différents sous-corpus. Si d'autres articles avaient été sélectionnés pour l'analyse exploratoire, la liste de marqueurs à étudier aurait pu être différente, notamment dans la mesure où les différents auteurs ont leurs marqueurs préférés, qui seraient donc fréquents chez certains, mais absents ou quasi-absents chez d'autres. Pour cette raison, nous avons jugé utile de compléter la sélection basée sur l'usage effectif dans le corpus échantillon avec quelques marqueurs supplémentaires. Il s'agit particulièrement de marqueurs français, ce qui s'explique d'une part par le nombre limité de marqueurs récurrents dans le corpus français – rendant difficile une sélection fondée sur la fréquence – et d'autre part par la (quasi)-absence de certains marqueurs courants dont on pourrait dire qu'ils sont les prototypes des marqueurs épistémico-modaux. Ainsi, certains marqueurs supplémentaires ont été retenus pour l'analyse quantitative, soit à cause de leur fréquence (il s'agit de *peut* et de *suggérer*, qui ont chacun quatre occurrences<sup>8</sup>), soit à cause de leur caractère prototypique ou de leur intérêt dans une perspective comparative (*possible*, *peut-être*, *peut* et *indiquer* pour le français, *possible* et *probably* pour l'anglais), soit à cause d'une combinaison de ces critères (*peut*). En ce qui concerne l'adjectif *possible*, l'adverbe *peut-être* et la forme verbale *peut*, on aurait attendu que ces marqueurs soient parmi les modalisateurs épistémiques les plus fréquents, étant donné que ce sont eux, nous semble-t-il, qui figurent le plus souvent dans les exemples des grammaires ou des études antérieures servant à illustrer ce qu'est la modalité épistémique. Néanmoins, les résultats issus du corpus exploratoire suggèrent que ces expressions ne sont pas fréquentes dans les articles français. Il sera intéressant de voir si l'analyse quantitative du grand corpus peut confirmer l'observation selon laquelle ces marqueurs ont peu de poids comme modalisateurs épistémiques. L'analyse exploratoire révèle que leurs correspondants en norvégien sont relativement fréquents. Outre

---

<sup>8</sup> Le subjonctif par contre, sera laissé de côté, bien qu'il apparaisse autant de fois que *peut* et *suggérer*. D'une part, le subjonctif ne se prête que très difficilement à une analyse automatique, les formes étant très diversifiées. D'autre part, comme nous l'avons vu dans la section 3.3, le subjonctif en lui-même ne sert que très rarement à marquer la modalité épistémique, et, dans tous les cas repérés dans le corpus échantillon, il s'agit d'un subjonctif qui suit un autre marqueur épistémico-modal. Ce sont donc les marqueurs dans la proposition principale plutôt que les formes au subjonctif en elles-mêmes qui sont retenus pour cette étude.

le fait qu'en tant que prototypes, elles sont particulièrement intéressantes et méritent une étude plus détaillée, il serait donc intéressant d'étudier ces expressions dans une perspective contrastive. Dans le même ordre d'idées ont été inclus aussi l'adjectif anglais *possible* et l'adverbe *probably*. Finalement, le verbe français *indiquer* a été retenu pour une analyse ultérieure, dans le but de le comparer avec le correspondant en anglais *indicate*, qui figurait tout en haut sur la liste de Hyland (1998 : 149, cf. *supra*, section 4.2) ainsi qu'avec la locution verbale en norvégien *tyde på*.

Les études de cas serviront aussi à compenser la sensibilité de l'analyse exploratoire aux variations internes de chaque groupe : par exemple, dans le corpus échantillon, les particules modales en norvégien se sont avérées peu fréquentes. Les recherches automatiques effectuées dans le but de contrôler cette observation ont révélé que ces marqueurs s'utilisent, mais dans un nombre restreint d'articles. Pour compenser le manque d'information sur les particules modales, un de ces articles a été choisi pour une étude de cas.

En plus de fournir une liste de marqueurs récurrents, l'analyse exploratoire donne une idée de la fréquence des modalisateurs épistémiques dans les articles de recherche. La fréquence relative varie entre 3,6 pour mille dans le corpus Frling et 4,8 dans le corpus Nomed. Les linguistes anglophones semblent utiliser plus de modalisateurs épistémiques que les linguistes norvégiens, qui à leur tour en utilisent plus que les linguistes francophones<sup>9</sup>. En médecine, par contre, ce sont les chercheurs norvégiens qui utilisent le plus d'atténuateurs. Les articles des médecins-chercheurs attestent d'une fréquence de marqueurs épistémico-modaux plus élevée que les linguistes, sauf dans le corpus anglais, où la fréquence est légèrement plus élevée chez les linguistes. On verra dans ce qui suit si l'analyse quantitative peut confirmer ou infirmer ces tendances.

### *Observations générales*

Avant de passer à l'analyse quantitative, signalons quelques traits généraux observés lors de l'analyse exploratoire. Le premier est que les médecins-chercheurs semblent faire usage, dans une plus large mesure que les linguistes, d'une stratégie d'atténuation qui consiste à faire référence à des connaissances limitées. Ce manque de connaissances vaut le plus souvent pour le domaine d'études en général, et les chercheurs en médecine semblent donc posséder une

---

<sup>9</sup> Etant donné que les auteurs des articles rédigés en français et en anglais ne sont pas nécessairement de nationalité française ou anglaise (ils peuvent être, par exemple, belges, américains etc.), il ne serait pas approprié d'utiliser les adjectifs *français* et *anglais*. Nous avons opté pour les termes *francophone* et *anglophone*, mais nous soulignons qu'ici, ces adjectifs ne signifient rien d'autre que, respectivement, « personne qui écrit en français » et « personne qui écrit en anglais ».



base commune de connaissances dont ils soulignent néanmoins les lacunes, à combler par des recherches ultérieures. La référence à des connaissances limitées va souvent de pair avec l'emploi de modalisateurs épistémiques d'incertitude (cf. *infra* 7.4.6, 7.6.4) :

19. Årsaken til den høye letaliteten blant pasienter som ble operert med Billroth II eller Billroth I (fire av seks pasienter døde), er **ikke kjent**. En sykere pasientgruppe med høyere alder, lengre operasjonstid og operatører med mindre trening i ventrikkkirurgi kan være mulige forklaringer. (nomed20)

'La cause du taux de léthalité élevé chez les malades opérés selon Billroth II ou Billroth I (quatre sur six malades sont morts), **n'est pas connue**. Un groupe de malades plus malades et plus âgés, une durée d'opération plus longue et des opérateurs moins exercés dans la chirurgie ventriculaire peuvent être des explications possibles.'

Cet exemple contient une série de modalisateurs épistémiques (cf. *kan*, *mulige*), mais le centre d'intérêt ici est la phrase *årsaken til ... er ikke kjent* ('la cause ... **n'est pas connue**'). La référence aux connaissances limitées est une stratégie d'atténuation signalée par Hyland (1998 : 141-148), mais qui n'entre pas dans notre définition de la modalité épistémique. Cette observation montre que, bien que la modalisation épistémique soit la catégorie dominante de l'atténuation, le nombre de modalisateurs épistémiques dans un texte n'est qu'une indication de la fréquence d'atténuateurs, dans la mesure où il existe aussi d'autres stratégies d'atténuation.

La seconde remarque concerne ce qu'on appelle – dans la littérature anglaise sur l'atténuation – *clusters*, c'est-à-dire des « entassements » de modalisateurs épistémiques. Plus nombreux sont les modalisateurs épistémiques utilisés, moins « certaine » apparaît la proposition qu'ils accompagnent. De tels entassements semblent être beaucoup plus fréquents dans les corpus norvégiens. Exemples :

20. En sykere pasientgruppe med høyere alder, lengre operasjonstid og operatører med mindre trening i ventrikkkirurgi **kan** være **mulige** forklaringer. (nomed20)

' Un groupe de malades plus malades et plus âgés, une durée d'opération plus longue et des opérateurs moins exercés dans la chirurgie ventriculaire **peuvent** être des explications **possibles**.'

21. En kooperativ fullføring blir introdusert utenfor et potensielt turskiftepunkt og **kan** dermed **se ut til** å bryte med turtakingsreglene. (noling15)

'Un accomplissement coopératif est introduit en dehors d'un moment potentiel pour un échange de tour de parole et, par conséquent, **peut sembler** rompre avec les règles de la prise de tour de paroles.'

Il reste à déterminer dans quelle mesure cette observation préliminaire peut être étayée. Cela sera un enjeu pour le chapitre 7, où nous reviendrons à la question des '*clusters*'. Nous y

examinerons dans quelle mesure les marqueurs norvégiens et français les plus fréquents sont combinés avec d'autres marqueurs épistémico-modaux.

Une troisième remarque porte sur le fait que l'anglais et le norvégien semblent disposer d'une série plus large de moyens linguistiques pour exprimer la modalité épistémique que ne le fait le français : les textes anglais et norvégiens attestent d'un emploi plus varié de modalisateurs épistémiques. Il n'est cependant pas sûr qu'il s'agisse d'une véritable différence : l'observation pourrait s'expliquer par la longueur des articles. Comme le montre le tableau 1, le nombre de mots dans le corpus exploratoire est plus élevé dans les corpus anglais et norvégien que dans le corpus français, et il est donc raisonnable de penser que la variation y est plus importante. En outre, il est possible qu'en tant que locuteur natif du norvégien, nous reconnaissons plus facilement un modalisateur épistémique en norvégien qu'en français ou en anglais.

## 6.4 Analyse quantitative

Le grand avantage d'une méthode quantitative et automatique est, évidemment, de pouvoir travailler sur des corpus d'une taille importante et obtenir ainsi des résultats plus représentatifs et généralisables. Certes, il y a aussi des inconvénients : on ne trouve que ce que l'on cherche, et les recherches automatiques par ordinateur se restreignent à des marqueurs aisément identifiables. On risque de fait de perdre un peu de vue l'ensemble, et cela peut contribuer à une insistance trop grande sur les marqueurs récurrents et les plus reconnaissables. Les études de cas (voir le chapitre 7) sont censées compenser partiellement cet inconvénient. Elles permettront de prendre en compte la structure argumentative des textes investigués ainsi que toutes les stratégies de modalisation utilisées. Dans le présent chapitre, en revanche, ce sont les résultats quantitatifs qui seront au centre du propos. Avant la présentation des résultats, quelques mises au point s'imposent concernant les principes méthodologiques adoptés.

### 6.4.1 Méthode

Dans cette section seront présentés les critères de sélection de marqueurs, la manière dont ont été effectuées les recherches automatiques, et les principes adoptés pour résoudre la polysémie des marqueurs polyvalents. Nous passerons assez vite sur le premier point, étant donné que les critères de sélection ont été plus amplement décrits à la section précédente.

#### 6.4.1.1 Critères de sélection

Comme déjà mentionné, la sélection d'expressions à étudier quantitativement s'est effectuée selon les critères suivants:

- 1) *Fréquence*, telle qu'elle est indiquée par l'analyse exploratoire : les modalisateurs épistémiques qui se sont avérés les plus fréquents dans les 30 articles étudiés manuellement ont été soumis à une analyse quantitative et automatique portant sur l'ensemble des articles.
- 2) *Caractère prototypique et/ou pertinence contrastive* : certaines expressions dont on dirait qu'elles sont les prototypes de marqueurs épistémico-modaux, ont été soumises à une analyse automatique bien qu'elles ne soient pas fréquentes dans les 30 articles analysés manuellement. Le but est de compléter la liste des marqueurs récurrents

fournie par l'analyse exploratoire. Certains marqueurs ont aussi été sélectionnés parce qu'il s'avère intéressant de les étudier dans une perspective contrastive.

Le critère de fréquence a donné comme résultat les marqueurs suivants (cf. tableau 2 ci-dessus) :

pour le français<sup>10</sup> :

*sembler* (22)  
*pourrait* (10)  
*paraître* (8)  
*probable* (5)  
*probablement* (5)  
*sans doute* (5)

pour le norvégien :

*kan* (22)  
*synes* (17)  
*sannsynligvis* (11)  
*mulig* (9)  
*se ut til/som* (9)  
*anta* (8)  
*kanskje* (7)  
*trolig* (7)  
*tyde på* (7)  
*antagelse* (6)  
*muligens* (5)

et pour l'anglais :

*may* (41)  
*suggest* (25)  
*appear* (22)  
*assume* (22)  
*might* (20)  
*seem* (16)  
*perhaps* (7)  
*indicate* (6)  
*could* (5)

Selon le second principe (caractère prototypique et/ou pertinence contrastive), nous avons complété la liste par les marqueurs suivants :

pour le français :

*peut*  
*suggérer*  
*peut-être*  
*possible*  
*indiquer*

et pour l'anglais :

*possible*  
*probably*

Rappelons que ces formes sont des *formes de base* (cf. section 6.3), et l'analyse inclut aussi les autres formes du même lexème. La forme de base *pourrait* par exemple, recouvre les formes *pourrais*, *pourrait*, *pourrions*, *pourriez* et *pourraient*, et la forme *anta* ('supposer') recouvre toutes les formes du verbe *anta* ('supposer'), par exemple *antar*, *antok* et (*har*) *antatt*, mais elle ne recouvre pas les occurrences de l'adjectif *antatt* ('supposé/présumé'). Plus précisément, la forme de base recouvre le radical avec toutes les diverses désinences qui peuvent s'y rattacher, mais les formes contenant un préfixe (par exemple *improbable*) ne sont pas prises en charge par la forme de base. Une exception a été faite pour le nom *antagelse*

---

<sup>10</sup> Entre parenthèses le nombre d'occurrences épistémico-modales au total, c'est-à-dire dans les articles de médecine et de linguistique réunis.

(‘supposition’), qui intègre aussi une occurrence *grunnantagelsen* (‘supposition de base’). Dans le cas où un marqueur est lui-même modifié, par exemple par la négation ou un adverbe tel que *peu* dans l’exemple 4, cela n’a pas mené à son exclusion, pourvu que la valeur reste épistémico-modale.

#### 6.4.1.2 Les recherches sur ordinateur

Dans ce qui suit, nous donnerons quelques exemples pour illustrer comment les recherches automatiques ont été effectuées. Les requêtes transmises au moteur de recherche ont été élaborées avec soin. Préférentiellement, la chaîne de caractères qui sert de requête permet d’inclure toutes les formes différentes du marqueur en question sans générer trop de « bruit ». Ainsi, pour rechercher *sembler*, nous avons choisi le mot de recherche *sembl.\*<sup>11</sup>*, qui permet d’identifier les diverses formes de ce verbe, les participes inclus, mais aussi les occurrences de l’adjectif *semblable*. Ces occurrences ont été manuellement mises de côté. Pour rechercher les occurrences de *suggérer*, il a fallu lancer deux recherches, *suggér.\** et *suggèr.\**, à cause de l’alternance de *é* et *è* qui caractérise ce type de verbes. Pour les unités lexicales composées de plusieurs lexèmes ou particules, il a fallu prendre en compte le fait que certains mots peuvent s’interposer entre les éléments constitutifs du marqueur. Ainsi, en recherchant les occurrences du marqueur norvégien *se ut (til/som)* (‘sembler’ / ‘paraître’), nous avons lancé les recherches “*s.\*<sup>12</sup> + ut + til*” et “*s.\* + ut + som*” en permettant jusqu’à 15 mots de s’interposer entre les deux premiers éléments pour éviter d’exclure des occurrences du type *ser dessuten ut til*, *ser det altså ut til* et même

22. Uansett årsak **ser** ikke den høyere injeksjonsfrekvensen og det større forbruket **ut til** å resultere i en høyere risiko for narkotikadødsfall [...]. (nomed 08)  
‘Quelle que soit la cause, la fréquence plus élevée d’injections et les dosages plus grands ne **paraissent** pas engendrer un risque plus élevé de surdosages [...].’

Pour les marqueurs norvégiens il a fallu prendre en considération les deux versions qui coexistent en norvégien : le *nynorsk* et le *bokmål*. Afin d’identifier toutes les occurrences de l’adverbe *trolig* (‘vraisemblablement’), il a donc fallu lancer la recherche *tr.\*l.\*g*, pour permettre l’inclusion des deux formes du *nynorsk* : *truleg* et *trulig*, aussi bien que la forme du *bokmål* : *trolig*.

---

<sup>11</sup> Le point et l’astérisque instruisent le programme de recherche à inclure toute chaîne qui commence par ‘sembl’. Par contre, le mot de recherche ‘semble’ (sans point-astérisque) aurait inclus uniquement les chaînes ‘semble’, excluant les formes semblent, semblé etc.

<sup>12</sup> Le *e* est omis pour permettre l’inclusion d’occurrences du passé : *så ut til/som*

Toutes les listes fournies en réponse par l'ordinateur ont été contrôlées afin de mettre de côté les occurrences qui n'entraient pas dans le cadre de l'étude. Ainsi, les occurrences de *semblable* ont été supprimées de la liste *sembler* et les occurrences de *antall* ont été supprimées de la liste *anta* dès le début. Une fois ce « bruit » éliminé, chaque occurrence a été étudiée dans son contexte afin de déterminer si le sens était un sens épistémico-modal ou non.

#### 6.4.1.3 Le classement des attestations

Comme il a déjà été mentionné, un grand nombre des marqueurs sélectionnés sont polysémiques et connaissent une multiplicité d'emplois. Selon Crompton (1997 : 278), la valeur atténuante n'est jamais un trait inhérent d'une unité lexicale, mais un effet qui naît de l'ensemble d'un énoncé. A notre avis, il existe certaines unités lexicales qui sont intrinsèquement épistémico-modales ; parmi les marqueurs étudiés ici nous dirons que *sembler*, *probablement*, *peut-être*<sup>13</sup> et *sans doute* en français, *seem*, *probably* et *perhaps* en anglais, *sannsynligvis*, *kanskje*, *muligens*, *trolig*, *tyde på* et *antagelse* en norvégien sont intrinsèquement ou quasi intrinsèquement épistémico-modaux<sup>14</sup>. Il reste pourtant une vingtaine de marqueurs susceptibles d'être utilisés dans plusieurs acceptions, dont certaines pourraient être assez proches et parfois difficiles à distinguer. Par conséquent, le classement des attestations a été un processus important mais complexe. Dans ce qui suit nous présenterons les principes qui ont guidé ce travail d'interprétation.

Avant de continuer avec le problème de la polysémie, une petite précision s'impose. Nous avons dit, à plusieurs reprises, qu'un modalisateur épistémique qualifie la valeur de vérité d'une proposition ou d'un contenu informatif. Cela implique que, pour qu'un marqueur soit étiqueté comme un modalisateur épistémique, il doit y avoir quelque chose à qualifier, un contenu informatif sur lequel on se prononce. Le marqueur en question ne doit pas nécessairement porter syntaxiquement sur une proposition entière, mais une telle proposition doit pouvoir être restituée. Dans l'exemple 23 ci-dessous, l'adjectif *possible* a été étiqueté comme épistémico-modal car il est censé qualifier la proposition « sous-jacente » *l'augmentation du temps passé à l'intérieur est une autre cause*. L'adjectif *possible* porte donc sur un contenu informatif :

---

<sup>13</sup> *Peut-être* et ses équivalents *perhaps* et *kanskje* peuvent parfois prendre une lecture sporadique (voir aussi Cotte 2002 : 4). Il nous semble pourtant qu'une telle lecture n'exclut jamais complètement une interprétation épistémique ; il peut y avoir indétermination entre les deux valeurs, qui sont, comme nous l'avons vu (cf. section 3.1.2.3), souvent traitées ensemble sous le terme de modalité épistémique à cause de leur parenté.

<sup>14</sup> Bien que ces marqueurs soient considérés comme intrinsèquement épistémico-modaux, les listes d'attestations fournies par l'ordinateur ont été contrôlées manuellement afin d'éliminer entre autres d'éventuelles occurrences d'usage métalinguistique.

23. Une autre cause **possible** est l'augmentation du temps passé à l'intérieur, ce qui augmente l'exposition aux allergènes tels que les acariens et les animaux domestiques. (frmed02)

Dans l'exemple suivant, en revanche, un tel contenu n'est pas présent, et le nom *antagelse* ne peut donc pas être vu comme ayant une fonction modalisante :

24. Her dreier det seg jo nettopp om å undersøke om analysen vår passer inn i lingvistens øvrige **antagelser** om språk. (noling06)  
'Ici il s'agit justement d'examiner si notre analyse s'harmonise avec les autres **suppositions** du linguiste sur le langage.'

Revenons maintenant à la question de la polysémie. Certains des marqueurs polysémiques possèdent des acceptions bien différentes, faciles à distinguer l'une de l'autre. Tel est le cas pour le verbe norvégien *synes*, par exemple, qui s'emploie tantôt comme un semi-auxiliaire épistémique dans le sens de *sembler* (exemple 25), tantôt comme un verbe d'opinion (exemple 26). Les occurrences de ce dernier type (peu nombreuses par ailleurs) ont été exclues sans difficultés :

25. Den totale komplikasjonsraten **synes** å ligge høyere i vår studie enn det som publiseres i de fleste serier (11). (nomed05)  
'Le taux total de complications **semble** être plus élevé dans notre étude...'

26. Vi **synes** at denne informasjonen er viktig, ettersom de preoperative plagene til enkelte pasienter er ubetydelige. (nomed05)  
Nous **trouvons** cette information importante, puisque les maux préopératoires de certains patients sont insignifiants.'

Pareillement, le verbe anglais *suggest* (comme l'équivalent français *suggérer*) s'utilise tantôt pour exprimer une modalité épistémique (exemple 27), tantôt dans des acceptions non-épistémiques, comme un synonyme de *propose* ('proposer') ou de *recommend* ('recommander' / 'conseiller') (exemples 28-29) :

27. Clinical reports **suggest** that it has a calming effect upon patients (engmed14)

28. Following his lead, we **suggest** this question for investigation (engling18)

29. Given the inconsistency among studies of diet and ovarian cancer and the inability to infer causality from associations found in retrospective epidemiologic studies, it may be premature to **suggest** that women modify their diets based on the evidence from our study. (engmed10)

Cependant, le classement d'occurrences n'est pas toujours simple et direct, étant donné que les différentes acceptions d'un marqueur polysémique sont toutes plus ou moins apparentées et qu'il n'y a donc pas toujours une distinction nette entre deux acceptions. Par exemple, l'auxiliaire *may* et ses correspondants *kunne* et *pouvoir* peuvent exprimer toute une série de valeurs : la capacité, la possibilité radicale sous-déterminée, la permission, la sporadicité et la possibilité épistémique. Parfois il y a indétermination entre deux ou plusieurs valeurs : plusieurs valeurs coexistent et contribuent ensemble à l'interprétation, comme dans l'exemple suivant, où les valeurs de sporadicité, capacité et épistémicité semblent coexister :

30. The goal for this study was to try to locate in the corpus cross-cultural differences which **may** ultimately influence the efficacy of the letters of application written by the American and European participants. (engling14)

Le cotexte aide souvent à sélectionner le sens à privilégier, comme dans l'exemple ci-dessous, où la seconde phrase sert à orienter l'interprétation vers le sens de capacité plutôt que vers celui d'éventualité :

31. de Hoop's (1996) analysis of presuppositionality **may** explain this issue. de Hoop shows that in languages such as Dutch with overt scrambling of presuppositional DPs, scrambling is optional (engling01)

Bien que le cotexte comporte souvent des indices qui aident à privilégier l'un ou l'autre sens, il est parfois utile d'avoir recours à des outils plus concrets. Trois tests ont été appliqués pour faciliter le travail de décision sur la valeur des marqueurs polysémiques :

Le premier test, utilisé aussi par Vihla (2000 : 212), consiste à examiner si le marqueur en question peut être remplacé par un marqueur intrinsèquement épistémico-modal tel que *vraisemblablement*, *peut-être*, *possibly*, *maybe*, *kanskje*, *trolig* etc. Si une telle substitution est possible sans altération appréciable de l'interprétation de l'énoncé, cela plaide en faveur d'une lecture épistémico-modale. Dans le cas contraire, une lecture épistémico-modale paraît moins appropriée. Considérons les deux paires d'exemples ci-dessous : dans 32, *might* semble remplaçable par *perhaps*. Dans 33, au contraire, *perhaps* ne semble pas transmettre le même sens que *might*. En revanche, *might* est ici remplaçable par *can*, sans que cela entraîne un changement sensible de sens, ce qui justifie une lecture non-épistémique :

32. The same results would pertain for paraphrases using 'it is false that P' and some readers **might** find this paraphrase more effective in prompting the intuitions I am trying to elicit here.



- 32' The same results would pertain for paraphrases using `it is false that P' and some readers will **perhaps** find this paraphrase more effective in prompting the intuitions I am trying to elicit here. (engling08)
33. We have adjusted for key factors that relate to ovarian cancer risk such as age, parity, oral contraceptive use and family history, as well as factors such as body mass index and total calories consumed which **might** be associated with diet. (engmed10)
- 33' ?<sup>15</sup>We have adjusted for key factors that relate to ovarian cancer risk such as age, parity, oral contraceptive use and family history, as well as factors such as body mass index and total calories consumed which are **perhaps** associated with diet.

Deuxième test : compte tenu du fait qu'un modalisateur épistémique d'incertitude met en débat ou du moins qualifie comme non vérifiée l'information transmise, il doit être possible d'enchaîner par une phrase d'incertitude, par exemple *mais je ne suis pas sûr* ou *mais cela n'a pas été vérifié*. Si l'adjonction d'une telle tournure est appropriée, cela plaide en faveur d'une lecture épistémico-modale.

34. Le complément de groupe verbal **peut** être un groupe prépositionnel, un adverbe ou un complément phrastique. (frling02)
- 34' ?Le complément de groupe verbal peut être un groupe prépositionnel, un adverbe ou un complément phrastique, *mais cela n'a pas été vérifié*.
35. L'utilisation de données auto-déclarées **peut** avoir faussé les estimations relatives à la clientèle des programmes de dépistage, (frmed05)
- 35' L'utilisation de données auto-déclarées **peut** avoir faussé les estimations relatives à la clientèle des programmes de dépistage, *mais ce n'est pas sûr*.

Le troisième test est emprunté à Crompton (1997 : 282) et consiste à essayer de reformuler d'une manière plus certaine l'énoncé en remplaçant le marqueur en question. Si on arrive à une version plus catégorique ou moins atténuée de l'énoncé, le marqueur substitué est un atténuateur/modalisateur épistémique. Par exemple, dans l'exemple 32 ci-dessus, on peut remplacer *might* par *will* et par là obtenir un énoncé plus catégorique.

Ces tests ne sont pas infaillibles. L'enchaînement avec une tournure d'incertitude peut sembler inapproprié après un modalisateur épistémique qui est simplement stylistique ou conventionnel (cf. chapitre 7). En outre, le test de remplacement n'est pas toujours applicable : d'une part il n'existe pas toujours un marqueur intrinsèquement épistémico-modal

---

<sup>15</sup> Le point d'interrogation ne signifie pas que la phrase est agrammaticale, mais que la reformulation ne semble pas rendre fidèlement le contenu de la phrase originale.

synonyme ou quasi-synonyme auquel on peut substituer le marqueur en question, et d'autre part, en cas d'indétermination, la substitution par un modalisateur intrinsèquement épistémique peut être possible sans pour autant que la valeur dominante soit nécessairement la valeur épistémico-modale. Néanmoins, ensemble, ces tests ont constitué un outil utile dans le classement des occurrences.

Cela dit, il faut admettre que ces tests ne résolvent pas le problème avec la polyvalence du verbe anglais *indicate* et son correspondant français *indiquer*. Ces verbes signifient tantôt *suggest / suggérer*, tantôt *show / montrer*, et ce dans un cotexte syntaxique identique, à savoir *indicate / indiquer* + une subordonnée introduite par *that / que*. Ce n'est que dans le premier sens que *indicate* ou *indiquer* marquent une incertitude par rapport à l'information véhiculée. Il est souvent presque impossible de déterminer la valeur exacte de ces verbes :

36. Again, the results **indicate** that oral communication skills are important for all of these aspects. (engling13)

37. Des études effectuées au Canada (13), en Grande-Bretagne (14) et aux États-Unis (15,16) **indiquent** toutes que le prix est un déterminant important du tabagisme. (frmed04)

La distinction n'est pas toujours pertinente, et les deux valeurs ne sont pas nécessairement mutuellement exclusives. Selon Johns (2001 : 61), le fait de choisir *indicate* permet à l'auteur de s'exprimer délibérément d'une façon vague dans les cas où il n'a pas envie de préciser son niveau de certitude. En raison de cette indétermination, les occurrences de *indicate* et de *indiquer* sont particulièrement difficiles à classer. On doit s'appuyer sur le cotexte pour décider de la valeur. Ainsi, dans les cas où ces verbes apparaissent en combinaison avec un autre marqueur d'incertitude, ou lorsque le cotexte suggère d'une manière ou l'autre qu'il y a une incertitude attachée à l'information transmise, ils ont été étiquetés comme épistémiques. De fait, aucun des marqueurs dans les exemples ci-dessus n'a été considéré comme marqueur d'incertitude, mais une occurrence comme celle que l'on observe dans l'exemple 38 l'a été.

38. Cela pourrait également **indiquer** que les personnes ont déjà utilisé de la théophylline, à laquelle ont été ajoutés des corticostéroïdes, ou pour lesquels ils ont été substitués au cours de l'année. (frmed03)

Cette solution est loin d'être optimale, car elle implique que *indicate* et *indiquer* ne sont pas en eux-mêmes vus comme des modalisateurs épistémiques à fonction atténuante, mais qu'ils peuvent assumer cette fonction dans certains cotextes, sous l'influence, pour ainsi dire,

d'autres marqueurs ou indices dans le cotexte. Il est cependant difficile d'envisager une meilleure manière de les classer.

Bien entendu, ce verbe a aussi d'autres acceptions. Dans les cas où *indiquer/indicate* est utilisé dans les sens de *signifier* ou de *dire*, les attestations ont été aisément identifiées comme non-épistémiques :

39. Les six propriétés qui opposent *en* quantitatif et *en* adnominal peuvent être résumées dans le tableau ci-dessous. Le 0 **indique** que la propriété n'est pas pertinente. (frling06)

40. Comme nous l'avons **indiqué** au départ, l'identité sémantique de *sur*, telle qu'elle est donnée par sa FS, est neutre pour ce qui est de la distinction entre valeurs spatiales, temporelles et figurées. (frling19)

Outre les tests mentionnés ci-dessus, qui sont d'ordre sémantique et qui s'appliquent à tous les marqueurs polysémiques, une série de critères syntaxiques se sont avérés utiles. Or, ceux-ci sont, évidemment, spécifiques aux différents marqueurs dans les différentes langues et nous ne pouvons pas les présenter tous ici. Nous revenons à ces critères dans les sections respectives du chapitre 7 pour les marqueurs français et norvégiens qui ont été sélectionnés pour des études plus approfondies. Ici seront donnés quelques exemples de critères syntaxiques qui s'appliquent à d'autres marqueurs.

Par exemple, en ce qui concerne le verbe anglais *assume*, qui peut signifier *supposer/croire*, mais aussi *adopter/prendre comme point de départ*, les différentes constructions syntaxiques vont de pair avec différentes lectures<sup>16</sup>. Ainsi, les constructions syntaxiques *assume that* et *be assumed to* impliquent une lecture épistémique, c'est-à-dire la signification *supposer/croire* (voir l'exemple 41), alors que *assume* + NOM a la propriété inverse (voir l'exemple 42) :

41. We therefore **assume** that these default preferences for the null complements in (19a), (20a) and (21a) have been established as part of the conventional meanings of the relevant verbs. (engling09)

42. I **assume** Hornstein's (1990) Reichenbachian approach to tense (engling 01)

La construction *assuming that* peut aller de pair avec une lecture épistémique, notamment dans les cas où la désinence *-ing* est simplement déclenchée par la présence d'une préposition

---

<sup>16</sup> *Assume* a aussi d'autres acceptions comme dans *assume responsibility* ('assumer la responsabilité'), mais ces acceptions se distinguent plus facilement de la valeur épistémique.

(voir l'exemple 43). Cependant, dans le corpus, cette construction est le plus souvent utilisée dans le sens de *pourvu que* ou *si l'on prend comme point de départ* (voir l'exemple 44) :

43. Nevertheless, there are good reasons for **assuming that** bare XP resultatives are not causatives, consistent with our proposal that they have a simple event structure. (engling18)
44. **Assuming that** the reciprocal *each other* must be c-commanded at LF by its antecedent, this contrast shows that whereas the subject of (36a) c-commands the reciprocal, the associate in (36b) does not c-command the reciprocal at LF. (engling01)

Quant au verbe *appear*, la construction *appear* + PREP (par exemple *in*, *with*) est liée à un sens non-épistémique (voir l'exemple 45), alors que les constructions *appear to be* ou *it appears that* sont liées à des valeurs épistémiques (exemple 46) :

45. The unergative use of these verbs is taken as basic, yet other unergative verbs rarely **appear** in both patterns, as Wechsler (1997:312-13) points out. (engling18)
46. This generalization **appears** to be exceptionless. (engling18)

L'adjectif *possible* tend pour sa part à recevoir une interprétation épistémique quand il apparaît dans la construction *it is possible that* ou, en français, *il est possible que*, alors qu'il reçoit une lecture non-épistémique dans la construction *it is possible to* ou *il est possible de*:

47. Ci-dessus nous avons posé que dans le cadre de la combinatoire Verbe / Préposition **il est possible de** distinguer différents degrés d'intrication entre le V (plus exactement sa représentation) et la Prép en tant que schéma de la forme XRY. (frling19)
48. **Il est possible que** d'autres hypothèses puissent se formuler, complétant ou contredisant notre présentation. (frmed19)

#### 6.4.2 Résultats<sup>17</sup>

Maintenant que les principes méthodologiques ont été présentés, il est temps de se reporter aux résultats. Le tableau 3 résume les résultats pour les articles français, le tableau 4 les résultats pour les textes norvégiens, et le tableau 5 ceux des articles anglais. Nous allons d'abord commenter l'aspect quantitatif, en comparant les fréquences relatives des divers sous-corpus (6.4.2.1) ; nous poursuivons ensuite avec des observations concernant la répartition de

---

<sup>17</sup> Les chiffres présentés ici diffèrent quelque peu de ceux présentés dans Vold 2006a. Ces différences minimes s'expliquent par un travail d'analyse minutieux qui a été effectué pour corriger et affiner les résultats du premier codage des marqueurs.

différents types de marqueurs (6.4.2.2). Tout d'abord s'imposent quelques brèves remarques explicatives sur les tableaux.

#### *Lire les tableaux*

Après une première colonne « discipline », la deuxième répertorie les différents marqueurs par ordre décroissant. La troisième colonne, intitulée *occurrences épistémico-modales*, se divise en deux : la première partie indique le taux d'occurrences épistémico-modales pour mille (f/1000) et la seconde partie donne le nombre absolu d'occurrences épistémico-modales du marqueur en question (N). La quatrième colonne, titrée *toutes occurrences*, indique les taux et les nombres absolus correspondants des occurrences du marqueur prises dans leur ensemble. En comparant les deux colonnes, on peut donc déterminer dans quelle mesure un marqueur peut être considéré comme principalement épistémico-modal. Par exemple, on voit qu'il y a en tout 64 attestations de l'adjectif français *possible* dans les articles de linguistique, dont uniquement deux sont étiquetées comme épistémico-modales. Cela indique que l'emploi épistémique de cet adjectif est très rare en linguistique. Cet emploi est un peu plus élevé dans les articles en médecine, où 7 sur 30 occurrences ont été classées comme épistémiques (cf. le tableau 3).

Le fait que les chiffres dans les deux colonnes ne sont pas toujours identiques pour les marqueurs considérés comme intrinsèquement épistémico-modaux (par exemple *sembler*, *probably*, cf. 6.4.1.3) est principalement dû à leurs usages métalinguistiques.

**Tableau 3 :** Marqueurs épistémico-modaux dans le corpus français

Discipline	Marqueur	Occ. épistémico-modales		Toutes occurrences	
		f/1000	N	f/1000	N
Linguistique	sembler	0.87	60	0.90	62
	paraître	0.33	23	0.39	27
	sans doute	0.26	18	0.26	18
	peut-être	0.16	12	0.16	12
	peut	0.06	4	5.24	360
	suggérer	0.06	4	0.09	6
	pourrait	0.06	4	0.76	52
	indiquer	0.03	2	0.35	24
	possible	0.03	2	0.93	64
	probablement	0.01	1	0.01	1
	probable	-	-	-	-
	<b>total</b>	<b>1.9</b>	<b>130</b>	<b>9.1</b>	<b>626</b>
	Médecine	sembler	0.75	46	0.75
pourrait		0.34	21	0.52	32
probablement		0.28	17	0.28	17
suggérer		0.28	17	0.28	17
paraître		0.13	8	0.13	8
probable		0.13	8	0.13	8
peut		0.13	8	2.58	158
possible		0.11	7	0.49	30
indiquer		0.08	5	0.51	31
peut-être		0.05	3	0.05	3
sans doute		0.03	2	0.03	2
<b>total</b>		<b>2.3</b>	<b>142</b>	<b>5.8</b>	<b>352</b>

**Tableau 4 :** Marqueurs épistémico-modaux dans le corpus norvégien

Discipline	Marqueur	Occ. épistémico-modales		Toutes occurrences		
		f/1000	N	f/1000	N	
Linguistique	kan	'peut'	0.68	62	8.25	747
	se ut	'paraître'	0.53	48	0.63	57
	anta	'supposer'	0.44	40	0.54	49
	synes	'sembler'	0.36	33	0.40	37
	kanskje	'peut-être'	0.33	30	0.33	30
	mulig	'possible'	0.22	20	1.00	91
	trolig	'vraisemblablement'	0.17	15	0.17	15
	tyde på	'suggérer'/'indiquer'	0.17	15	0.18	16
	antagelse	'supposition'	0.14	13	0.20	18
	sannsynligvis	'probablement'	0.09	8	0.09	8
	muligens	'peut-être'	0.06	5	0.06	5
	<b>total</b>		<b>3.2</b>	<b>289</b>	<b>11.85</b>	<b>1073</b>
	Médecine	kan	'peut'	1.04	45	5.27
synes		'sembler'	0.46	20	0.56	24
tyde på		'suggérer'/'indiquer'	0.37	16	0.37	16
sannsynligvis		'probablement'	0.30	13	0.30	13
mulig		'possible'	0.28	12	0.74	32
anta		'supposer'	0.21	9	0.25	11
muligens		'peut-être'	0.19	8	0.19	8
trolig		'vraisemblablement'	0.19	8	0.19	8
se ut		'paraître'	0.16	7	0.16	7
kanskje		'peut-être'	0.09	4	0.09	4
antagelse		'supposition'	0.02	1	0.02	1
<b>total</b>			<b>3.3</b>	<b>143</b>	<b>8.14</b>	<b>352</b>

**Tableau 5 : Marqueurs épistémico-modaux dans le corpus anglais**

Discipline	Marqueur	Occ. épistémico-modales		Toutes occurrences	
		f/1000	N	f/1000	N
Linguistique	seem	0.80	137	0.80	137
	suggest	0.71	121	1.03	176
	assume	0.54	92	0.81	138
	may	0.50	86	1.64	280
	appear	0.43	74	0.64	109
	might	0.34	58	0.67	114
	perhaps	0.22	37	0.22	37
	indicate	0.12	20	0.45	77
	probably	0.08	13	0.10	17
	possible	0.06	11	0.87	148
	could	0.06	10	0.53	91
	<b>total</b>	<b>3.9</b>	<b>659</b>	<b>7.7</b>	<b>1324</b>
Médecine	may	1.28	76	1.75	104
	suggest	0.81	48	0.93	55
	might	0.37	22	0.44	26
	could	0.29	17	0.72	43
	possible	0.20	12	0.39	23
	indicate	0.19	11	0.42	25
	appear	0.10	6	0.10	6
	assume	0.07	4	0.08	5
	perhaps	0.07	4	0.07	4
	probably	0.05	3	0.05	3
	seem	0.02	1	0.02	1
<b>total</b>	<b>3.4</b>	<b>204</b>	<b>4.97</b>	<b>295</b>	

#### 6.4.2.1 Fréquence

Revenons à présent à la question posée au début : quel facteur, discipline ou langue, semble influencer le plus sur l'emploi des modalisateurs épistémiques ? Considérons d'abord la variable disciplinaire : le test bilatéral de Mann-Whitney a été effectué sur les différentes séries de données pour déterminer la significativité statistique des différences observées ( $\alpha = 0.05$ ). Dans le corpus français, les marqueurs étudiés constituent 1.9 pour mille des mots dans les articles de linguistique et 2.3 pour mille dans ceux de médecine. Il semble a priori y avoir une légère différence entre ces deux disciplines en ce qui concerne la fréquence de modalisateurs épistémiques, mais cette différence ne s'avère pas statistiquement significative ( $p = 0.425$ ).

Dans le corpus norvégien, il n'y a pas de différences sensibles entre les deux disciplines en ce qui concerne la fréquence des marqueurs retenus. La fréquence relative est de 3.2 en linguistique et de 3.3 en médecine. En toute logique, cette différence n'est pas statistiquement significative ( $p = 0.914$ ).



En ce qui concerne les textes anglais, on note une différence entre disciplines, qui va dans la direction inverse que celle observée dans le corpus français : la fréquence relative des marqueurs étudiés est de 3.9 dans les articles linguistiques et de 3.4 dans les articles en médecine. Les linguistes et les chercheurs en médecine ne diffèrent pourtant pas de manière statistiquement significative dans l'usage qu'ils font des marqueurs ( $p = 0.204$ ).

Le tableau 6 indique les différences entre disciplines quand les trois langues sont considérées en bloc. La fréquence relative des marqueurs étudiés est de 3.3 dans le corpus linguistique et de 3.0 dans le corpus médical, suggérant que les linguistes utilisent ces marqueurs un peu plus que les médecins. Cependant, comme nous l'avons vu, la différence est loin d'être statistiquement significative ( $p = 0.830$ ).

**Tableau 6** : Différences entre disciplines

Discipline	f/1000	Nombre d'occ.	Nombre de mots
Linguistique	3.3	1078	330287
Médecine	3.0	489	163824

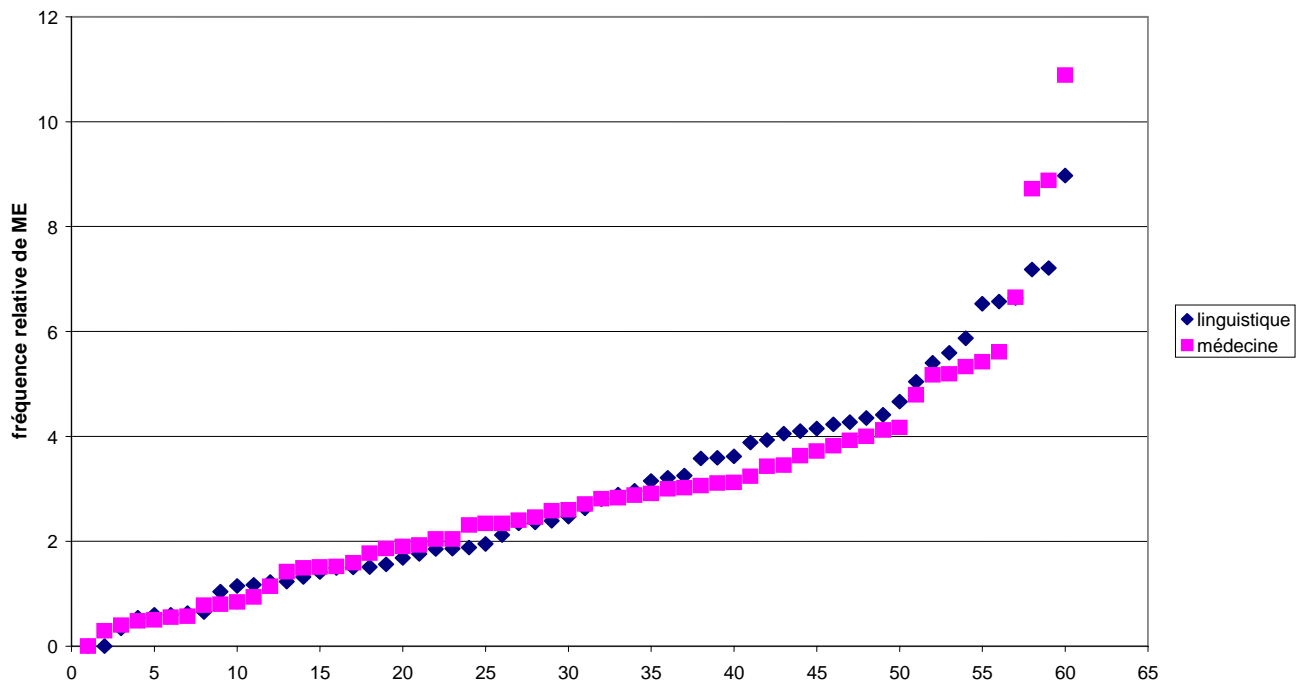
Ces résultats nous amènent donc à conclure qu'au niveau de la fréquence, il n'y a pas de différences nettes entre les deux disciplines investiguées. Il est vrai que quand les trois langues sont considérées en bloc, la fréquence des modalisateurs épistémiques examinés est légèrement plus élevée dans le corpus linguistique que dans le corpus médical, mais cette différence résulte entièrement de la partie anglaise du corpus. Dans les sous-corpus français et norvégien, on ne la retrouve pas. L'hypothèse posée au départ, selon laquelle les articles en médecine utiliseraient moins de modalisateurs épistémiques que ceux des linguistes, étant donné que ces premiers représentent une science plus « dure » que les seconds, se trouve donc difficile à valider en tant que telle, car elle n'est ni étayée ni infirmée par ces résultats.

Les différences (ou plutôt les ressemblances) entre les deux disciplines sont illustrées dans le graphique 1. Ce graphique montre le taux individuel de modalisateurs épistémiques de chacun des articles, donnant une idée de la variation individuelle entre articles<sup>18</sup>. Les similarités entre les disciplines sont frappantes : les deux courbes se recouvrent presque complètement. Comme mentionné au début de ce chapitre, on pourrait s'attendre à ce que la variation individuelle soit moins importante parmi les médecins-chercheurs, étant donné que ceux-ci sont réputés écrire tous d'une manière assez similaire, alors que les linguistes sont

<sup>18</sup> Pour obtenir une idée de la variation individuelle, on pourrait aussi se reporter à l'appendice 2.

supposés avoir un style plus personnel. Cependant, le graphique montre que la variation entre articles est à peu près équivalente dans les deux disciplines.

**Graphique 1: Fréquence relative de ME dans les articles en linguistique et en médecine**



Alors que les différences entre disciplines sont faibles, les résultats présentés dans les tableaux 3, 4 et 5 suggèrent qu'il existe des différences considérables entre les trois langues : les marqueurs épistémico-modaux d'incertitude sont plus fréquents dans les corpus anglais et norvégien que dans le corpus français, et cette différence est particulièrement sensible dans la partie linguistique du corpus. Le tableau 7 ci-dessous indique les différences entre les langues quand les deux disciplines sont traitées en bloc. On notera que les chercheurs anglophones emploient le plus ces marqueurs, avec une fréquence relative de 3.7. Ils sont suivis des chercheurs norvégiens, chez qui la fréquence relative de ces marqueurs est de 3.2. Dans le corpus français, la fréquence relative est de 2.1. Les chercheurs francophones et les chercheurs anglophones constituent donc deux pôles, avec les chercheurs norvégiens se situant entre les deux. Il est toutefois clair que sur ce point, les norvégiens tendent plutôt vers le style anglo-américain que vers le style français.

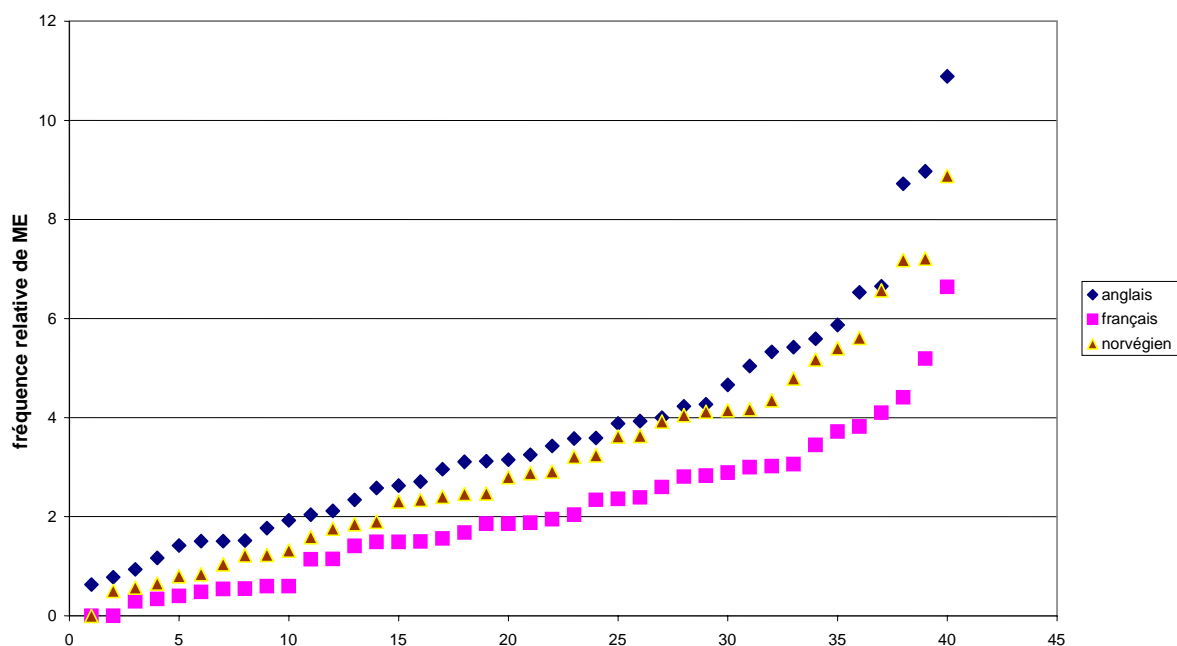
**Tableau 7 : Différences entre langues**

<b>Langue</b>	<b>f/1000</b>	<b>nombre d'occ.</b>	<b>nombre de mots</b>
Français	2.1	272	129907
Norvégien	3.2	432	133813
Anglais	3.7	863	230391

Pour déterminer la signification statistique de ces différences, le test de Kruskal-Wallis a été mis en œuvre sur les trois langues en bloc, et le test bilatéral de Mann-Whitney ensuite sur les paires de langues. De manière très nette, les différences entre les langues se sont avérées statistiquement significatives : les auteurs anglophones diffèrent significativement des auteurs francophones ( $p = 0.000$ ), les auteurs norvégiens diffèrent significativement aussi des auteurs francophones ( $p = 0.019$ ), les auteurs anglophones et les auteurs norvégiens, cependant, ne diffèrent pas de manière statistiquement significative ( $p = 0.290$ ). On peut en tirer la conclusion que les marqueurs épistémico-modaux à fonction atténuante sont nettement plus fréquents dans les articles scientifiques écrits en anglais et en norvégien que dans ceux rédigés en français. Cela corrobore les deux hypothèses posées au départ : on trouve un usage plus fréquent de modalisateurs épistémiques dans les articles anglais que dans les articles français, et le style norvégien est à cet égard similaire au style anglo-américain.

Le graphique 2 illustre les différences entre les trois langues, ainsi que la variation individuelle entre articles. Une certaine variation individuelle est prévisible, étant donné que la fréquence de modalisateurs épistémiques est censée varier selon le sujet traité, l'âge et le statut professionnel de l'auteur, etc. Mais cette variation normale et escomptée ne devrait pas différer d'une langue à l'autre. Le graphique montre pourtant que les auteurs francophones constituent un groupe plus homogène à cet égard que les auteurs anglophones et norvégiens : on observe que dans les articles rédigés en français, la fréquence relative de modalisateurs épistémiques tend à se situer entre zéro et quatre pour mille, alors que dans les articles rédigés en norvégien et en anglais elle tend à se situer entre zéro et six pour mille, indiquant une variation plus grande dans ces groupes. En outre, il y a un bon nombre d'articles norvégiens et anglais qui font preuve d'une fréquence relative encore plus élevée (cf. le graphique).

**Graphique 2: Fréquence relative de ME dans les articles anglais, français et norvégiens**



Avant d’interpréter et discuter plus amplement ces résultats, nous allons examiner l’influence éventuelle du troisième facteur qu’on s’est proposé d’étudier : celui du sexe de l’auteur. Rappelons que le classement d’articles selon le sexe de l’auteur était possible uniquement pour le corpus de linguistique, puisque les articles médicaux sont en général rédigés collectivement, et associent le plus souvent des auteurs féminins et masculins. Dans le même sens, six articles du corpus anglais et un du corpus français ont été exclus à cause d’une auctorialité mixte. Les 53 articles qui restent ont été divisés en un groupe « masculin » et un groupe « féminin », dont les taux de modalisateurs épistémiques ont été comparés à l’aide du test bilatéral de Mann-Whitney. Les différences repérées sont indiquées dans le tableau 8 ci-dessous ; ce tableau présente les résultats pour chacune des langues individuellement ainsi que pour les langues considérées en bloc.

**Tableau 8 :** Différences selon le sexe de l’auteur (articles de linguistique uniquement)

langage	f/1000		No. d’occurrences		No. de mots		No. d’articles	
	Masc.	Fém.	Masc.	Fém.	Masc.	Fém.	Masc	Fém.
Français	1,9	1,9	82	41	43 657	21 350	11	8
Norv.	3,0	3,6	182	107	61 089	29490	12	8
Anglais	4,1	3,9	191	293	46 495	74 437	6	8
Toutes	3,0	3,5	455	441	151 241	125 277	29	24

Comme le tableau l'indique, les différences de fréquence de modalisateurs épistémiques sont faibles entre auteurs masculins et auteurs féminins, et elles ne vont même pas dans la même direction dans les différents sous-corpus. Il semble y avoir une différence sensible entre les sexes dans le corpus norvégien, où le taux de modalisateurs épistémiques est plus élevé chez les femmes, mais cette différence apparente est loin d'être statistiquement significative ( $p = 0,970$ ). L'éventuel impact du sexe de l'auteur sur le taux de modalisateurs épistémiques n'est donc statistiquement significatif dans aucune des langues ( $p = 0,492$  et  $0,491$  pour le français et l'anglais, respectivement), ni pour les langues considérées en bloc ( $p = 0,592$ ). On peut en conclure que la variable du sexe n'est pas pertinente dans l'usage de ces marqueurs.

### *Interprétation et discussion*

A partir des résultats présentés ci-dessus, on peut tirer la conclusion que le facteur le plus décisif en ce qui concerne la fréquence de modalisateurs épistémiques est l'appartenance langagière du chercheur et non pas l'appartenance disciplinaire. Cette observation nuance les résultats principaux du projet KIAP, selon lesquels le facteur de la discipline prime sur le facteur de la langue pour la plupart des phénomènes étudiés, c'est-à-dire que la plus grande partie de la variation observée est imputable au facteur disciplinaire. Le projet KIAP conclut donc que les pratiques d'écriture scientifique dépendent davantage des cultures disciplinaires que des cultures linguistiques (Fløttum et al. 2006). En ce qui concerne la modalité épistémique, la situation est plutôt inverse : nos résultats suggèrent que pour ce phénomène, les pratiques d'écriture varient plus entre langues qu'entre disciplines : nous venons de voir que, en ce qui concerne l'aspect quantitatif de l'emploi de modalisateurs épistémiques, les différences entre disciplines sont assez faibles en comparaison des différences entre langues. L'appartenance disciplinaire semble cependant influencer sur l'importance des différences langagières et vice versa ; on observe donc une interdépendance de ces deux facteurs. Par exemple, en ce qui concerne les différences entre langues, nous avons vu que celles-ci sont plus grandes parmi les linguistes que parmi les médecins-chercheurs (cf. les tableaux 3, 4 et 5), ce qui peut être compris comme un effet du plus grand degré de standardisation qui caractérise le discours médical (cf. *supra*, chapitre 2). Il semble raisonnable de conclure que les différences entre langues sont moins importantes dans le corpus médical puisque les normes qui gouvernent le discours écrit médical dépassent souvent les frontières linguistiques, la médecine étant une discipline hautement internationalisée. Bien que la variation individuelle entre articles se soit avérée aussi importante dans le corpus médical que dans le corpus linguistique (cf. graphique 1), il n'en reste pas moins que les trois sous-corpus

médicaux sont plus similaires entre eux (fréquences relatives de 2.3, de 3.3 et de 3.4, cf. les tableaux 3-5) que les trois sous-corpus linguistiques (fréquences relatives de 1.9, de 3.2 et de 3.9, cf. les tableaux 3-5). De ce point de vue, on pourrait conclure que les médecins-chercheurs constituent un groupe plus homogène que les linguistes.

En ce qui concerne les différences entre disciplines, il convient de noter que si l'hypothèse de départ, selon laquelle les chercheurs en médecine utiliseraient moins de modalisateurs épistémiques que les linguistes, tient toujours pour le corpus anglais (mais non de façon significative), elle est en fait contredite par les résultats pour le corpus français, où ce sont les chercheurs en médecine qui utilisent le plus de modalisateurs épistémiques (voir les tableaux 3-5). Une explication raisonnable de cette observation est que les chercheurs en médecine sont influencés dans une plus grande mesure que les linguistes par la pratique d'écriture anglo-américaine. Les linguistes tendent à s'en tenir à un taux bas de modalisateurs épistémiques, propre au style français, alors que les médecins-chercheurs s'approchent du taux plus élevé qui caractérise le style anglo-américain. On pourrait soupçonner que ce rapprochement entre médecins-chercheurs francophones et auteurs anglophones est dû aux quelques articles canadiens du corpus français (Frmed01-05) (cf. *supra* 2.2.2), mais ceux-ci n'ont pas de fréquences relatives particulièrement élevées (cf. appendice 2).

On observe aussi que les différences entre disciplines sont les plus saillantes dans le corpus anglais (cf. tableaux 3-5). Cela peut indiquer que c'est dans la recherche publiée dans les revues internationales que se manifestent le plus clairement les lignes de démarcation entre les différents domaines scientifiques. Il est bien connu que les articles médicaux écrits pour la communauté scientifique internationale doivent obéir à des normes de rédaction bien précises, mais les règles de la rédaction médicale semblent être plus mouvantes quand il s'agit d'articles écrits pour une communauté nationale. Cela se manifeste dans l'utilisation du format IMRAD – tous les articles médicaux écrits en anglais suivent rigoureusement ce format, alors que dans les corpus français et norvégien certains articles le suivent moins rigoureusement (cf. *infra*, les études de cas). Le fait que les articles médicaux anglais suivent une structure à ce point standardisée et adoptent un style scientifique traditionnel contribue sans doute aux plus grandes différences observées dans le corpus anglais.

L'observation de différences faibles entre les disciplines étaye la conclusion de Markkanen et Schröder, affirmant que « ... the differences in the use of hedges between texts in different fields are not so great as has been often assumed » (Markkanen and Schröder 1997 : 10). Cependant, il faut souligner qu'il est tout à fait possible de trouver des différences plus grandes entre d'autres disciplines – ici n'en ont été comparées que deux. Par exemple,

Varttala (2001 : 154) a observé que la fréquence d'atténuateurs est plus élevée dans la discipline de l'économie que dans celles de la médecine et de la technologie – ces dernières ayant approximativement la même fréquence d'atténuateurs. Et comme nous l'avons vu au chapitre 4, des différences quantitatives entre disciplines ont été constatées dans d'autres études aussi (Hyland 1999b, Ventola 1997). En outre, une discipline ne constitue évidemment pas une culture bien homogène. Chaque discipline scientifique consiste en plusieurs sous-disciplines susceptibles de représenter des traditions théoriques et méthodologiques différentes, voire contradictoires. Afin de pouvoir répondre de manière plus approfondie à la question de savoir dans quelle mesure l'appartenance disciplinaire influence la fréquence d'atténuateurs, d'autres disciplines et même des sous-disciplines devraient être comparées dans des études ultérieures. Il serait donc hâtif de conclure, à partir d'une comparaison entre deux disciplines uniquement, que le facteur de discipline n'est pas important. Cependant, pour les disciplines et les langues étudiées ici, il est clair que c'est le facteur de langue qui l'emporte. Les résultats suggèrent qu'en ce qui concerne la fréquence d'atténuateurs, les sciences naturelles et les sciences humaines, représentées ici par la médecine et la linguistique, ne diffèrent pas toujours autant qu'on peut le croire. Il s'avère que cela dépend des disciplines (et de la langue) investiguées, et cela incite à beaucoup de nuance.

En ce qui concerne les différences entre les trois langues, celles-ci se manifestent de manière statistiquement significative dans tous les groupes : que l'on compare les linguistes entre eux ou les médecins-chercheurs entre eux, les chercheurs anglophones et norvégiens utilisent significativement plus de modalisateurs épistémiques que les chercheurs francophones. La fréquence de ces marqueurs est la plus élevée dans les articles anglais, mais les articles norvégiens sont très proches. La ressemblance entre les articles anglais et norvégiens est en accord avec l'hypothèse posée au début, selon laquelle le discours scientifique anglais et norvégien d'un côté et le discours scientifique français de l'autre constitueraient deux pôles opposés qui représenteraient deux styles différents. Une série de résultats issus d'autres études du projet KIAP corroborent cette observation : les articles écrits en anglais ou en norvégien attestent d'une fréquence plus élevée que les articles écrits en français de phénomènes linguistiques tels que les pronoms personnels de la première personne (Fløttum 2003b ; Fløttum et al. 2006), le métatexte (Dahl 2003, 2004, Fløttum et al. 2006), la négation et les marqueurs concessifs (Fløttum 2005a, 2005c ; Fløttum et al. 2006), et les références bibliographiques (Fløttum 2003c). La fréquence importante de tels phénomènes linguistiques suggère que les auteurs anglophones et norvégiens se manifestent plus explicitement dans leurs textes que les auteurs francophones, et qu'ils sont plus explicitement

polémiques, dans le sens où ils utilisent des moyens linguistiques pour marquer qu'ils souhaitent entrer en discussion avec leurs pairs (Fløttum et al. 2006).

Soulignons une fois de plus que si les modalisateurs épistémiques sont plus fréquents dans les corpus anglais et norvégien, il ne faut pas automatiquement en tirer la conclusion que les chercheurs anglophones et norvégiens atténuent plus leurs énoncés que les chercheurs francophones. La modalité épistémique n'est qu'une des stratégies d'atténuation ; il en existe d'autres, comme par exemple les compléments adverbiaux d'opinion et certains adverbes de fréquence (cf. *supra*, section 6.2.1). Cependant, comme la modalité épistémique est reconnue comme une stratégie dominante et essentielle de l'atténuation (voir Hyland 1998 : 149 ; Salager-Meyer 1994 ; Vihla 2000 ; Varttala 1999), il est raisonnable d'avancer que les articles anglais et norvégiens sont plus fortement atténués que ceux rédigés en français. Si l'atténuation est généralement considérée comme un trait caractéristique du discours scientifique (voir par exemple Lewin 2005 : 164), les résultats présentés ici suggèrent qu'il ne s'agit pas nécessairement d'un trait caractéristique du discours scientifique en général, mais peut-être plutôt d'un phénomène qui est lié au discours scientifique de certains groupes de langues.

Quelles peuvent être les explications de ces différences entre les articles anglais et norvégiens d'un côté et les articles français de l'autre ? La réponse se trouve peut-être dans les différentes cultures et traditions scientifiques. La culture scientifique norvégienne est fortement influencée par la culture anglo-américaine et des différences culturelles entre chercheurs anglophones et chercheurs francophones ont été décrites dans des recherches antérieures. A partir de leur étude contrastive des figures de rhétorique employées dans des articles de médecine écrits en anglais, français et espagnol, Salager-Meyer et al. (2003) concluent que les chercheurs français et espagnols tendent à s'exprimer d'une manière plus critique et plus autoritaire que leurs collègues anglophones. Ces auteurs remarquent que leurs observations sont en accord avec celles de Motchane (1990) et celles de Régent (1994), suggérant que « French scientists are much more prescriptive, authoritative and categorical than their English-speaking colleagues » (Salager-Meyer et al. 2003 : 232). De telles différences culturelles peuvent expliquer pourquoi Beaufrère-Bertheux (1997 : 232) fait référence à ce qu'elle appelle l'*hypermodestie* des chercheurs anglais et pourquoi Sionis (1997 : 211) attribue le refus par une revue internationale d'un article écrit en anglais par un groupe de chercheurs français à ce qu'il appelle la *confiance exagérée* des auteurs. Dans la mesure où de telles différences culturelles existent, il est raisonnable de supposer qu'elles se reflètent dans les choix linguistiques que font les auteurs.



Une autre explication possible est que les chercheurs anglophones et norvégiens utilisent des atténuateurs pour neutraliser leur emploi relativement fréquent de verbes tels que *argue* et *claim* en combinaison avec des pronoms personnels de la première personne, c'est-à-dire des constructions qui signalent la position personnelle de l'auteur. Certaines études de KIAP ont montré que les auteurs anglophones et norvégiens utilisent ce type de verbes plus que les auteurs francophones (Fløttum 2004 ; Fløttum et al. 2006), et les atténuateurs pourraient être un moyen de contrebalancer l'impression d'un auteur trop affirmatif ou trop polémique.

En ce qui concerne la variable du sexe de l'auteur, celle-ci ne joue pas de rôle significatif sur la quantité de modalisateurs épistémiques utilisés dans les articles examinés. A première vue, les chiffres semblent indiquer que dans l'ensemble, les auteurs féminins utilisent un peu plus de modalisateurs que les auteurs masculins, mais cette différence résulte uniquement du corpus norvégien. En outre, au sein de chacun des deux groupes comparés, la variation individuelle est grande : dans le groupe féminin, la fréquence relative des marqueurs sélectionnés va de 0,3 à 9,0, et dans le groupe masculin, elle va de 0 à 7,2. Cette variation intragroupe est particulièrement importante dans le corpus norvégien. Les différences apparentes entre auteurs féminins et auteurs masculins dans le corpus norvégien pourraient ainsi relever de quelques articles atypiques. Les résultats de notre étude ne peuvent donc pas étayer la conception traditionnelle évoquée entre autres par R. Lakoff (1975) et Preisler (1986), selon laquelle l'atténuation serait un trait caractéristique du discours féminin. Au contraire, les résultats indiquent que le sexe de l'auteur n'a aucun effet sur la fréquence de modalisateurs épistémiques, mais le nombre limité d'articles étudiés empêche de tirer des conclusions définitives à ce sujet.

#### 6.4.2.2 Types de marqueurs

L'analyse exploratoire a révélé que les trois langues font dans une large mesure usage des mêmes types formels de marqueurs. A part les auxiliaires de possibilité, que nous avons choisi de mettre dans la catégorie des moyens grammaticaux, il s'agit presque exclusivement de moyens lexicaux. Certains des moyens morphologiques ont été attestés dans le corpus français et certaines occurrences de subordinées conditionnelles à valeur épistémico-modale ont été attestées dans les corpus norvégien et anglais, mais de manière rare. Il ressort des tableaux 3-5 que ce sont les verbes de perception, les auxiliaires modaux, les adverbiaux de phrase et les verbes épistémiques qui sont les types de marqueurs les plus récurrents. Les adjectifs et les noms épistémiques sont moins fréquents. Cette observation vaut pour les trois

langues. En ce qui concerne les types formels de marqueurs, il ne semble donc pas y avoir des différences remarquables entre ces langues. On pourrait toutefois noter que l'anglais dispose d'une plus large série d'auxiliaires de possibilité que le norvégien et le français (cf. *supra*, section 3.3), de sorte que les auxiliaires modaux représentent un moyen d'expression de la modalité épistémique particulièrement important en anglais. Pour ce qui est des types sémantiques, on peut noter que les auteurs anglophones et norvégiens font usage des verbes de croyance *assume* et *anta*, qui privilégient l'emploi de pronoms personnels sujets, alors qu'il n'y a pas de correspondant français sur la liste.

Si les différences entre langues sont faibles, qu'en est-il des différences entre disciplines en ce qui concerne les types de marqueurs préférés ? Rappelons que les résultats de l'analyse exploratoire ont suggéré qu'un bon nombre des marqueurs français et anglais étaient spécifiques à l'une des deux disciplines; nous allons voir que l'analyse quantitative confirme en grande partie cette tendance. Dans notre tentative d'expliquer les différences observées, nous mettrons l'accent sur les traits sémantiques des marqueurs plutôt que sur leurs caractéristiques formelles ; ceux-ci sont, par ailleurs, souvent liés. Résumons d'abord les différences et les ressemblances observées entre les disciplines, et ce pour chaque langue :

Dans le corpus français, le semi-auxiliaire *sembler* est le modalisateur épistémique le plus fréquent, et de loin, tant dans les articles de linguistique que dans les articles de médecine. Voici quelques exemples :

49. Le budénoside, un corticoïde à action topique plus élevée et un passage systémique plus faible, **semble** être efficace avec moins d'effets secondaires [21],[22]. (frmed06)

50. Il **semble** donc que les théories du paradigme B se répartissent en deux groupes (frling11)

Mis à part *sembler*, les linguistes et les médecins-chercheurs ne font pas usage des mêmes marqueurs. Comme il ressort du tableau 3, les linguistes préfèrent *paraître* et *sans doute*, alors que les médecins-chercheurs préfèrent *pourrait* et *probablement*. Quant à la fréquence relativement élevée de *suggérer* dans les articles de médecine, celle-ci est largement due à un seul article dans lequel on trouve 10 occurrences sur un total de 17 dans l'ensemble du corpus Frmed (cf. appendice 2), et on ne saurait donc dire si *suggérer* est en fait plus utilisé par les médecins-chercheurs que par les linguistes. Il est clair cependant qu'un grand nombre de marqueurs est relativement spécifique à l'une des deux disciplines. Les médecins-chercheurs emploient par exemple les expressions *probable* et *probablement*, tandis que ces expressions sont très rares chez les linguistes : il n'y a aucune occurrence de *probable* dans les articles de

linguistique et seulement une occurrence de *probablement*. *Sans doute*, en revanche, est une expression habituelle dans les articles de linguistique, mais que les médecins-chercheurs tendent à éviter. *Peut-être* s'avère aussi relativement fréquent dans les articles de linguistique, mais non dans ceux de médecine. En ce qui concerne *pouvoir*, l'analyse quantitative met en évidence que *peut* est très rare comme modalisateur épistémique, quelle que soit la discipline (ce qui confirme par ailleurs l'observation faite lors de l'analyse exploratoire), alors qu'il y a de grandes différences entre les deux disciplines en ce qui concerne l'emploi du conditionnel (*pourrait*). L'analyse révèle une quasi-absence de *pourrait* épistémique chez les linguistes (4 sur 52 occurrences ont une valeur épistémique), mais un usage massif chez les médecins-chercheurs (21 sur 32 attestations sont étiquetées comme épistémiques, ce qui donne un pourcentage de presque 66 %) :

51. Il reste qu'une certaine proportion de ces personnes sont traitées plutôt en fonction d'un stade 3, car 60 % d'entre elles ont également pris de la théophylline à action prolongée. Cela **pourrait** signifier qu'au lieu d'une augmentation du dosage du corticostéroïde en inhalation, le traitement ferait appel à des doses accrues de médicament au même dosage (frmed03)

L'emploi épistémique de *pourrait* est donc beaucoup plus courant dans les articles de médecine que dans les articles de linguistique, mais force est de constater que les variations individuelles sont là encore importantes. Bien que *pourrait* soit le deuxième marqueur épistémico-modal dans le corpus médical, un examen de la distribution d'occurrences entre articles révèle qu'il n'est utilisé que dans huit articles (voir l'appendice 2). Comme les articles français contiennent relativement peu de marqueurs épistémico-modaux, les résultats pour ce corpus sont susceptibles d'être influencés dans une plus large mesure par les variations individuelles entre articles ; c'est ce qui explique le poids de *pourrait* en dépit de sa distribution. Nous revenons par ailleurs aux différences disciplinaires observées par rapport à l'emploi de *pourrait* au chapitre 7, où sera présentée une étude plus approfondie de l'emploi de *pouvoir*.

La valeur épistémico-modale de *indiquer* et de *possible* s'avère peu utilisée, ce qui confirme les résultats préliminaires du corpus échantillon. Dans la plupart des cas, *indiquer* n'implique pas d'incertitude, mais plutôt la certitude ou l'affirmation. Ce verbe se distingue donc nettement de la locution verbale *tyde på* en norvégien, souvent présentée par les dictionnaires comme le correspondant d'*indiquer*. Il se distingue aussi en partie de son quasi-équivalent en anglais, *indicate*, qui semble plus facilement accepter le sens d'incertitude.

L'adjectif *possible* de son côté, est parfois utilisé comme modalisateur épistémique dans les articles de médecine, ce qu'illustre l'exemple ci-dessous :

52. **Il est possible que** des taux plus élevés d'anomalies congénitales soient simplement le reflet d'une propension plus grande à indiquer un code d'anomalies congénitales. (frmed01)

Ces cas sont pourtant assez rares, et le plus souvent les occurrences de *possible* expriment une affirmation simple d'une (ou plusieurs) possibilité(s) (possibilité radicale) :

53. Deux réalisations morphologiques sont également **possibles** (frling01)

En résumé, on constate qu'il y a des ressemblances aussi bien que des différences évidentes entre les linguistes et les médecins-chercheurs en ce qui concerne le choix d'expressions épistémico-modales. Il est vrai que les deux groupes préfèrent *sembler*, et la dominance de ce marqueur est très nette dans les deux disciplines. Cependant, certaines expressions différencient les deux groupes. *Pourrait*, *probable* et *probablement*, en tant que marqueurs épistémico-modaux, semblent être des marqueurs spécifiques à la discipline de médecine, alors que *sans doute*, et dans une certaine mesure *peut-être*, semblent être propres au discours des linguistes.

Dans le corpus norvégien, c'est la forme *kan* (le présent de l'auxiliaire modal *kunne*) qui est manifestement le marqueur épistémico-modal le plus fréquent dans les deux disciplines. Le verbe *synes* et le verbe à particule *tyde på* sont aussi des expressions importantes dans les deux disciplines. Une différence se révèle pourtant en ce qui concerne le verbe *anta* et le nom *antagelse*. Fréquentes chez les linguistes, ces expressions sont peu utilisées par les médecins-chercheurs. Une recherche électronique sur *anta*.\* dans le corpus Nomed, donne, symptomatiquement, beaucoup d'attestations de *antall* ('nombre'), mais peu d'attestations de *antagelse*, *antar*, etc. D'autres différences importantes concernent la locution verbale *se ut til/som* et l'adverbe *kanskje*, qui apparaissent souvent dans les articles de linguistique (fréquence relative de 0,53 et de 0,33, respectivement), mais qui sont assez rares dans les articles de médecine (fréquence relative de 0,16 et de 0,09, respectivement). Les médecins-chercheurs de leur côté, emploient l'adverbe *sannsynligvis* beaucoup plus que les linguistes (0,30 vs 0,09). Ces dernières observations concordent avec celles faites sur le corpus français, où *peut-être* et *probablement* se sont révélés spécifiques en linguistique et en médecine respectivement.

Quelques différences se manifestent donc en ce qui concerne les marqueurs choisis pour exprimer de la modalité épistémique, mais les différences sont moins importantes dans le corpus norvégien que dans le corpus français.

Les différences les plus nettes à cet égard se manifestent dans le corpus anglais. En comparant les résultats pour les corpus Engling et Engmed, on note qu'un grand nombre de marqueurs tendent à être spécifiques à l'une des disciplines : *seem*, *appear*, *assume* et *perhaps* sont tous utilisés presque exclusivement dans les articles de linguistique, alors que l'emploi de *could* et de *possible* se restreint plutôt aux articles de médecine.

Ces tendances devraient être testées sur un corpus plus large afin d'être éventuellement vérifiées, mais elles indiquent tout de même des différences intéressantes entre les disciplines par rapport aux choix de marqueurs. Tout marqueur épistémico-modal présuppose, d'un point de vue pragmatique, une instance évaluatrice, mais les propriétés sémantiques des marqueurs diffèrent sensiblement. Les verbes *seem*, *appear* et *assume* sont tous relativement « subjectifs » en ce sens qu'ils présupposent, de par leur sémantisme même, une instance modalisante spécifique, à savoir un sujet de conscience. *Assume* appartient à la catégorie traditionnellement appelée « *mental state predicates* » (Nuyts 2001) ou « *tentative cognition verbs* » (Varttala 2001 : 122), verbes qui possèdent une subjectivité inhérente en ce qu'ils dénotent un processus mental personnel, une activité cognitive subjective (Varttala 2001 : 122). Les semi-auxiliaires *seem* et *appear* impliquent aussi une instance modalisante spécifique, bien que la source de l'évaluation demeure, dans la plupart des cas, implicite. Cette source peut aisément être ajoutée : *it seems that* -> *it seems to me that*, ce qui vaut aussi pour l'emploi passif de *assume* : *x is assumed to be* -> *x is assumed by y to be...* Par la référence aux processus mentaux (*assume*) ou perceptifs (*seem* et *appear*), ces marqueurs présupposent une évaluation personnelle. *Could*, *may*, *might* et *possible*, par contre, renvoient à la notion de possibilité et peuvent être entendus comme des marqueurs d'une simple éventualité sans présupposer d'instance modalisante spécifique. Cette propriété se reflète dans le fait que tous ces marqueurs peuvent exprimer une possibilité radicale aussi bien qu'une possibilité épistémique, ce qui semble les draper d'un voile d'objectivité y compris dans leurs emplois épistémiques. *Seem*, *assume* et *appear* se présentent comme plus subjectifs, étant donné qu'ils n'ont pas l'aptitude d'exprimer la possibilité radicale et qu'ils présupposent toujours une instance modalisante. Ils portent la marque d'une évaluation personnelle, alors que *could*, *may*, *might* et *possible*, dans leurs emplois épistémiques, servent à dissimuler la source de l'évaluation. Evidemment, il y a une évaluation personnelle dans ces cas-là aussi – quelqu'un a choisi de mettre ces marqueurs d'incertitude dans le texte – mais l'éventualité

qu'ils expriment peut être interprétée comme un fait objectif plutôt que comme un jugement personnel sur la valeur de vérité de l'information véhiculée. Choisir ce dernier type de marqueurs peut être considéré comme une manière de favoriser l'impersonnalité par opposition à la subjectivité, en créant ainsi une impression d'objectivité (voir Hyland 2001). Le premier type de marqueurs est peut-être conçu comme trop subjectif ou trop personnel pour les chercheurs en médecine, qui de manière plus générale tendent à éviter les marqueurs de subjectivité dans leurs textes (Fløttum 2003b : 40, 2006 ; Fløttum et al. 2006).

Le verbe *suggest* est quant à lui susceptible de se placer dans les deux catégories : avec un sujet personnel (*I / we suggest*, voir l'exemple 54) il rejoint la catégorie de *assume*, mais avec un sujet passif ou non-animé (voir l'exemple 55), il apparaît comme un marqueur « objectif », moins dépendant d'un sujet modal. Alors que les médecins-chercheurs utilisent presque sans exception ce dernier type de constructions, les linguistes utilisent les deux (cf. Vold 2006b) :

54. How then did the control sense derive from the semantic network associated with *over*? **We suggest** that this sense is due to an implicature becoming conventionally associated with *over*, from an independently motivated experiential correlation between control and vertical elevation. (engling17)
55. We chose to examine variation by menopausal status because of **data suggesting** that breast cancer risk associated with carotenoid consumption <5> and ovarian cancer risk associated with caffeine consumption may be modified by menopausal status. <33> (engmed10)

Cette tendance qu'ont les linguistes et les médecins-chercheurs à choisir des types de marqueurs nettement différents semble valable uniquement pour le corpus anglais. Quelques signes de cette même tendance sont repérables dans les corpus français et norvégien, par exemple le fait que *pourrait* soit plus utilisé par les médecins-chercheurs que par les linguistes, et que *anta* ('assume') le soit plus par les linguistes que par les médecins-chercheurs. Cependant, la tendance est moins nette qu'en anglais, et il ne semble pas possible de relier les différences dans le choix de marqueurs à des préférences pour différents types sémantiques de marqueurs. Encore une fois, on observe que les différences entre disciplines se dessinent le plus clairement dans le corpus anglais.

Pour finir, nous pouvons comparer les résultats de la présente étude avec les chiffres obtenus par Vihla (2000). Rappelons que Vihla a examiné la fréquence d'une sélection de marqueurs de possibilité épistémique dans des articles de recherche médicaux (voir *supra* 4.2). Quatre des sept marqueurs étudiés par Vihla réapparaissent dans notre corpus : *may*,

*might*, *perhaps* et *possible*. Les ressemblances sont frappantes, comme il ressort du tableau ci-dessous :

**Tableau 9** : Fréquences relatives de *may*, *might*, *perhaps* et *possible*, dans le corpus de Vihla (2000) et dans notre corpus

	Dans l'étude de Vihla f/1000		Dans la présente étude f/1000	
	toutes occ.	occ.ép.	toutes occ.	occ.ép.
may	1.9	1.8	1.75	1.23
might	0.4	0.4	0.44	0.35
perhaps	0.1	0.1	0.1	0.1
possible	0.4	0.1	0.39	0.2

On pourrait expliquer les ressemblances frappantes par la standardisation de la rédaction médicale (et de fait, on trouvera à peu près les mêmes fréquences dans n'importe quel corpus d'articles scientifiques médicaux), mais elles indiquent aussi que notre conception de la possibilité épistémique et de ses frontières avec d'autres types de possibilité concorde assez bien avec celle de Vihla.

## 6.5 Bilan

Dans ce chapitre, nous avons vu qu'en ce qui concerne l'aspect quantitatif de l'emploi de marqueurs épistémico-modaux d'incertitude, la langue a un effet plus important que la discipline. La fréquence des marqueurs examinés s'est révélée significativement plus élevée dans les corpus anglais et norvégien que dans le corpus français, suggérant que les chercheurs anglophones et norvégiens ont tendance à faire plus de réserves dans leurs textes scientifiques que leurs collègues francophones. Cette observation étaye donc les hypothèses posées au départ, qui se basaient sur des recherches antérieures de phénomènes apparentés, et qui prévoyaient d'une part que les modalisateurs épistémiques seraient plus nombreux dans le discours scientifique anglais que dans le discours scientifique français, et d'autre part que les corpus anglais et norvégien seraient assez proches l'un de l'autre. Bien que cette étude ne porte pas sur l'atténuation en général, mais sur la modalité épistémique uniquement, les différences interlinguistiques repérées invitent à mettre en question la conception traditionnelle selon laquelle l'atténuation constitue un trait caractéristique du discours scientifique dans son ensemble. Si l'atténuation est un trait caractéristique du discours

scientifique anglo-américain, il n'en va pas nécessairement de même pour les discours scientifiques d'autres cultures.

En ce qui concerne les différences quantitatives entre les deux disciplines investiguées, celles-ci se sont révélées très peu marquées. La fréquence de modalisateurs épistémiques est à peu près la même dans les deux disciplines, à l'exception peut-être du corpus anglais, où nous avons observé une différence non négligeable (mais non statistiquement significative) entre les deux groupes, suggérant que les médecins-chercheurs emploient moins de modalisateurs épistémiques que les linguistes. L'hypothèse posée au début, selon laquelle la fréquence de modalisateurs épistémiques serait plus élevée dans le corpus linguistique que dans le corpus médical (selon une différence entre sciences humaines et sciences naturelles) ne peut être validée, étant donné qu'elle ne vaut pas pour les corpus français et norvégien. Cela ne veut pas dire que la différence entre sciences naturelles et sciences humaines, si souvent observée dans d'autres études, n'est pas réelle, mais simplement qu'elle ne vaut pas pour les disciplines de médecine et de linguistique, du moins pas en général, c'est-à-dire toutes langues confondues. Comme nous l'avons vu, les différences disciplinaires dépendent de la langue ; elles sont les plus importantes dans le corpus anglais – et c'est justement sur l'anglais que portent les études antérieures. Cela peut indiquer que c'est dans la recherche internationale et hautement compétitive plutôt qu'à l'échelle nationale que se dessinent le plus clairement les différences entre les domaines scientifiques.

On observe donc une interdépendance des deux variables investiguées. Les différences entre disciplines sont dans l'ensemble faibles, mais plus nettes dans le corpus anglais. En outre, les différences entre langues sont plus importantes dans le corpus linguistique que dans le corpus médical. Cela corrobore en partie l'hypothèse que les médecins-chercheurs constitueraient un groupe plus homogène que les linguistes. Cependant, la variation individuelle de la fréquence de modalisateurs épistémiques est importante au sein des articles de médecine comme de ceux de linguistique.

La variation individuelle se dessine encore plus clairement quand on passe des fréquences relatives pour les modalisateurs épistémiques en général aux fréquences de certains marqueurs spécifiques. Nous avons vu que certains marqueurs se placent assez haut dans la liste en raison d'un usage extensif de ces marqueurs dans un nombre limité d'articles. Les occurrences d'un marqueur peuvent être assez irrégulièrement réparties entre les divers articles. Il semblerait donc que chaque auteur ait des marqueurs préférés ; la dimension des styles individuels, i.e. des préférences linguistiques de l'auteur, ne doit donc pas être négligée. Plus généralement, l'importance de la variation individuelle dans le discours scientifique a été



soulignée auparavant, par le projet KIAP (Fløttum et al. 2006) aussi bien que par Poudat (2006) et Rinck (2006). Cette variation incite à la prudence dans l'interprétation des résultats, et elle rappelle en outre que même au sein d'un genre bien défini et aussi standardisé que l'est l'article de recherche, il y a place pour les styles personnels, du moins dans une certaine mesure.

En résumé, en ce qui concerne les différents types de marqueurs, des ressemblances importantes se manifestent au niveau des langues. Il s'avère que les trois langues investiguées font à peu près usage des mêmes types de marqueurs, les verbes épistémiques étant particulièrement fréquents. On note cependant que l'usage épistémique d'auxiliaires modaux est plus important en anglais et en norvégien qu'en français. Au niveau des disciplines, il s'avère que plusieurs des marqueurs sont spécifiques à l'une des disciplines ; linguistes et médecins-chercheurs semblent donc privilégier des marqueurs différents. Dans le corpus anglais, ces préférences peuvent être rattachées au sémantisme des marqueurs : les médecins-chercheurs préfèrent des marqueurs « objectifs » qui ne présupposent pas un sujet de conscience et qui peuvent masquer l'instance modalisante, alors que les linguistes font un usage extensif de marqueurs « subjectifs » tels que les verbes de perception et le verbe de cognition *assume*, c'est-à-dire des marqueurs qui présupposent un sujet de conscience et qui explicitent l'instance modalisante plutôt que de la dissimuler.



## 7 ANALYSES QUALITATIVES

Les différences de fréquence présentées au chapitre précédent permettent de comparer le type de marqueurs et leur poids dans chaque langue et chaque discipline, mais elles ne disent que peu du *rôle* que jouent les modalisateurs épistémiques dans un texte. Dans le présent chapitre, nous passons de l'approche quantitative à une approche plus qualitative, centrée sur le fonctionnement pragmatique de la modalité épistémique. Nous avons vu que, par leur contenu sémantique, les modalisateurs épistémiques situent l'information transmise sur une échelle de certitude. Ils peuvent cependant comporter une série d'effets pragmatiques différents, comme nous allons le voir par la suite. Dans le présent chapitre, il s'agit donc de coupler l'aspect sémantique à un aspect pragmatique ; le centre d'intérêt passe du sens du langage à l'usage du langage<sup>1</sup>.

Les fonctions pragmatiques remplies par les modalisateurs épistémiques dépendent, bien entendu, du cotexte dans lequel ils apparaissent. Nous étudierons pour cette raison les occurrences des modalisateurs épistémiques dans leurs cotextes respectifs. Il s'agit de mener une étude approfondie d'un nombre restreint de marqueurs et de leurs attestations dans le corpus (sections 7.2, 7.3, 7.4), afin de répondre aux questions suivantes : Dans quels cotextes les modalisateurs épistémiques apparaissent-ils et quels sont les effets pragmatiques qu'ils engendrent ? Nous compléterons cette approche par quelques études de cas qui nous permettent d'examiner l'interaction entre les modalisateurs épistémiques et le texte pris dans son ensemble (section 7.6). Tout d'abord, un compte rendu succinct de ce que l'on connaît sur les fonctions pragmatiques des atténuateurs est nécessaire.

### 7.1 Fonctions pragmatiques de modalisateurs épistémiques

Pour traiter du domaine fonctionnel de la modalité épistémique, différents termes sont en usage. Certains parlent de fonctions communicatives (Varttala 1999, Salager-Meyer 1994), d'autres parlent de valeurs communicatives (Kourilová 1994), de fonctions (pragmatiques) (Hyland 1998 : 156 sqq) ou d'usage communicatif (Hyland 1998 : 189). Quel que soit le terme utilisé, c'est l'aspect pragmatique de la modalité épistémique qui est pointé par ces chercheurs ; plus précisément, il s'agit d'une part des intentions qu'a l'auteur quand il emploie des marqueurs atténuants et d'autre part des effets que ces marqueurs sont supposés

---

<sup>1</sup> Bien que l'aspect pragmatique soit au centre de ce chapitre, nous inclurons aussi un examen du fonctionnement syntaxique de certains marqueurs sélectionnés, afin de fournir une base pour les études des fonctions pragmatiques.

exercer sur le lecteur. C'est avec le même sens que nous proposons de parler ici de *fonctions pragmatiques*.

On distingue traditionnellement deux types majeurs de fonctions : un premier type est lié au contenu, au texte lui-même, – c'est ce que nous appellerons une *fonction contenu* –, alors que le second s'oriente plus directement vers l'allocataire – en l'occurrence le lecteur – c'est-à-dire qu'il s'agit d'une *fonction interpersonnelle*. Plusieurs chercheurs adoptent une telle dichotomie, mais en utilisant des termes variés : Hyland par exemple (1996, 1998), distingue les atténuateurs orientés vers le contenu (« *content-oriented hedges* ») et les atténuateurs orientés vers le lecteur (« *reader-oriented hedges* »). On retrouve des distinctions similaires chez Mauranen (1997), qui parle d'usage épistémique (« *epistemic use* ») et d'usage interpersonnel (« *interpersonal use* ») et chez Varttala (1999), qui parle de fonctions communicatives textuelles (« *communicative textual functions* ») et de fonctions communicatives interpersonnelles (« *communicative interpersonal functions* »). Soulignons tout d'abord qu'en pratique, il est très difficile, voire souvent impossible de distinguer nettement ces deux types de fonctions. La distinction est avant tout théorique, mais elle pointe un fait important : les fonctions pragmatiques des modalisateurs épistémiques peuvent aller, à des degrés divers, dans ces deux directions que sont le contenu et le lecteur. Si un modalisateur épistémique est « orienté vers le contenu », cela signifie que sa présence a avant tout pour rôle de transmettre avec exactitude le degré de certitude que l'auteur souhaite attacher à la proposition ; l'auteur l'utilise pour être précis quant au statut de vérité de la proposition. En cela, la fonction contenu se lie principalement à l'aspect informatif de l'article de recherche (cf. *supra*, section 2.1), mais elle peut aussi être rattachée à l'aspect persuasif, dans la mesure où la précision et l'honnêteté contribuent à l'image d'un chercheur sérieux et par conséquent fiable. Les items qui assument une fonction contenu sont souvent employés pour signaler une incertitude réelle ; cela est illustré par l'exemple 1, où il est très clair qu'on a affaire à une véritable incertitude de la part de l'auteur, et les modalisateurs épistémiques servent à rendre compte avec honnêteté de ce niveau de certitude :

1. De eldste fragmenter av Gulatingsloven er fra noe før 1200, men **det er mulig** loven ble skriftfestet alt på 1000-tallet. **Kanskje** var *-a* og *-at* gått ut av vanlig bruk i Norge alt den gang. **Det er vanskelig å si noe mer nøyaktig om dette.** (noling08)  
'Les fragments les plus anciens de la loi du tribunal de Gulating datent d'environ 1200, mais **il est possible que** la loi ait été mise par écrit déjà au onzième siècle. **Peut-être que** *-a* et *-at* étaient déjà sortis de l'usage général à ce temps-là. **Il est difficile de dire quelque chose de plus exact à ce propos.'**

Quand un marqueur épistémico-modal est « orienté vers le lecteur », sa présence est motivée par le besoin de tenir compte des réactions potentielles du lecteur plutôt que par la nature du contenu informatif lui-même. Sur la base des occurrences de modalisateurs épistémiques repérées dans la présente étude, nous avançons que la fonction interpersonnelle peut à son tour être divisée en deux : un usage qui est motivé par le besoin de se présenter comme poli, modeste et prudent (voir les exemples 2-3), et un autre usage qui est motivé par le désir d'anticiper la critique ou de convaincre le lecteur (voir l'exemple 4 : *we consider this unlikely...*).

2. Ainsi, si le langage a principalement une fonction cognitive, et accessoirement une fonction communicative, comment peut-on définir le rapport que le langage entretient avec la communication ? Il n'est **peut-être** pas inutile de rappeler les deux propositions suivantes: [...] (frling11)
3. Om dei to sanningane om Smiths aktivitet kona har formidla, ville **truleg** William James hevda at dei er produkt av fru Smiths "trade on Betty's and Mr. Jones' truths". (noling13)  
'Sur les deux vérités concernant les activités de Mr. Smith que sa femme a transmises, William James dirait **probablement** qu'elles sont le produit du "trade on Betty's and Mr. Jones' truths" de Mme Smith.'
4. Quality-of-life data were missing at 1 or more time points in 19% of participants, and **we cannot entirely exclude the possibility that** differential loss of participants **may** have affected our results. However, **we consider this unlikely** because treatment assignment was not significantly different among women with complete vs incomplete data (engmed11)

Par leur nature interpersonnelle, les deux usages se lient à l'aspect interactif de l'article de recherche (cf. *supra*, section 2.1). Le premier usage (se présenter comme poli, modeste et prudent) est courant quand l'auteur fait des suppositions concernant les idées d'autres chercheurs (voir l'exemple 3) et quand il émet des remarques critiques sur d'autres travaux (nous en verrons plusieurs exemples dans les sections à venir) ; *l'Autre* est donc directement impliqué dans le texte. Dans le second usage (anticiper la critique ou convaincre le lecteur), on prend en compte et on anticipe les points de vue potentiels de l'Autre : c'est donc là aussi une manifestation claire de l'aspect interactif de l'article de recherche. Ce second usage concerne plus directement encore l'aspect persuasif : il s'agit de prendre position, d'avancer un argument, sans pour autant imposer au lecteur une interprétation unique. Dans cette optique, la modalité épistémique représente pour le chercheur-auteur une stratégie rhétorique qui aide à promouvoir sa propre recherche. Dans l'exemple 4 ci-dessus, cette fonction est illustrée par la troisième marque d'atténuation (*we consider this unlikely*), alors que les deux

premières (*we cannot entirely exclude the possibility that* et *may*) s'orientent plutôt vers le contenu, et ne consistent pas tant à prendre position qu'à signaler une possibilité.

Comme déjà mentionné, il n'y a pas de frontière nette entre les deux types de fonctions, et le fait que les deux premières marques de modalisation dans l'exemple 4 s'orientent plutôt vers le contenu ne veut pas dire qu'elles sont dépourvues de toute fonction stratégique. En envisageant différentes possibilités et les facteurs qui interviennent potentiellement sur le phénomène étudié, les auteurs confèrent à leur propos un certain degré de scientificité : rien n'a été négligé, tout a été considéré. Cela contribue à une image des auteurs en tant que chercheurs sérieux.

Comme nous l'avons vu au chapitre 4, certains des chercheurs qui ont travaillé sur les atténuateurs insistent plus que d'autres sur la fonction rhétorique ou interpersonnelle. Mais alors que certains la considèrent comme une stratégie positive qui témoigne de politesse, de modestie ou encore d'un esprit ouvert (voir par exemple Hyland 1998 et Myers 1989), d'autres la dépeignent comme une stratégie de détour, un artifice, autrement dit une stratégie qui se ramène presque à la ruse (voir Foullioux et Tejedor de Felipe 2004 : 114 et Beaufrère-Bertheux 1997). Ces différentes manières de concevoir le même phénomène relèvent peut-être de différences culturelles : il est raisonnable de supposer que les chercheurs issus d'une culture scientifique où l'usage de ces marques est abondant auront tendance à considérer l'atténuation comme une stratégie positive alors que ceux issus d'une culture scientifique où l'usage de tels marqueurs est restreint seront enclins à la considérer d'une manière moins favorable. Quoi qu'il en soit, c'est cette fonction rhétorique ou interpersonnelle qui retient l'intérêt des chercheurs (voir aussi Varttala 2001 : 14) et plusieurs sont ceux qui voient principalement dans le phénomène d'atténuation une stratégie de politesse (Brown et Levinson 1987, Myers 1989). Néanmoins, comme Salager-Meyer (1997) le souligne à juste titre, il ne faut pas oublier qu'un atténuateur est souvent utilisé pour véhiculer une incertitude réelle, et qu'il peut être employé par l'auteur dans le but de transmettre avec exactitude ses observations et ses conclusions. En fait, une enquête menée par Lewin (2005) a dégagé que d'après les chercheurs-auteurs eux-mêmes, c'est l'incertitude réelle qui constitue le motif principal d'utilisation des atténuateurs : les participants de ce sondage maintenaient que leur emploi d'atténuateurs n'était pas motivé par la politesse ou la modestie, mais simplement par un désir d'être précis et de transmettre exactement leur degré de certitude. Cependant, il convient de rappeler que les auteurs ne sont évidemment pas toujours conscients de leurs propres intentions dans l'emploi d'atténuateurs (Hyland 1998 : 158).

La dichotomie présentée ci-dessus, entre une fonction “contenu” et une fonction interpersonnelle, englobe plusieurs usages plus concrets. Bien qu’on ne puisse pas, en tant que lecteur, se prononcer sur les motifs qui ont conduit un auteur à employer des modalisateurs spécifiques, on peut étudier dans quels co(n)textes les modalisateurs épistémiques apparaissent et réfléchir sur les fonctions pragmatiques que ceux-ci sont censés produire. C’est ce que Hunston propose dans son article de 1994. A partir d’une étude de 10 articles expérimentaux en linguistique appliquée, elle établit une liste de contextes dans lesquels les modalisateurs épistémiques sont particulièrement fréquents. Ces contextes incluent entre autres la présentation de l’objectif de l’étude et de ses résultats, l’interprétation et la discussion de résultats, la description de limites et d’implications, et les propositions et préconisations formulées sur la base des résultats de l’étude.

En ce qui concerne les articles en médecine, il est bien connu que les atténuateurs s’accumulent dans les sections Discussion et Introduction. Salager-Meyer (1994) observe que dans l’introduction de leurs articles, les médecins-chercheurs utilisent des atténuateurs dans les comptes rendus de l’état de l’art afin d’indiquer une lacune dans les recherches antérieures et de justifier ainsi leur propre recherche. Dans la section Discussion, les atténuateurs font partie de la formulation des hypothèses et ils servent à présenter des conclusions et à évaluer des résultats.

Les contextes et fonctions pragmatiques mentionnés par Hunston (1994) et Salager-Meyer (1994) se retrouvent dans le corpus KIAP. Dans l’exemple 5 par exemple, les modalisateurs *suggest* et *may* servent à atténuer la force d’une conclusion qui repose sur les résultats obtenus :

5. Our study **suggests** that a diet high in fruits and vegetables containing carotenoids, including raw carrots and tomato products, **may** be important in the prevention of a very lethal form of cancer in women. (engmed10)

Dans l’exemple 6, l’objectif est de signaler une limite de l’étude :

6. Comme nous n'avons analysé que les hospitalisations de 30 jours et moins dans les centres hospitaliers de courte durée, il est **possible** que notre étude ne tienne pas compte de certains coûts d'hospitalisation pour l'asthme. (frmed02)

Dans l’exemple 7, l’auteur présente quelques préconisations fondées sur les résultats de l’étude :

7. Enda bedre hadde det **muligens** vært med en etterundersøkelse utført av en helt uhildet person, et krav man i fremtiden **kanskje** bør stille i forbindelse med kliniske studier. (nomed05)

‘**Peut-être** aurait-il été encore mieux si on avait des examens de contrôle effectués par une personne complètement impartiale, une condition qu’on devrait **peut-être** poser à l’avenir...’

Dans l’exemple 8 enfin, *suggest* est employé dans un cotexte qui sert à signaler le caractère indéfini ou contradictoire des conclusions issues d’études antérieures. Cette difficulté permet aux auteurs de justifier leur propre recherche. Dans ce cotexte, caractéristique de l’introduction, les modalisateurs épistémiques apparaissent typiquement quand l’auteur rend compte des idées et des conclusions d’autres chercheurs :

8. Some studies have **suggested** that first teenage pregnancies have a higher frequency of adverse perinatal outcomes. 1-2 However, there is argument about whether this is an independent association 1-2 or explained by confounding factors 3-5. [...] The aims of this study were to determine whether teenage pregnancy was associated with increased rates of adverse perinatal outcome, whether the association differed by parity, and whether any associations were independent of confounding factors. (engmed01)

Une tentative de classification des exemples 5-8 pourrait être que 5 (présenter une conclusion d’une manière atténuée) et 6 (signaler des limites d’une étude) appartiennent à la catégorie « fonction contenu », alors que 7 (donner des préconisations) appartient à la catégorie « fonction interpersonnelle ». L’acte de donner des préconisations présuppose qu’on a pris position et qu’on souhaite influencer sur l’autre, mais comme le montre l’exemple 7, ces préconisations sont cependant émises d’une manière prudente et quelque peu hésitante. L’exemple 8 pourrait relever des deux catégories : le verbe *suggest* sert ici à rapporter les résultats d’autres chercheurs (ces résultats sont qualifiés de non vérifiés) et il remplit ainsi une fonction contenu. L’exemple relève aussi de l’interpersonnel, dans la mesure où le modalisateur sert à préserver la face des autres chercheurs (leurs résultats sont en partie contestés).

Dans 5, la motivation dans l’emploi des modalisateurs n’est pas nécessairement une incertitude réelle chez les auteurs, ceux-ci peuvent les avoir employés simplement pour répondre aux exigences du style scientifique. Selon un grand nombre de chercheurs (voir par exemple Banks 1994, Salager-Meyer 1997, Varttala 2001 : 244-245), l’atténuation fait partie intégrante d’un discours scientifique standardisé : elle s’est conventionnalisée jusqu’à être routinisée et l’emploi d’un atténuateur peut donc être motivé par des raisons simplement stylistiques, plus précisément par un désir de répondre aux attentes du genre. Dans les articles



de médecine rédigés en anglais, par exemple, il existe une structure quasiment standardisée pour présenter des hypothèses et des conclusions : c'est celle qui est exemplifiée dans l'exemple 5, à savoir une structure qui se compose d'un sujet + le verbe *suggest* + une subordonnée introduite par *that* et contenant l'auxiliaire modal *may* (Vold 2006b). L'institutionnalisation de certaines tournures constitue donc un facteur important dans l'emploi d'atténuateurs. D'après Myers (1989 : 13) et Kourilová (1994 : 643), l'atténuation est à ce point standardisée dans le discours scientifique anglais qu'elle est une composante obligatoire de toute nouvelle assertion : celles-ci doivent être présentées d'une manière atténuée, quelle que soit la force des preuves ou la certitude de l'auteur. Ce n'est qu'une fois que les nouvelles assertions ou observations ont été acceptées par la communauté scientifique qu'elles peuvent être reprises sans marque d'atténuation<sup>2</sup>.

Comme déjà mentionné, les études sur le fonctionnement pragmatique des atténuateurs dans le discours scientifique se limitent dans une large mesure à des études portant sur l'anglais. Les articles étudiés par Salager-Meyer (1994) et Hunston (1994), par exemple, sont tous rédigés en anglais. Pareillement, il existe peu d'études qui comparent le fonctionnement pragmatique des atténuateurs à travers différentes disciplines. L'étude de Varttala (2001), qui compare trois disciplines et qui aborde l'aspect fonctionnel aussi bien que l'aspect quantitatif, constitue un cas à part de ce point de vue. On parle cependant souvent du fonctionnement pragmatique en termes généraux (en disant, par exemple, que l'atténuation est une stratégie rhétorique employée par l'auteur pour faire des réserves, se prémunir contre d'éventuelles remarques critiques, apparaître poli et modeste etc.), mais cet aspect est rarement décrit à partir d'une étude systématique des occurrences attestées dans un corpus, comme nous allons le faire ici. Les études de Varttala (2001) et de Hyland (1998) font figure d'exceptions, en ce qu'elles abordent de manière plus approfondie le domaine fonctionnel et se basent sur des exemples authentiques relevant d'un corpus. Cependant, comme déjà mentionné, ces études portent toutes deux sur des articles de recherche anglais. En raison de la prépondérance d'études sur l'anglais, nous limiterons dans ce qui suit nos considérations à des exemples français et norvégiens. L'approche va principalement du bas vers le haut ; c'est-à-dire que c'est à partir des exemples attestés dans le corpus que se sont réalisées nos descriptions des fonctions pragmatiques. Si nous sommes partie de la dichotomie traditionnelle évoquée ci-

---

<sup>2</sup> Cependant, il n'est pas facile pour le chercheur individuel de savoir justement *quand* quelque chose est à considérer comme accepté par la communauté scientifique, ce qui implique qu'on aura souvent des atténuateurs « superflus », attachés par précaution à des assertions encore longtemps après qu'elles se sont mises à gagner du terrain.

dessus, les catégories plus fines (cf. les sections 7.2.6, 7.3.5 et 7.4.6) n'ont pu être établies qu'à partir des exemples attestés.

En abordant le domaine fonctionnel, nous nous centrons sur un nombre limité de marqueurs, étant donné qu'une étude détaillée de l'ensemble des occurrences de chaque marqueur n'est pas envisageable. Les marqueurs sélectionnés seront *sembler*, *paraître* et *pouvoir* pour le français, et *synes*, *se ut* et *kunne* pour le norvégien. Ces marqueurs correspondent aux marqueurs principaux de la modalité épistémique dans les deux langues : comme nous l'avons vu, *sembler* et *paraître* dominent en linguistique française, *sembler* et *pourrait* en médecine, *kan* et *se ut* dominent en linguistique norvégienne et *kan* et *synes* en médecine (cf. les tableaux 3 et 4 dans la section 6.4.2). En même temps, ces marqueurs sont des quasi-équivalents : *kunne* est l'équivalent usuel de *pouvoir*, et *synes* et *se ut (til/som)* sont deux correspondants habituels de *sembler* et de *paraître*. Cela invite à comparer non seulement le fonctionnement pragmatique de ces marqueurs, mais aussi leurs fréquences et leur emploi général à travers les deux langues investiguées afin de déterminer dans quelle mesure ces marqueurs s'utilisent effectivement comme des équivalents. Les sections qui suivent abordent donc, outre le domaine fonctionnel, l'aspect quantitatif de l'emploi de ces marqueurs ainsi que la question du cotexte immédiat.

Le fait de se limiter à un nombre restreint de marqueurs empêche de brosser un tableau complet des fonctions pragmatiques de la modalité épistémique, dans la mesure où celles-ci dépendent du marqueur ; les différents effets pragmatiques transmis par un marqueur tel que *possible*, par exemple, ne correspondent pas tout à fait à ceux transmis par un marqueur tel que *sembler*. Néanmoins, le fait que les marqueurs sélectionnés correspondent aux marqueurs les plus récurrents nous permet de postuler que les fonctions observées représentent les principales fonctions pragmatiques remplies par la modalité épistémique dans les deux disciplines retenues. En outre, les études de cas nous permettront de compléter cette analyse des fonctions pragmatiques par quelques observations sur d'autres marqueurs épistémico-modaux.

## 7.2 Sembler et paraître

*Sembler* et *paraître* figurent en tête de la liste de modalisateurs épistémiques employés dans les articles de recherche français (cf. chapitre 6). Ces verbes, qui se caractérisent par des constructions syntaxiques analogues, ainsi que des sémantismes similaires, sont souvent traités simultanément (voir par exemple Nølke 2001, Riegel et al. 1998), un procédé que nous suivrons aussi dans ce chapitre où nous étudierons de plus près leur emploi dans les articles de linguistique et de médecine. Comme déjà mentionné, ce sont le cotexte immédiat et les fonctions pragmatiques qui seront au centre de ces analyses plus approfondies des marqueurs. Avant d'aborder ces questions sera donnée une présentation des différentes fonctions syntaxiques de *sembler* et *paraître*, ainsi qu'une présentation de leur contenu sémantique et de leurs configurations polyphoniques. Cela fournira une base pour l'étude des attestations de ces verbes dans le corpus.

### 7.2.1 Fonctionnement syntaxique

Schématiquement, les emplois de *sembler* et de *paraître* se scindent en trois constructions syntaxiques différentes (voir aussi Bourdin 1989 : 47) :

#### 1) *Semi-auxiliaires modaux*

En tant que semi-auxiliaires modaux, *sembler* et *paraître* modifient un verbe lexical à l'infinitif. Leur rôle est d'affecter d'une modalité épistémique le procès exprimé par l'infinitif, indiquant ainsi qu'il s'agit d'une impression subjective (voir Riegel et al. 1998 : 254). Bourdin nomme cet emploi « l'emploi infinitif » (Bourdin 1989 : 47).

1. Et puis, la thèse benvenistienne suppose au moins clarification. Elle **semble** éliminer toute autre forme de médiation sémiotique (frling12)
2. Ainsi l'instruction topologique, même purement configurationnelle et déspatialisée (c'est-à-dire déliée de l'espace perçu), **paraît** céder le pas devant un principe plus ouvert et plus riche de définition-délimitation de deux 'segments' ou 'phases' par le biais de leur 'mise en contact'. (frling15)

A deux occasions, *sembler* employé comme semi-auxiliaire entre dans une construction impersonnelle :

3. Il **semble** aller de soi pour un projet "scientifique" sérieux de poser des barrières; on propose, par exemple, de s'occuper des modes de fonctionnement du langage et non de ses origines. (frling12)

Quand le verbe lexical est *être*, celui-ci est parfois omissible, et la construction avec *sembler* comme auxiliaire peut alors alterner avec une construction où *sembler* fonctionne comme verbe attributif, cas que nous traitons au point suivant :

4. Le budénoïde, un corticoïde à action topique plus élevée et un passage systémique plus faible, **semble** être efficace avec moins d'effets secondaires [21], [22]. (frmed06)

## 2) *Verbes attributifs* :

*Sembler* et *paraître* peuvent être employés comme verbes attributifs, et se construisent alors avec un sujet et un attribut du sujet. L'attribut du sujet prend le plus souvent la forme d'un adjectif, mais il peut aussi prendre la forme d'un participe ou d'un nom :

5. L'intention de l'actant **semble** moins déterminante. (frling07)
6. Les informations fournies par les douanes **paraissent** corrélées aux autres données exposées (frmed19)

Dans cet emploi, *sembler* et *paraître* font parfois partie d'une construction impersonnelle :

7. Dans ces conditions, il **paraît** concevable qu'un certain nombre de SD aient échappé aux statistiques. (frmed19)

Employés comme verbes attributifs, *sembler* et *paraître* ne sont que des variantes du verbe copule *être*, mais, bien entendu, des variantes plus riches en sens : là où *être* n'exprime que la simple existence d'une relation, *sembler* et *paraître* ajoutent à cette idée de relation une précision modale et épistémique, en introduisant une forme de réticence et d'atténuation (voir aussi Riegel et al. 1998 : 237)<sup>3</sup>.

---

<sup>3</sup> Les occurrences de *sembler* et *paraître* dans les exemples suivants ont aussi été rangées dans cette catégorie (verbes attributifs), bien que leur rôle syntaxique ne soit pas évident :

Ces exemples ont pourtant une interprétation temporelle; celle-ci **semble** induite par les contraintes sémantico-lexicales du verbe *annoncer* [20]. (frling10)

La sémantique prépositionnelle me **paraît** d'abord déterminée par des valeurs renvoyant à la dépendance, au contrôle, à l'appropriation réciproque, à l'anticipation et à l'attente (frling15)

On peut considérer que *sembler* et *paraître* remplacent ici *être*, c'est-à-dire qu'ils sont des variantes atténuées du verbe *être*, et il paraît donc approprié de les analyser comme des verbes attributifs, dont l'attribut du sujet correspond à un syntagme participial. D'un autre côté, il est clair que nous avons ici affaire à des constructions passives. La séquence verbe au présent + participe passé pourra être conçue comme une forme composée d'un verbe au passif, et les groupes prépositionnels introduits par *par* sont des compléments d'agent. En effet, s'il n'y avait pas eu *sembler* et *paraître*, mais *être*, une analyse où *être* était conçu comme auxiliaire dans une forme verbale au passif aurait été tout à fait convenable. Cependant, il ne semble pas tout à fait approprié d'analyser

### 3) Constructions impersonnelles

Enfin, *sembler* et *paraître* s'emploient dans des constructions impersonnelles du type *il semble que, il paraît que...* Dans ces constructions, *il* impersonnel est le sujet formel, et le sujet logique prend la forme d'une complétive (voir les exemples 8 et 9) ou, plus rarement d'une construction infinitive (voir l'exemple 10)<sup>4</sup>. La construction permet l'adjonction d'un objet datif, normalement sous forme d'un pronom datif, mais à part l'éventuelle présence d'un tel objet, il n'y a aucun complément lié au verbe.

8. Il **semble** que cette alternance singulier/pluriel dépende du renvoi anaphorique supporté par le déterminant déictique (frling20)
9. Loin de considérer, comme on le fait le plus souvent, les valeurs aspectuelles, subjectives et qualitatives comme des suppléments que la reconstruction linguistique devrait dériver dans un deuxième temps, il nous **paraît** qu'il faut les inscrire au cœur des "motifs" les plus originels attribués aux prépositions. (frling15)

---

*sembler* et *paraître* comme des auxiliaires quand ils ne sont pas suivis d'un infinitif (cf. Riegel et al.1998 : 252, 254). On pourrait certes facilement insérer un tel infinitif, à savoir l'infinitif *être*, sans que cela change l'interprétation : *La sémantique prépositionnelle me paraît d'abord être déterminée par... ; celle-ci semble être induite par...* On arguera alors qu'il y a ellipse de *être*, et que *sembler* et *paraître* sont à considérer comme des semi-auxiliaires modaux.

Deux analyses sont donc possibles face à de telles occurrences (seulement le premier exemple est analysé ici, mais la même analyse vaut pour le second). Elle inclut uniquement le niveau de la *fonction syntaxique*) :

1.) Sujet	verbe	attribut du sujet		
CELLE-CI	SEMBLE	INDUITE PAR LES CONTRAINTES SEMANTICO-LEXICALES DU VERBE ANNONCER		
2.) Sujet	forme verbale composée		complément d'agent	
	forme finie du verbe (auxiliaire)	forme non finie du verbe (infinitif passé)		
CELLE-CI	SEMBLE	Ø (ETRE)	INDUITE	PAR LES CONTRAINTES SEMANTICO-LEXICALES DU VERBE ANNONCER

L'analyse de telles phrases se prête clairement à la discussion, mais nous avons choisi de traiter ces occurrences de *sembler* et *paraître* comme des verbes attributifs. Cette analyse est en concordance avec celle proposée par Faarlund et al. (2002 : 850-853) au sujet de constructions similaires. Faarlund et al. voient ces constructions comme des paraphrases d'une proposition principale contenant une subordonnée qui à son tour contient un verbe au passif : *Il semble que celle-ci soit induite par les contraintes...* Dans la paraphrase, le sujet de la subordonnée est devenu sujet principal, et le participe de la subordonnée se présente comme un attribut du sujet (Faarlund et al. 2002 : 851).

<sup>4</sup> Une remarque terminologique s'impose : bien que cet emploi de *sembler* et *paraître* soit traité comme construction impersonnelle, il convient de noter que ces verbes peuvent tout à fait admettre un sujet impersonnel dans d'autres emplois aussi, cf. les exemples 3 et 7 ci-dessus. Par la catégorie *construction impersonnelle*, nous entendons ici les constructions du type *il semble/paraît que...*

10. *Il me semble l'avoir déjà vu quelque part* (Pedersen et al.1992 : 390)<sup>5</sup>

Les incises *semble-t-il*, *paraît-il*, éventuellement précédées d'un pronom datif, sont de simples variantes des constructions *il semble que* et *il paraît que* :

11. Ce problème a sa source, me **semble-t-il**, dans la méthode elle-même, qui n'envisage pas de façon systématique une solution de type instructionnel, bien que D.L. l'évoque à plusieurs reprises. (frling13)

L'insertion d'un pronom datif est en principe possible dans les trois types de construction syntaxique évoqués ci-dessus, bien qu'elle semble rare devant *sembler* / *paraître* employés comme auxiliaires modaux.

Outre les emplois mentionnés ci-dessus, *paraître* connaît au moins deux autres emplois, à savoir *paraître* dans le sens de *devenir/être visible*, *apparaître* et dans le sens de *être publié* (le Petit Robert 1995 : 1582). Non modaux, ces emplois ne sont pas pertinents pour la présente étude. Par ailleurs, il n'y a aucune manifestation dans le corpus de *paraître* dans le sens de *devenir/être visible*, *apparaître*. Le verbe apparaît trois fois dans le sens de *être publié*, mais étant donné qu'il s'agit là d'une signification tout à fait dépourvue de valeur épistémico-modale, ces attestations ont été exclues de l'analyse.

### 7.2.2 Sémantisme

Nous avons vu que les deux verbes sont susceptibles d'exprimer une modalité épistémique. Selon Kibbee (1995), le verbe *sembler* se situe à la frontière entre les expériences de sens et les jugements de l'esprit. Il peut donc rendre compte d'impressions sensorielles, comme dans *Pierre semble fatigué* et de jugements de l'esprit : *il me semble qu'elle a raison*. D'après Bourdin (1989 : 55), la construction attributive va de pair avec les impressions sensorielles alors que les constructions infinitive et impersonnelle sont plus appropriées quand il s'agit d'un jugement de l'esprit. Dans tous les cas, *sembler* exprime une modalité épistémique en ce qu'il marque, par son sémantisme même, qu'il s'agit d'une *impression* ou d'une *supposition* et non d'un fait. *Ssembler* peut donc être considéré comme un marqueur intrinsèquement épistémico-modal, peut-être avec l'exception d'un certain cas de figure (cf. exemple 13) auquel nous reviendrons ci-dessous.

---

<sup>5</sup> L'exemple 3 ci-dessus (*Il semble aller de soi ...*) n'entre pas dans la même catégorie que 10, parce qu'il est une variante modalisée de l'expression figée « il va de soi ». Cette occurrence de *sembler* est donc classée dans la catégorie auxiliaire modal.

*Paraître* se situe aussi quelque part entre les sens et l'esprit, mais l'aspect sensoriel semble ici l'emporter sur celui de l'esprit (Bourdin 1989). En outre, le sémantisme de *paraître* est plus compliqué que celui de *sembler*, étant donné que *paraître* connaît des emplois plus variés. Les différentes nuances de sens de *paraître* sont dans une certaine mesure liées aux différentes constructions syntaxiques. Les constructions infinitive et attributive sont des variantes modalisées d'énoncés plus catégoriques, tout comme le sont les constructions correspondantes avec *sembler* (voir Riegel et al. 1998 : 237, 254). Ainsi, l'exemple 2 ci-dessus (avec *paraît céder*) est une variante atténuée de l'énoncé 2' suivant :

2'. Ainsi l'instruction topologique, même purement configurationnelle et déspatialisée (c'est-à-dire déliée de l'espace perçu), **cède** le pas devant un principe plus ouvert et plus riche de définition-délimitation de deux 'segments' ou 'phases' par le biais de leur 'mise en contact'.

et l'exemple 7 ci-dessus est une variante atténuée de

7'. Dans ces conditions, il **est** concevable qu'un certain nombre de SD aient échappée aux statistiques.

La construction attributive privilégiée, comme nous l'avons vu, une lecture du type « expérience des sens ». Il nous semble cependant qu'il existe certains cas où *paraître* suivi d'un attribut n'a pas de valeur épistémico-modale, ou du moins une valeur épistémico-modale très affaiblie. Tel est le cas dans 12, où *paraître* n'est pas une variante atténuée de *être*, mais où il semble simplement renvoyer à une sensation perceptive, sans pour autant mettre en question la valeur de vérité de la proposition :

12. Il est également difficile de conserver cette préposition devant l'autre préposition, appartenant à l'interrogative qui suit, car ces combinaisons **paraissent** toujours "bizarres":  
Je l'ai interrogé sur de quoi il était question (frling01)

Il existe un exemple similaire avec *sembler* :

13. La quantification par unité de mesure pose un problème particulier, dans ce sens qu'elle s'applique difficilement à des référents humains. Ainsi les exemples suivants **semblent-ils** franchement bizarres: [40] ?\*Des enfants, il y en avait 150 kilos dans l'ascenseur. [41] ?\*Des voyageurs, on en a trouvé une tonne dans le train. (frling06)

Ces occurrences ne passent pas vraiment les tests décrits au chapitre 6. Il ne semble pas approprié d'enchaîner avec une phrase d'incertitude :

12' ?? ces combinaisons **paraissent** toujours "bizarres", mais ce n'est pas nécessairement ainsi

et *paraître* et *sembler* ne sont pas remplaçables par d'autres modalisateurs épistémiques sans que cela entraîne un changement sensible de sens. Finalement, il est difficile de formuler des variantes plus catégoriques de ces énoncés : si l'on substitue *être* à *sembler* ou *paraître*, les énoncés ne deviennent pas forcément plus sûrs ou plus catégoriques ; par contre, le sens reste à peu près le même. Ces cas sont pourtant très rares dans notre corpus et se limitent aux deux exemples cités ci-dessus. Il convient de noter que les adverbes *toujours* et *franchement* jouent aussi un rôle dans ces exemples ; ils semblent orienter l'interprétation dans la direction « certitude ». Dans la grande majorité des cas, les occurrences de *paraître* (et de *sembler*) comme verbe attributif affectent la relation entre le sujet et la qualité (l'attribut) d'une modalité épistémique, et sont employées comme des variantes atténuées de *être*.

Dans la locution impersonnelle *il paraît que*, *paraître* est un marqueur d'évidentialité indiquant que l'évidence est du type ouï-dire, et par conséquent, que l'information est incertaine. Dans cet emploi, *paraître* exprime donc une modalité épistémique, en ce qu'il véhicule un certain degré d'incertitude du locuteur : le locuteur ne sait pas (ou bien il fait semblant de ne pas savoir), il ne fait que répéter ce qu'il a entendu dire. La non-prise en charge va souvent de pair avec l'expression d'une certaine méfiance par rapport à la fiabilité de ce qui est rapporté, et c'est effectivement le cas pour *il paraît que*. Nølke (2001) montre que la locution *il paraît que*, paraphrasable par *on dit que*, perd sa valeur d'ouï-dire dès qu'on ajoute un pronom datif, ou, même, dès que *paraître* est modifié (par exemple par l'adjonction d'un attribut du sujet). Ainsi, tandis que *il semble que* et *il paraît que* sont deux locutions qui diffèrent en sens – la première étant paraphrasable par *j'ai l'impression que* et la seconde par *on dit que* – *il semble que* et *il paraît que* avec l'intercalation d'un pronom datif sont des locutions presque synonymes (Nølke 2001 : 26) : *sembler* indique que la conclusion exprimée par la proposition est tirée sur la base d'inférences, et *paraître* indique qu'elle est tirée sur la base d'impressions visuelles (ibid. : 27). L'adjonction d'un pronom datif fait donc disparaître la valeur d'ouï-dire, mais l'idée d'atténuation reste, et *paraître* correspond donc à une modalité épistémique dans *il me paraît que* aussi bien que dans *il paraît que*, bien que la nature de l'évidence soit différente : un raisonnement subjectif dans le premier cas et un bruit ou le témoignage d'un tiers dans le second. Ces locutions sont donc des marqueurs évidentiels, mais elles sont aussi des marqueurs épistémico-modaux, dans la mesure où elles véhiculent un sens d'atténuation et de réticence.



Outre ces emplois épistémico-modaux, il existe, comme mentionné plus haut, certains autres emplois de *paraître* qui ne peuvent pas être assimilés à la modalité épistémique. (Plus précisément, il s'agit de *paraître* dans le sens de *devenir/être visible*, *apparaître* et dans le sens de *être publié*). *Paraître* n'est donc pas intrinsèquement épistémico-modal dans la même mesure que *sembler*, qui est plus monosémique.

Pour conclure, on peut dire qu'alors que la modalité épistémique est un trait pratiquement inhérent au verbe *sembler*, qui possède un noyau de sens assez bien déterminé, *paraître* est un verbe polysémique susceptible de véhiculer des valeurs diversifiées. Cependant, les trois constructions syntaxiques présentées ci-dessus s'associent toutes à la modalité épistémique, mais de diverses manières. En tant que verbe attributif ou verbe modal, *paraître* fait de l'énoncé dont il relève une variante modalisée de l'énoncé « catégorique » correspondant. Dans la construction impersonnelle *il (pronom datif) paraît que*, *paraître* transmet une épistémicité modale et évidentielle.

### 7.2.3 *Ssembler et paraître* – marqueurs polyphoniques

Dans son article datant de 2001, Nølke propose une analyse détaillée des locutions impersonnelles *il (me) semble que* et *il (me) paraît que* dans le cadre d'une théorie polyphonique. Une telle analyse nous aidera à expliquer les fonctions pragmatiques de ces locutions (voir *infra*, section 7.2.6) ainsi qu'à détailler leur contenu sémantique et à élucider les fines différences de sens qui existent entre elles. Pour ces raisons, nous donnerons dans ce qui suit un bref résumé de l'analyse polyphonique proposée par Nølke.

Selon Nølke, il est possible de distinguer deux points de vue (ci-après *pdv*) dans des énoncés *il semble que p* et *il paraît que p*, où *p* correspond à un contenu propositionnel / une proposition, par exemple *Marie est malade* :

- a) *Il semble que Marie soit malade*
- b) *Il paraît que Marie est malade*

Ces structures sont polyphoniques dans le sens où plusieurs « voix » se laissent entendre. D'un côté, on a le point de vue exprimé par le contenu propositionnel (*pdv*<sub>1</sub>), et de l'autre côté, on a un second point de vue (*pdv*<sub>2</sub>) qui est un commentaire du premier. Cette configuration polyphonique se schématise par la formule suivante (tirée de Nølke, Fløttum et Norén 2004), où *p* est un contenu propositionnel et *K* est un commentaire à propos de *p* :

pdv<sub>1</sub> : *p*

pdv<sub>2</sub> : (K) *p*

Les exemples a et b correspondent bien tous deux à cette formule polyphonique, mais la polyphonie de *il semble que p* est pourtant différente de celle de *il paraît que p*. Dans le premier cas, a), il s'agit d'une polyphonie interne, c'est-à-dire que les voix qui se laissent entendre sont toutes deux les voix du locuteur, manifestées à travers les deux images du locuteur, *L* et *l* (voir ci-dessous) (Nølke 2001 : 23). Dans le second cas, b), la polyphonie est de nature externe, c'est-à-dire que le point de vue exprimé par la proposition n'est pas associé au locuteur, mais à quelqu'un d'autre (ibid. : 19).

Regardons de plus près en quoi consistent ces différences. Dans le cadre de la théorie polyphonique de Nølke et ses collègues scandinaves (appelée *la théorie scandinave de la polyphonie linguistique (ScaPoLine)*, voir Nølke et al. 2004), deux questions sont censées particulièrement importantes : *qui est responsable du point de vue exprimé par la proposition ?* et *quelle est l'attitude du locuteur par rapport au point de vue exprimé ?* Commençons par la polyphonie interne : dans ce cas, les deux voix qui se laissent entendre appartiennent, comme il a déjà été mentionné, au locuteur. Cela s'explique par le fait que le locuteur puisse créer deux images de lui-même (Nølke 2001 : 23) dont l'une est le locuteur de l'énoncé actuel (*le locuteur de l'énoncé*, abrégé *l*), et l'autre est *le locuteur textuel* (abrégé *L*). L'image du locuteur de l'énoncé est restreinte au moment précis de l'énonciation (Nølke et al. 2004 : 38) : le locuteur de l'énoncé n'existe que *hic et nunc*, alors que le locuteur textuel existe indépendamment de l'événement énonciatif. Ces deux instances du locuteur peuvent s'associer à deux points de vue différents<sup>6</sup>. Dans un énoncé *il semble que p*, *p* est le résultat d'un dialogue intériorisé (Nølke 2001 : 22), c'est-à-dire un dialogue que le locuteur mène avec lui-même<sup>7</sup>. Le locuteur textuel s'associe normalement au point de vue exprimé dans la proposition, mais ce point de vue est une conclusion tirée sur la base d'indices peu précis. Comme il ne peut pas assumer sans réserves la vérité de *p*, le locuteur de l'énoncé introduit *p* par la locution épistémico-modale *il semble que*. Le locuteur textuel s'associe donc (dans

---

<sup>6</sup> Dans son article datant de 2001, Nølke utilise les termes *locuteur-en-tant-que-tel* et *locuteur-en-tant-qu'individu*, mais dans l'ouvrage de 2004, ces termes d'origine ducrotienne sont quittés en faveur des termes plus simples *locuteur de l'énoncé* et *locuteur textuel*, respectivement.

<sup>7</sup> Cela corrobore d'ailleurs l'observation faite par Bourdin (1989) selon laquelle la construction impersonnelle va de pair avec un jugement de l'esprit.

l'interprétation par défaut<sup>8</sup>) au point de vue exprimé par *p* (pdv<sub>1</sub>), tandis que le locuteur de l'énoncé s'associe au point de vue exprimé par l'énoncé *il semble que p* (pdv<sub>2</sub>). Il y a donc un *lien de responsabilité* entre le locuteur textuel et le point de vue exprimé, mais un *lien de non-responsabilité* entre pdv<sub>1</sub> et le locuteur de l'énoncé (Nølke et al. 2004 : 45). L'analyse polyphonique de l'énoncé *Il semble que Marie soit malade* peut être schématisée comme suit :

pdv<sub>1</sub> (pris en charge par le locuteur textuel) : Marie est malade

pdv<sub>2</sub> (pris en charge par le locuteur de l'énoncé) : Il semble que Marie soit malade.

Ainsi est décrite la responsabilité du point de vue exprimé. Quant à l'attitude du locuteur par rapport à ce point de vue, on peut dire que le locuteur se distancie dans une certaine mesure de son propre point de vue en exprimant un sens d'incertitude et de réticence. Ce sens est corroboré par l'emploi du mode subjonctif, qui correspond ici à l'idée d'une possibilité non confirmée. Selon Nølke, cette distance diminue dès l'adjonction d'un pronom datif à la première personne, car un tel pronom indique que le locuteur « prend [... la] vérité [de *p*] à son propre compte » (Nølke 2001 : 24). L'adjonction d'un pronom datif à la première personne ancre donc explicitement la responsabilité de pdv<sub>1</sub> chez le locuteur, de la même manière qu'un pronom datif à la troisième personne l'ancre chez quelqu'un d'autre, comme dans *il semble au médecin qu'une nouvelle crise va survenir* (Nølke 2001 : 25). L'adjonction d'un objet datif fait passer du subjonctif à l'indicatif.

Avec *il paraît que* en revanche, la polyphonie exprimée est de nature *externe*, c'est-à-dire que le point de vue exprimé par la proposition n'est pas associé au locuteur, mais à une autre instance. Cela est caractéristique de *il paraît que*, qui est un marqueur évidentiel du type ouï-dire. La locution *il paraît que* marque clairement que le locuteur n'est pas *responsable* du contenu posé, mais le locuteur de l'énoncé est naturellement, comme toujours, responsable de l'énoncé. Cependant, il ne fait que rapporter le point de vue d'une autre instance, et il ne donne pas son propre avis sur la vérité ou la fausseté de ce point de vue. Il ne dit pas s'il l'accepte ou s'il le réfute, et dans la terminologie de la ScaPoline (Nølke et al. 2004), il y a donc un *lien de non-responsabilité non réfutatif* :

pdv<sub>1</sub> (pris en charge par quelqu'un de différent du locuteur) : Marie est malade

pdv<sub>2</sub> (pris en charge par le locuteur de l'énoncé) : Il paraît que Marie est malade

---

<sup>8</sup> La locution *il semble que* (dépourvue d'objet datif) peut en théorie introduire un point de vue dont le locuteur se dissocie, mais ces cas sont extrêmement rares (voir Nølke 2001 : 21).

Il n'y a aucun exemple de *il paraît que* dans le corpus. Cela n'est pas très surprenant, étant donné que la recherche scientifique n'est pas un lieu où il convient de discuter des ouï-dire. Il y a, en revanche, certaines occurrences de *il paraît que* modifié par un pronom datif. Comme nous l'avons déjà vu (voir la section précédente), Nølke a montré que *paraître* dans cet emploi perd sa valeur d'ouï-dire et que les locutions *il (pronom datif) paraît que* et *il (pronom datif) semble que* sont interchangeable et presque synonymes dans beaucoup de cas. Il reste quand même une différence en ce qui concerne la nature de l'évidence pour *p* : si celle-ci provient de la vision, on emploie *paraître*, si elle provient d'indices plus ou moins inconscients, on emploie *sembler* (Nølke 2001 : 26, 27). Ainsi, Nølke montre que *Il lui semble qu'il connaît Marie depuis longtemps* est acceptable, tandis que l'acceptabilité de *Il lui paraît qu'il connaît Marie depuis longtemps* est jugée douteuse, étant donné que l'évidence pour *p* (*il connaît Marie depuis longtemps*) ne peut pas être de nature visuelle. Bien que le type d'évidence soit différent, la polyphonie exprimée par *il me paraît que* est du même type que celle exprimée par *il me semble que* : il s'agit d'une polyphonie interne où le locuteur textuel est responsable de  $pdv_1$ , mais où le locuteur de l'énoncé n'assume pas la vérité de ce point de vue ; il le présente comme non vérifié. En tant que marqueurs évidentiels, *il me semble que* et *il me paraît que* fonctionnent différemment, mais leur valeur modale reste la même.

Maintenant que nous avons mis en évidence le fonctionnement syntaxique de *sembler* et de *paraître*, leur contenu sémantique et leurs structures polyphoniques, passons aux occurrences de *sembler* et de *paraître* relevées dans le corpus.

#### 7.2.4 La fréquence de *sembler* et de *paraître* dans le corpus

Dans les articles de linguistique, la recherche de '*sembl.\**' produit 66 attestations, dont quatre sont à considérer comme du « bruit » (plus précisément, il s'agit de quatre occurrences de l'adjectif *semblable*). Il nous reste donc 62 occurrences, ce qui donne une fréquence relative de ce verbe de 0,9 pour mille (cf. tableau 3, section 6.4.2). Deux de ces occurrences ont été classées comme non épistémico-modales ; l'une est l'exemple 13 discuté ci-dessus (... *Ainsi les exemples suivants semblent-ils franchement bizarres...*) et l'autre est une occurrence de *sembler* employé métalinguistiquement. Chez les médecins-chercheurs, on trouve 46 attestations du verbe *sembler*, ce qui donne une fréquence relative de 0,75 pour mille. Toutes ces occurrences ont une valeur épistémico-modale.

Ces résultats montrent que l'emploi de *sembler* est légèrement plus fréquent chez les linguistes. La répartition des différentes formes du verbe est la suivante :

Linguistique : 48 *semble*, 10 *semblent*, 1 *semblerait*, 1 *semblaient*,

Médecine : 37 *semble*, 6 *semblent*, 1 *sembler*, 1 *semblé* et 1 *semblant*

On note que ce sont presque exclusivement les formes du présent qui sont employées, ce qui n'est pas étonnant, car le présent est le temps le plus fréquemment employé dans les articles de linguistique (voir Fløttum et al. 2006 : 262) ainsi que dans les parties des articles médicaux qui contiennent le plus d'atténuateurs (voir Beaufrère-Bertheux 1997)<sup>9</sup>.

Le verbe *paraître* est beaucoup moins fréquent que *sembler*. Son emploi est particulièrement rare dans les articles de médecine : on y a repéré 8 occurrences – ce qui donne une fréquence relative de 0,13 pour mille –, contre 27 occurrences dans les articles de linguistique, ce qui donne une fréquence relative de 0,39 pour mille. Alors que les huit attestations dans les articles de médecine relèvent toutes de la catégorie épistémico-modale, quatre des occurrences dans les articles de linguistiques ont été classées comme non-épistémiques. Il s'agit de trois occurrences de *paraître* dans le sens de *être publié*, et d'un cas limite, du type de celui discuté ci-dessus (avec l'exemple 12).

Il ressort de ces chiffres que les linguistes emploient le verbe *paraître* presque trois fois plus souvent que les médecins-chercheurs. Il convient cependant de noter qu'il s'agit, même chez les linguistes, d'un nombre assez restreint d'occurrences, par rapport au nombre d'occurrences de *sembler*. Cela pourrait s'expliquer par l'importance de l'aspect sensoriel ; celui-ci domine avec le verbe *paraître* alors que pour *sembler*, c'est l'aspect de jugement de l'esprit qui domine (cf. la section précédente). La prédominance de *sembler* sur *paraître* reflète le fait que les opinions et assertions formulées dans les textes scientifiques se basent sur des jugements de l'esprit plutôt que sur des impressions visuelles.

La répartition des différentes formes du verbe est la suivante :

Linguistique : 19 *paraît*, 2 *paraissent*, 2 *paraître*

Médecine : 6 *paraît*, 1 *paraïsse*, 1 *paraissent*.

---

<sup>9</sup> Les articles de médecine comportent souvent beaucoup de formes du passé aussi, mais celles-ci se trouvent pour la plupart dans les sections Méthode et Résultats, c'est-à-dire dans les sections les plus pauvres en atténuateurs.

Sur les 8 occurrences repérées dans le corpus médical, 6 sont issues d'un même article, l'article Frmed19 (voir l'appendice 2). Cela indique que l'emploi de *paraître* comme modalisateur épistémique dans les articles de médecine peut être encore moins fréquent que ces chiffres ne le suggèrent, ou du moins qu'il dépend largement des préférences stylistiques des auteurs.

### 7.2.5 Le cotexte immédiat

Passons maintenant aux études du cotexte immédiat. L'étude du cotexte de *sembler* et de *paraître* sera menée à partir de trois entrées : la répartition des constructions syntaxiques, la combinaison avec d'autres modalisateurs épistémiques et l'adjonction de pronoms datifs qui explicitent la source de l'évaluation exprimée.

Pour ce qui est de la distribution des différentes constructions syntaxiques, celle-ci ne révèle pas grand chose en elle-même. Nous avons vu que *sembler* et *paraître* s'emploient entre autres comme verbes attributifs et comme semi-auxiliaires, et plus que la construction syntaxique, ce sont les types d'attribut ou les types de verbes principaux qui peuvent intéresser l'analyse de la manière dont ces verbes sont utilisés dans le texte. Nous aborderons donc les différentes constructions syntaxiques en présentant quelques observations concernant le type d'adjectifs et le type de verbes principaux avec lesquels ces verbes sont combinés.

En ce qui concerne la combinaison de modalisateurs épistémiques (les '*clusters*', cf. *supra* 6.3), l'étude de ce phénomène est motivée par une observation faite par plusieurs chercheurs selon laquelle les modalisateurs épistémiques apparaissent souvent en combinaison, rendant ainsi l'énoncé moins sûr encore (cf. *supra*, 3.3.3). Cette observation a été faite pour le norvégien (Breivega 2003) aussi bien que pour l'anglais (voir par exemple Salager-Meyer 1994, qui parle de '*compound hedges*'). Par une étude systématique de telles combinaisons dans notre corpus, nous cherchons à examiner dans quelle mesure cette observation est valide pour le français aussi. Dans la présente étude, pour qu'une combinaison de marqueurs soit considérée comme telle, deux ou plusieurs modalisateurs épistémiques doivent apparaître dans la même période syntaxique et porter sur le même contenu informatif.

Enfin, l'étude des pronoms datifs vise à dégager dans quelle mesure la source de la qualification épistémique exprimée par *sembler* ou *paraître* est explicitement marquée. Il est inscrit dans le contenu sémantique de *sembler* et de *paraître* qu'il y a une instance énonciative responsable de l'évaluation et de la modalisation, et cette instance est par défaut le locuteur. La source de l'évaluation reste implicite dans la plupart des cas, mais elle peut être rendue explicite par des pronoms datifs. La présence ou l'absence de pronoms datifs dans de telles

constructions est un indice de la manifestation explicite de l'auteur dans son texte. Le fait d'expliciter l'instance modalisante par les pronoms *me* ou *nous* a pour effet de rendre l'énoncé plus personnel, mais peut aussi le rendre plus sûr (cf. Nølke 2001), ce qui expliquerait le changement des modes. Il semble y avoir une nette différence avec l'équivalent usuel en anglais, puisque la locution *it seems to me that* ne serait probablement pas conçue comme plus sûre que *it seems that* (cf. Johansson 2001 : 239).

#### 7.2.5.1 Répartition des constructions syntaxiques

Dans les articles de linguistique, sur les 60 occurrences de *sembler*, 21 sont des attestations de *sembler* comme semi-auxiliaire modal (exemples 14 et 15 ci-dessous), dont une entre dans une construction impersonnelle (cf. exemple 3 ci-dessus), 15 sont des attestations de *sembler* comme verbe attributif (exemples 16 et 17), et 24 font partie d'une construction impersonnelle du type *il semble que* (exemple 18).

14. Magali Rouquier (1988, 1990) a étudié le développement diachronique de cette tournure *en ce que*, qui **semble** s'être restreinte au cours de l'histoire. (frling01)
15. Si on compare le sens de *sur* et de *contre*, le trait qui **semble** opposer ces deux prépositions, c'est l'axe le long duquel se fait le contact: axe horizontal pour *contre*, axe vertical pour *sur*. (frling16)
16. Ainsi, toute une série de contraintes syntaxiques **semblent** liées à ce qu'on pourrait appeler une "allergie" à la préposition. (frling04)
17. Dire qu'une préposition introduit un groupe prépositionnel **semble** une lapalissade. (frling08)
18. Il me **semble** par exemple que chaque fois qu'on relit le Discours de la Méthode ou les Méditations, on est d'abord confronté à l'hétérogénéité des modes de présence du "je" dans le texte. (frling12)

Dans les articles de médecine, sur les 46 occurrences de *sembler*, 22 sont des attestations de *sembler* comme semi-auxiliaire modal (exemples 19 et 20 ci-dessous), dont une entre dans une construction impersonnelle. Il y a 14 occurrences de *sembler* comme verbe attributif (exemples 21 et 22), dont trois entrent dans une construction impersonnelle (voir l'exemple 21), et il y a 10 attestations de *sembler* dans des constructions impersonnelles du type *il semble que* (exemple 23).

19. Le délai habituel avant le diagnostic des SP ne **semble** pas dépasser une moyenne de 2 mois pour 80% des patients [4], ce qui est le cas dans notre série. (frmed15)

20. La vascularite leucocytoclasique **semble** être aussi fréquente [12], probablement secondaire à un afflux de polynucléaires neutrophiles ou en réponse à une atteinte vasculaire primitivement lymphocytaire. (frmed16)
21. En raison des similarités entre les trois ensembles nationaux de données, il **semble** raisonnable de regrouper les valeurs annuelles de ces sources (frmed04)
22. Plus rarement, la SC **semble** l'unique localisation, tout au moins diagnostiquée avec les investigations actuelles [18], [45], [46], [48], [49]. (frmed18)
23. Il nous **semble** que c'est la force d'entretien et d'appauvrissement de ces conduites qui donne le sentiment que des formes névrotiques évolutives sont à classer dans des formes limites. (frmed13)

En pourcentage, on obtient les chiffres suivants (chiffres arrondis) :

**Tableau 1** : Répartition des constructions syntaxiques de *sembler*

Constr.syntaxique	linguistique	médecine
Semi-auxiliaire	35 %	48 %
Verbe attributif	25 %	30 %
Constr. impersonnelle	40 %	22 %

Dans les articles de linguistique, *paraître* est employé comme semi-auxiliaire 4 fois (exemple 24) et comme verbe attributif dans 18 cas (exemples 25 et 26), dont deux dans des constructions impersonnelles. Il y a 1 occurrence de *paraître* dans une construction impersonnelle du type *il paraît que* (voir l'exemple 9 ci-dessus).

24. Par contre, une voûte **paraît** surplomber, en [21], le sujet S de perception. (frling17)
25. Par contre *il est de genre noble* **paraît** acceptable. (frling20)
26. La réponse me **paraît** positive pour le français, dans la mesure où nombre de phénomènes syntaxiques et sémantiques doivent y faire référence. (frling04)

Dans les articles de médecine, il y a 1 occurrence de *paraître* comme semi-auxiliaire modal (exemple 27) et 7 occurrences de *paraître* comme verbe attributif (exemple 28), dont 3 se trouvent dans des constructions impersonnelles (voir l'exemple 7 ci-dessus). On ne trouve aucun exemple de *paraître* dans la construction *il paraît que* :



27. Les dimensions psychologiques qui les différencient sont variables, selon le type de conduites addictives : si l'alexithymie **paraît** être une dimension commune à toutes les conduites addictives, il semble, chez les toxicomanes, que ce soit la recherche de sensations (plus que la dépendance affective ou la dépression) qui est liée au passage de la consommation régulière à la conduite addictive. (frmed13)

28. Cependant, l'énoncé même de l'hypothèse **paraît** délicat. (frmed19)

En pourcentage, on obtient les chiffres suivants (chiffres arrondis) :

**Tableau 2** : Répartition des constructions syntaxiques de *paraître*

Constr.syntaxique	linguistique	médecine
Semi-auxiliaire	17,4 %	12,5 %
Verbe attributif	78,3 %	87,5 %
Constr. impersonnelle	4,4 %	0 %

On pourrait en conclure qu'en ce qui concerne *sembler*, les trois types de constructions syntaxiques sont couramment employés. Pour *paraître*, en revanche, c'est avant tout la construction attributive qui est utilisée, par les linguistes ainsi que par les médecins-chercheurs.

Passons maintenant à la question de savoir avec quels types de verbes (emploi infinitif) et quels types d'attributs (emploi attributif) *sembler* et *paraître* sont combinés. En ce qui concerne l'emploi infinitif, il s'est vite avéré que le verbe le plus fréquent est *être* ; cela est largement dû aux occurrences de *sembler* dans les articles de médecine, dont plus de la moitié est suivie du verbe *être*. Ce verbe est souvent omissible, et dans ces cas, *sembler* (et dans une moindre mesure *paraître*<sup>10</sup>) fonctionnent quasiment comme dans leur emploi attributif. Pour les occurrences de *sembler* (et de *paraître*) comme semi-auxiliaires où le verbe principal est *être*, l'attribut qui éventuellement suit *être* a été pris en considération dans l'étude des types d'attributs.

En ce qui concerne les autres verbes principaux, il n'est pas possible de détecter des tendances spécifiques chez les deux groupes. *Ssembler* se combine, pour ne citer que quelques exemples, avec des verbes comme *présenter*, *correspondre* et *dépendre* (linguistique) et *indiquer*, *représenter*, *expliquer* (médecine), et il ne semble pas y avoir de différences sensibles entre médecins-chercheurs et linguistes en ce qui concerne les types de verbes

<sup>10</sup> Comme nous l'avons vu, l'emploi de *paraître* comme semi-auxiliaire est rare dans le corpus, et il n'y a qu'une occurrence où *paraître* est suivi du verbe *être*.

employés. La seule différence appréciable est que les linguistes utilisent une gamme plus large de verbes principaux que ne le font les médecins-chercheurs ; en raison de la prépondérance du verbe *être*, ces derniers font preuve d'une variation lexicale moins importante dans cette construction spécifique.

En ce qui concerne l'emploi attributif par contre, certaines tendances se dessinent. Dans les articles de médecine, l'adjectif le plus souvent employé avec *sembler* (que celui-là soit directement postposé à *sembler* ou s'y associe par l'intermédiaire de *être*) est l'adjectif *fréquent* avec six occurrences. Selon la classification de Kerbrat-Orrechioni (2002, voir ci-dessous), cet adjectif est *évaluatif* mais *non axiologique*, c'est-à-dire qu'il implique une évaluation quantitative ou qualitative de l'objet dénoté, sans énoncer de jugement de valeur, ni d'engagement affectif du locuteur (ibid. : 96-97) :

29. Par ailleurs, un infiltrat inflammatoire granulomateux **semble être fréquent** dans les formes à évolution chroniques et végétantes [4], [13]. (frmed16)

30. Certains articles récents mettent l'accent sur une augmentation croissante du nombre de cas de SC [11], mais aussi sur un très possible facteur racial, puisque cette localisation **semble particulièrement fréquente** chez les Japonais. (frmed18)

Deux autres concepts apparaissent à plusieurs reprises, sous forme d'attribut (emploi attributif) ou de verbe (emploi infinitif), celui d'*efficacité* et celui de *relation causale*:

31. Le budénoïde, un corticoïde à action topique plus élevée et un passage systémique plus faible, **semble être efficace** avec moins d'effets secondaires [21],[22]. (frmed06)

32. En revanche, les immunoglobulines intraveineuses **semblent inefficaces** (1 réponse sur 12 malades traités dans la série de Mariette), tout comme les plasmaphères [15],[17]. (frmed08)

33. Cette pointe saisonnière **ne semble pas être causée par** l'inclusion de patients souffrant de bronchopneumopathie chronique obstructive (dont les problèmes peuvent s'accroître à l'automne) et ayant fait l'objet d'un diagnostic erroné... (frmed02).

34. L'impulsivité chez ces patients **semble par ailleurs expliquer** en partie leur prédisposition à consommer des toxiques. (frmed13)

Dans les articles de linguistique, aucun concept spécifique ne s'avère particulièrement fréquent, mais on note une fréquence relativement élevée d'adjectifs évaluatifs axiologiques, c'est-à-dire des adjectifs qui transmettent un jugement de valeur, positif ou négatif, de la part du locuteur, qui ainsi prend position en faveur ou à l'encontre de l'objet ou du procès qualifié (cf. Kerbrat-Orechioni 2002 : 102). Dans notre corpus, les adjectifs axiologiques sont

souvent eux-mêmes modifiés par un adverbe modificateur qui renforce l'aspect axiologique (voir l'exemple 35) :

35. Cette dernière hypothèse **semble particulièrement appropriée** pour l'emploi de à-datif (frling08)
36. Sans prétendre ni bien sûr que cette option soit vaine, ni que notre effort pour l'élargir soit entièrement neuf (je pense par exemple aux travaux de C. Vandeloise qui mettent l'accent sur la dimension plus fonctionnelle des repérages spatiaux), il me **paraît essentiel** d'aller plus loin en insistant sur le fait que les emplois des prépositions spatiales sont conditionnés par des valeurs tout autres que configurationnelles (frling15)
37. Les locutions prépositionnelles *à travers* et *au travers (de)* confrontent la recherche linguistique à un défi qui nous **paraît exemplaire** à plus d'un titre. (frling17)
38. Ce phénomène nous **semble crucial** pour rendre compte de l'opération effectuée par ces items (frling09)

Du fait de leur subjectivité manifeste, ces adjectifs seraient, selon Kerbrat-Orecchioni (2002 : 103) proscrits de tout discours à vocation scientifique. Néanmoins, nous en avons trouvé plusieurs occurrences dans le corpus, mais ces occurrences relèvent principalement du corpus linguistique. Par ailleurs, il est fort possible que ces occurrences atténuées soient les seules repérables ; sans doute ces adjectifs sont-ils à ce point « subjectifs » qu'ils demandent l'adjonction d'un atténuateur afin d'éviter à l'auteur d'être trop catégorique. C'est en tout cas le rôle de *sembler* et *paraître* dans les exemples ci-dessus : ils servent d'équilibre à la qualification relativement forte transmise par le syntagme adjectival. On note aussi que dans plusieurs de ces exemples, l'adjonction d'un pronom datif à la première personne souligne l'aspect personnel de l'évaluation et signale ainsi que celle-ci ne concorde pas nécessairement avec celle du lecteur.

Il semble donc que les linguistes emploient *sembler* et *paraître* avec des syntagmes adjectivaux assez fortement marqués subjectivement, alors que les médecins-chercheurs ont tendance à employer ces verbes en relation avec des concepts plus aisément mesurables comme la fréquence et l'efficacité. Les adjectifs liés à ces notions sont aussi subjectifs et évaluatifs, mais leur subjectivité est moins apparente (cf. Kerbrat-Orecchioni 2002 : 94 sqq). Le corpus est cependant trop restreint et les tendances pas suffisamment nettes pour qu'on puisse tirer des conclusions définitives sur ce point. On trouve ainsi quelques adjectifs de type axiologique dans les articles de médecine également (cf. les exemples 21 et 28), bien qu'ils restent beaucoup plus rares ici que dans les articles de linguistique.

### 7.2.5.2 L'emploi de modalisateurs accumulés

Comme déjà mentionné, il a été observé que les modalisateurs épistémiques apparaissent souvent en combinaison entre eux dans les discours scientifique norvégien et anglais. Nous avons maintenu plus haut qu'un tel « entassement » de modalisateurs épistémiques a pour effet de rendre l'énoncé moins sûr. Si cela est vrai dans beaucoup de cas, il ne faut pas oublier non plus que la combinaison d'atténuateurs peut aussi être une question de conventions : il s'agit souvent de combinaisons de mots relativement figées, et les langues diffèrent dans leur emploi de telles combinaisons, préférant, à des degrés divers, des lexèmes isolés ou des co-occurrences. Néanmoins, ce phénomène s'est avéré rare dans les articles de recherche français, du moins en ce qui concerne les combinaisons incluant *sembler* ou *paraître*. Deux cas uniques dans le corpus attestent de cette combinaison des verbes avec un autre modalisateur épistémique ; il s'agit d'un exemple de Frling04 où *paraître* se combine avec *pouvoir* et un exemple tiré de Frmed02 où *sembler* se combine avec *pouvoir* :

39. Ainsi, la relation sémantique entre le verbe et le complément **peut paraître** identique pour autoriser et permettre, mais cela n'empêche pas l'un de se construire sans préposition et l'autre avec. (frling04)
40. Le coût annuel des hospitalisations pour l'asthme, chiffré entre 18 et 21 millions de dollars, **peut sembler** élevé, mais cette estimation est sans doute de beaucoup inférieure au coût réel (frmed02)

### 7.2.5.3 Le rôle du datif à la première personne

Comme nous l'avons vu, l'adjonction d'un pronom datif à la première personne avant *sembler* ou *paraître* peut avoir pour effet de rendre l'énoncé plus sûr (Nølke 2001 : 24). *Il me semble que* indiquerait donc un degré de certitude plus haut que *il semble que*. Le locuteur qui énonce *il me semble que* s'attache plus nettement à la vérité du contenu informatif que celui qui énonce *il semble que*, puisqu'il prend la vérité de *p* à son propre compte. L'emploi des modes souligne les degrés de certitude : *il me semble que* se construit avec l'indicatif, et *il semble que* avec le subjonctif. Il ne s'agit cependant pas exclusivement d'une question de degré de certitude, car l'emploi du datif à la première personne reflète aussi une tendance de la part de l'auteur à se manifester dans son texte. Par l'emploi d'un tel pronom, l'auteur se rend visible et révèle la personne derrière les mots écrits. Il sera donc intéressant de voir dans quelle mesure les occurrences de *sembler* et de *paraître* sont précédées d'un pronom datif à la première personne, en tant qu'indice de la manifestation explicite des auteurs dans leur texte.

Les tableaux 3 et 4 montrent la répartition des pronoms datifs précédant *sembler* dans les articles de linguistique et dans ceux de médecine<sup>11</sup> :

**Tableau 3** : Répartition des pronoms datifs précédant *sembler* dans les articles de linguistique

Construction syntaxique	Nombre total d'occurrences	Occ. précédées d'un datif →	<i>me</i>	<i>nous</i>	%
Semi-auxiliaire	21	0	0	0	0
Verbe attributif	15	5	4	1	33
Constr. impersonnelle	24	14	11	3	56
Total	60	19	15	4	32

**Tableau 4** : Répartition des pronoms datifs précédant *sembler* dans les articles de médecine

Constr. syntaxique	Nombre total d'occurrences	Occ. précédées d'un datif →	<i>me</i>	<i>nous</i>	%
Verbe attributif	14	0	0	0	0
Semi-auxiliaire	22	0	0	0	0
Constr. impersonnelle	10	4	0	4	29
Total	46	4	0	4	9

Les résultats révèlent que les linguistes tendent à utiliser un pronom datif à la première personne plus souvent que les médecins-chercheurs : sur 60 occurrences de *sembler* chez les linguistes, 15 sont précédées d'un pronom datif à la première personne au singulier (*me*), et 4 sont précédées d'un pronom datif à la première personne au pluriel (*nous*). 19 occurrences sont alors marquées par un pronom datif à la première personne, ce qui nous donne un pourcentage de presque un tiers (32 %).

Exemples :

41. Il **me semble** que cette interprétation est fondée sur des bases peu solides (frling01)

42. H. Arendt part d'une évidence qui l'a frappée et qu'elle rend, **me semble-t-il**, en quelque sorte présente pour nous. (frling12)

Dans les articles de médecine, 4 sur 46 occurrences seulement sont marquées par un pronom datif. Il s'agit à chaque occurrence du pronom datif à la première personne *au pluriel*, et à l'exception d'une seule, elles relèvent toutes du même article.

<sup>11</sup> Aucune occurrence de pronoms datifs à la deuxième ou à la troisième personne n'a été relevée.

Exemple :

43. Il nous **semble** qu'elle est essentiellement le reflet d'une structuration psychique (frmed13)

L'absence de *me* dans les articles de médecine peut s'expliquer par le fait que les médecins-chercheurs rédigent le plus souvent leurs articles en coopération ; il est rare qu'un chercheur en médecine publie un article tout seul. Cependant, le fait que seulement 9 % des occurrences de *sembler* dans les articles de médecine soient précédées d'un pronom datif à la première personne, tandis que le pourcentage est d'environ 32 % chez les linguistes, semble indiquer, à première vue, une nette différence dans les pratiques d'écriture des linguistes et des chercheurs en médecine. Cette différence contribue à l'impression que les linguistes tendent à s'exprimer d'une manière plus manifeste que les médecins-chercheurs (Fløttum et al. 2006), en ce sens que ces premiers sont plus présents dans leurs textes ; la personne derrière le texte est plus visible. Il est toutefois important de noter que sur les 15 occurrences qui sont précédées de *me* dans les articles de linguistique, 11 relèvent d'un même article, l'article Frling12. Cet article contribue donc à une augmentation considérable de la moyenne, et les différences entre linguistes et médecins-chercheurs ne sont pas si remarquables qu'il y paraît. Cette observation souligne encore une fois l'importance de l'aspect individuel : il y a, au sein de chaque sous-corpus, des variations importantes entre articles. Il n'en reste pas moins que, exception faite de l'article Frling12, l'emploi de *sembler* précédé d'un pronom datif à la première personne reste plus fréquent chez les linguistes que chez les médecins-chercheurs, mais il ne s'agit que d'une faible différence. Cette différence peut aussi s'expliquer par la fonction syntaxique de *sembler*, étant donné que l'adjonction d'un pronom datif est rare quand *sembler* s'emploie comme semi-auxiliaire modal. Or, dans les articles de médecine, *sembler* fonctionne comme semi-auxiliaire modal dans environ 48 % des cas (22 sur 46 occurrences), alors que le pourcentage correspondant chez les linguistes est de 35%. Le choix de construction syntaxique peut à son tour refléter une tendance à expliciter plus ou moins la présence de l'auteur : c'est surtout la construction impersonnelle *il semble que* qui est susceptible de se voir modifiée par un pronom datif (cf. les tableaux 3 et 4, le présent chapitre).

En ce qui concerne *paraître*, le nombre restreint d'attestations empêche de parler de tendances spécifiques. On pourrait cependant noter que dans les articles de linguistique, 4 sur 22 occurrences sont précédées d'un pronom datif à la première personne au singulier (*me*), et

4 sont précédées du pronom datif à la première personne au pluriel (*nous*). Cela veut dire que 8 sur 22 occurrences sont marquées d'un pronom datif à la première personne<sup>12</sup>. Exemples :

44. Loin de considérer, comme on le fait le plus souvent, les valeurs aspectuelles, subjectives et qualitatives comme des suppléments que la reconstruction linguistique devrait dériver dans un deuxième temps, il **nous paraît** qu'il faut les inscrire au cœur des "motifs" les plus originels attribués aux prépositions. (frling15)
45. La sémantique prépositionnelle **me paraît** d'abord déterminée par des valeurs renvoyant à la dépendance, au contrôle, à l'appropriation réciproque, à l'anticipation et à l'attente, même s'il leur arrive aussi de s'effacer au profit de critères configurationnels, voire strictement géométriques. (frling15)

Pour ce qui est des articles de médecine, il n'y a pas d'exemples où le verbe *paraître* est précédé d'un tel pronom. La différence entre les deux disciplines rejoint donc ce qui a été observé pour *sembler*.

#### 7.2.6 Les fonctions pragmatiques de *sembler* et de *paraître*

Quelles sont alors les fonctions pragmatiques de *sembler* et de *paraître*, et dans quelle mesure existe-t-il des différences entre linguistes et médecins-chercheurs en ce qui concerne ces fonctions ? Comme nous l'avons mentionné plus haut, l'analyse des fonctions pragmatiques se fonde sur une lecture attentive de chaque attestation dans son cotexte, c'est-à-dire le paragraphe dans lequel elle apparaît et souvent les deux paragraphes adjoints également. En examinant les cotextes d'emploi de *sembler* et de *paraître*, nous essayerons de décrire les fonctions pragmatiques que ces verbes remplissent. La dichotomie entre modalisateurs orientés vers le contenu (fonction contenu) et modalisateurs orientés vers le lecteur (fonction interpersonnelle) constituera un arrière-fond pour le classement des exemples, mais nous allons ici aller un peu plus loin et considérer des usages plus spécifiques.

Commençons par la fonction contenu. L'exemple type de cette catégorie est celui où le modalisateur exprime une incertitude réelle. C'est ce que fait *sembler* dans les exemples 46 et 47:

46. Il existe peu d'études épidémiologiques permettant de calculer l'incidence de la tuberculose ostéo-articulaire dans les pays développés, il **semble** exister une augmentation des cas dans ces pays à partir des années 90 [3], [8] mais ceci ne **semble** pas être le cas dans d'autres séries [9]. (frmed15)

---

<sup>12</sup> Il y a en outre une occurrence d'un pronom datif à la troisième personne (*leur*).

47. Les signes cliniques dominés par la fièvre, les douleurs rachidiennes et l'altération de l'état général ne constituent pas des critères distinctifs; en revanche, les signes neurologiques associant l'abolition d'un réflexe tendineux, avec ou sans signes de compression médullaire, **sont** surtout observés dans les ST (53,8% vs 15,3%), avec une différence statistiquement significative ( $p < 0,05$ ). Par ailleurs, l'atteinte du rachis dorsal ou dorso-lombaire **est** plus fréquente dans les ST que dans les SP (61,5% vs 38,4%), ( $p < 0,05$ ). Biologiquement, l'existence d'une hyperleucocytose **semble** plus fréquente dans les SP (38,5% vs 19,3%) mais la différence n'est pas statistiquement significative tableau III. (frmed15)

L'exemple 47 illustre bien la différence entre *être* et *sembler*. Le degré de certitude, et par conséquent le choix du verbe, est ici directement lié à la signification statistique : les différences observées sont exprimées par les verbes *être* ou *sembler* selon qu'elles se sont avérées statistiquement significatives ou non. Cet exemple illustre également la distinction que Skelton (1997) fait entre l'« evidential truth », typique de la section Résultats, et l'« interpretative truth », typique de la section Discussion, dans le discours médical. Le premier type d'assertion est lié aux résultats non ambigus, souvent exprimés à l'aide de chiffres et de fréquences, alors que le second est lié à l'interprétation des résultats et aux spéculations qu'ils autorisent.

La question de savoir s'il y a en effet incertitude réelle, ou s'il s'agit simplement de s'adapter à un style scientifique est, évidemment, souvent insoluble, puisque il est impossible d'entrer dans l'esprit de l'auteur. Cependant, dans 46 et 47, il y a des indices évidents dans le cotexte qui signalent une incertitude réelle : dans 46, l'incertitude résulte des résultats contradictoires d'études antérieures, et dans 47 elle résulte, comme nous l'avons vu, du fait que la tendance observée ne soit pas confirmée de façon univoque par les tests statistiques. L'exemple 48 illustre la difficulté de savoir s'il s'agit d'une véritable incertitude :

48. Ces contraintes apparaissent en présence de la classe (2), "finales", et **semblent** dépendre de l'opération sémantique effectuée par des V-n particuliers associés à certaines Conjs de cette classe. C'est ce que montrent, en effet, les exemples où figurent le V-n départ: (frling09)

Une lecture possible de cet exemple est que *sembler* marque une hypothèse qui juste après est présentée comme vérifiée (par l'emploi du verbe *montrer* dans *C'est ce que montrent...*), ce qui nous indique qu'il n'y a pas d'incertitude réelle, ou, plus précisément, qu'il n'y a plus d'incertitude réelle au moment de la rédaction, car une incertitude peut avoir été levée au cours du procès. L'auteur semble inclure le lecteur dans le procès de recherche, créant ainsi l'impression que le lecteur découvre les mêmes choses que l'auteur et au même moment. Il



est également possible que ce que nous avons appelé une hypothèse soit plutôt la conclusion de l'observation faite par les exemples cités après ; la structure du texte et de l'argumentation ne coïncide pas toujours avec celle du procès de recherche. Dans ce cas, il est plus envisageable qu'il s'agisse d'une véritable incertitude.

Un cas de figure où le modalisateur assume une fonction contenu est donc celui où il s'agit d'avancer une hypothèse. Peu importe si l'auteur est effectivement convaincu ou non de l'exactitude de son hypothèse, il est presque de règle de présenter cette information sous la forme d'une affirmation atténuée, étant donné qu'il s'agit d'une hypothèse qui n'a pas (encore) été confirmée. Voici quelques exemples :

49. De ce point de vue, le *en* dans la partition interne **semble** donc pouvoir être rapproché du *en* adnominal. (frling06)
50. Selon nous celle-ci ne doit pas orienter la réflexion de manière univoque vers l'hypothèse génétique d'une vulnérabilité biologique primaire commune aux TCA, aux troubles de l'humeur et de la personnalité. Il nous **semble** qu'elle est essentiellement le reflet d'une structuration psychique, toujours ouverte à l'adolescence (frmed13)

Selon Nølke (2001 : 21), la construction impersonnelle *il semble que*, illustrée dans l'exemple 50, véhicule le résultat d'un raisonnement interne : la proposition introduite par *il semble que* exprime la conclusion qu'a tirée le locuteur par la suite d'un dialogue intériorisé. Dans notre corpus, l'emploi d'adverbes comme *donc* et *toutefois* renforce l'impression qu'il s'agit d'une conclusion tirée sur la base d'un raisonnement interne (voir les exemples 51-52). Dans la plupart de ces cas, l'atténuateur s'oriente plus vers le contenu que vers le lecteur, étant donné qu'il s'agit de transmettre un contenu informatif pour lequel on a des indices mais non des preuves :

51. On peut dire de même: il y a un volant dans une voiture, car l'opération thétiq ue du *il y a*, et l'emploi en promotion générique, permettent de dissocier le volant avant de le réintroduire). Il **semble** donc que la relation d'inclusion portée par *dans* soit entièrement un effet de cette dynamique de mise en phase de la relation de dépendance. (frling15)
52. Ils sont en moins grande quantité au niveau des autres tissus. Ils proviennent de l'alimentation et de la synthèse endogène. Dans l'adrénoleucodystrophie, on ne peut déterminer précisément laquelle des 2 voies prédomine. Il **semble** toutefois que ce soit la synthèse endogène qui prime selon les essais thérapeutiques (voir: § "Traitement"). (frmed11)

Nølke (2001 : 21) maintient que le locuteur « n'est pas prêt à fournir ses raisons pour dire *p*<sup>13</sup> [...] », qu' « il les présente comme non repérables » et que ces raisons sont plus ou moins inconscientes. Il est pourtant raisonnable de supposer que cette tendance dépend dans une certaine mesure du genre. Dans le discours scientifique, il faut que l'auteur soit prêt à justifier et fonder ses points de vue. Il n'y a donc rien de surprenant à ce que les chercheurs, dans notre corpus, présentent parfois les raisons qui permettent de tirer la conclusion introduite par *il semble que*. Dans l'exemple ci-dessous, ce sont les résultats d'une étude scientifique qui ont amené l'auteur à tirer la conclusion en question :

53. En revanche, **il semble que** l'administration de GTO-GTE avant l'apparition des premiers symptômes de la maladie ralentisse sa progression; en effet, les premiers résultats de l'étude menée par Moser sur 52 malades asymptomatiques suivis pendant 33 mois, montrent que 10 % d'entre eux développent une atteinte neurologique sévère, comparés aux 30 % prévus approximativement sur les données de l'histoire naturelle de l'adrénoleucodystrophie [5], [10], [36], [39], [40]. (frmed11)

Il est vrai cependant que les raisons qu'a l'auteur de tirer sa conclusion restent non-spécifiées dans nombre de cas. Cela est très explicite dans 54, où la conclusion est présentée comme une intuition. Cette intuition est néanmoins, à son tour, qualifiée comme *confirmée* :

54. Il nous **semble** néanmoins que la seule interprétation possible pour [51] est qu'un tiers de chaque manuscrit a été publiée et que l'on a donc affaire à un cas de partition interne. **Cette intuition** sémantique se voit confirmée par deux faits: les phénomènes d'accord et le comportement de la plupart. (frling06)

Il y a là une différence avec les autres emplois impersonnels ; lorsque les auteurs emploient *sembler / paraître* dans d'autres constructions impersonnelles, ils tendent en même temps à expliciter les raisons au fondement du point de vue exprimé, que ce soit avant (55) ou après (56) la proposition elle-même.

55. Les quantificateurs de la série *une dizaine* sont les seuls à permettre sans réserves l'utilisation de *dont*, mais comme ils partagent les cinq autres propriétés de *en* quantitatif et que le résultat de ce critère varie de toute façon selon le quantificateur utilisé, il **paraît** légitime d'admettre que le *en* dans l'exemple [9] est également *en* quantitatif. (frling06)

56. Les études internationales rendent compte d'une surconsommation de psychotropes chez les patients boulimiques (hétéroprescription et autoprescription); mais il **semble** difficile de différencier les consommateurs réguliers des abuseurs (problèmes méthodologiques, de définition notamment). (frmed13)

---

<sup>13</sup> C'est-à-dire la proposition.

Dans l'exemple 55, les raisons données au point de vue *il est légitime d'admettre que le en dans l'exemple [9] est également en quantitatif* sont exprimées dans la subordonnée introduite par *comme*. Dans cet exemple, l'atténuation assume une fonction interpersonnelle plutôt qu'une fonction contenu, car l'auteur cherche à justifier son point de vue auprès des lecteurs. Dans 56, les raisons données au point de vue *il est difficile de différencier les consommateurs réguliers des abuseurs* sont précisées après coup entre parenthèses.

Comme nous l'avons vu plus haut, le locuteur textuel prend en charge, dans l'interprétation par défaut, le point de vue exprimé par la proposition introduite par *il semble que*. Les cas inverses sont très rares (Nølke 2001). On note ainsi une différence entre la construction impersonnelle étudiée par Nølke et les autres emplois de *sembler* (et de *paraître*). Dans ces emplois, il arrive souvent que le point de vue exprimé par la proposition modalisée soit réfuté, totalement ou en partie (exemples 57-58). Dans ces cas, le modalisateur apparaît souvent dans des cotextes concessifs (exemple 58) :

57. Ce tableau général, qui donne une importance capitale aux programmes de Chomsky, de Montague, et d'un philosophe du langage comme Searle, serait incomplet si une autre contribution, d'un philosophe également, n'avait pas eu pour conséquence de mélanger complètement les cartes et de rendre compatibles des positions qui **semblaient** initialement incompatibles. (frling11)

58. Chez 50 % des malades environ, l'atteinte reste localisée au niveau des nerfs périphériques et de la moelle. C'est pourquoi le principal diagnostic différentiel de cette affection est la sclérose en plaques. Les fonctions intellectuelles **semblent** conservées, mais des tests psychologiques précis permettent de détecter dans environ la moitié des cas des troubles cognitifs qui peuvent ensuite progresser. (frmed11)

Dans 57 et 58, le locuteur ne prend pas en charge les points de vue exprimés par les propositions sous-jacentes *ces positions sont incompatibles* dans 57 et *les fonctions intellectuelles sont conservées* dans 58. *Ssembler* est ici employé pour transmettre une impression qui est alors qualifiée de fautive. Un contraste est ainsi établi entre l'apparence et la réalité (cf. Kibbee 1995).

Par ailleurs, la provenance du jugement introduit par *sembler / paraître* est presque toujours l'auteur. Mais il existe quelques rares cas de modalisation rapportée (voir *supra*, section 6.2.3) :

59. L'étude remarquablement claire, documentée, et argumentée de D. Leeman s'appuie fondamentalement sur Gross (1986). Celui-ci remarque que la construction *en tout(e) N* d'une part ne **semble** pas fonctionner de façon simple, et d'autre part dissimule plusieurs structures, en particulier adverbiales, celles qui m'intéresseront ici. (frling13)

Dans 59, *sembler* fait partie d'un discours indirect, et – pourvu que la représentation soit fidèle au texte rapporté – ce n'est pas le locuteur qui est responsable de la modalisation, mais le *locuteur représenté* (Nølke et al. 2004 : 57), c'est-à-dire celui qui a produit l'énoncé rapporté. Le locuteur peut à son tour exprimer son accord ou son désaccord avec le point de vue du locuteur représenté. Dans le cas présent, le locuteur n'est pas très explicite en ce qui concerne son propre point de vue, mais il semble accepter le point de vue exprimé par le locuteur représenté.

Dans sa fonction interpersonnelle, l'atténuation a été mise en relation avec la notion de politesse. Dans nos données, l'aspect de politesse se manifeste le plus clairement quand *sembler* et *paraître* sont utilisés pour atténuer la critique envers d'autres chercheurs. L'auteur peut se servir de *sembler* et *paraître* afin de critiquer avec prudence les théories ou les résultats d'autres chercheurs ou simplement pour marquer son désaccord :

60. Le LEX la répertorie sous une division d'expressions et le GL la note dans une sous-entrée. La préposition à *propos de*, en revanche, est consignée sous une division d'expressions, sauf dans le TLF. Pour justifier une telle répartition, le critère de l'autonomie de l'unité lexicale ne nous **paraît** pas assez robuste. L'une des expressions serait-elle plus abstraite que l'autre et donc moins compositionnelle? (frling10)

61. Il a été noté dans ce travail 57 % d'hypothyroïdies sub-cliniques (TSH de base comprises entre 5,4 et 7,6 uU/mL). Il nous a **semblé** que la fréquence de l'hypothyroïdie définitive dans cette affection est sous-estimée. (frmed12)

Parfois, les modalisateurs ne servent pas vraiment à atténuer la critique, mais plutôt à masquer celle-ci sous une prudence apparente, tandis qu'elle reste en réalité assez acerbe. Ainsi, dans l'exemple 62, les adjectifs et la tournure introductive semblent produire le contraire de l'effet attendu du modalisateur :

62. Toute volonté d'offense mise à part, cette analyse me **paraît** aussi tortueuse que peu convaincante. (frling13)

La critique de théories et de points de vue existants correspond à une occasion pour l'auteur de se positionner, et dans ces cas, les aspects de politesse et de persuasion sont co-présents :

63. Comme il arrive très souvent dans ce type de présentation, tout se passe comme si un "bon complot" avait guidé l'évolution historique pour l'amener vers la seule bonne forme normative. Il me **semble** que cette interprétation est fondée sur des bases peu solides et

qu'on doit accepter l'idée que les trois tournures coexistent actuellement, et qu'elles ont sans doute coexisté à d'autres époques. (frling01)

64. Au moins chez les linguistes, est souvent répété un discours qui me **semble** justement discutabile, celui selon lequel la langue conditionne la forme de la pensée possible sur le horslangage. (frling12)

Ces emplois, où *sembler* et *paraître* servent à atténuer la critique, sont à quelques exceptions près absents dans les articles de médecine. Il ne faut pas en tirer la conclusion que les médecins-chercheurs critiquent d'une manière plus brusque, ni qu'ils ne critiquent pas du tout ; cela pourrait simplement indiquer que les médecins-chercheurs ont d'autres traditions pour présenter leur critique, par exemple en signalant des faiblesses méthodologiques dans des études antérieures.

On pourrait aussi maintenir que *sembler* et *paraître* fonctionnent comme des marqueurs de politesse lorsqu'ils sont utilisés pour atténuer un énoncé paraissant trop abrupte, bien qu'il soit peut-être plus approprié de décrire cet usage en faisant référence à la notion de modestie plutôt qu'à celle de politesse. L'idée est que, par l'adjonction d'un modalisateur dans de tels énoncés, l'auteur évite d'imposer au lecteur une opinion spécifique et le modalisateur sert à contrebalancer un énoncé qui pourrait être conçu comme trop catégorique. Il s'agit surtout d'énoncés qui transmettent un haut degré de certitude, par exemple à l'aide d'un amplificateur comme l'adjectif *évident* :

65. En conclusion, alors qu'il **semble** évident que les AGTLC jouent un rôle dans la pathogénie de l'adrénoleucodystrophie, que le gène de l'adrénoleucodystrophie a été récemment identifié, et que l'étude de la protéine de membrane du peroxyosome pour laquelle il code permet de grands progrès dans la compréhension des mécanismes physiopathologiques, l'accumulation des AGTLC ne peut à lui seul tout expliquer. (frmed11)

D'autres fois, les stratégies d'atténuation produisent plutôt l'effet contraire et mènent à un ton presque faussement modeste, car l'auteur apparaît assez sûr de lui :

66. Il n'en est pourtant pas toujours ainsi. En effet, nous **croyons pouvoir montrer** que [50-51] sont tous deux des exemples de partition interne. Pour [50], cela **paraît** assez évident, puisque *en* réfère à un singulier. (frling06)

Il s'agit ici non seulement d'apparaître modeste et poli, mais de prendre position, de promouvoir sa propre recherche et de convaincre le lecteur.

Un autre usage où *sembler* et *paraître* assument une fonction interpersonnelle est celui où le locuteur les emploie pour parler d'impressions ou d'opinions possibles, pour ainsi devancer d'éventuelles objections :

67. Une autre cause possible est l'augmentation du temps passé à l'intérieur, ce qui augmente l'exposition aux allergènes tels que les acariens et les animaux domestiques. Cette pointe saisonnière ne **semble** pas être causée par l'inclusion de patients souffrant de bronchopneumopathie chronique obstructive (dont les problèmes peuvent s'accroître à l'automne) et ayant fait l'objet d'un diagnostic erroné d'asthme. (frmed02)

Le locuteur anticipe ici les réactions et les remarques potentielles du lecteur, et y répond par avance. Mais on pourrait aussi arguer qu'on a affaire à une fonction contenu : l'auteur n'a pas assez d'éléments pour étayer son affirmation (*Cette pointe saisonnière n'est pas causée par l'inclusion de patients souffrant de bronchopneumopathie chronique obstructive...*) à laquelle il donne donc un statut hypothétique.

Finalement, nous commenterons brièvement le rôle des incisives. Celles-ci semblent souligner la subjectivité de ce qui est dit, et parfois, il n'y a pas de grande différence entre ces constructions et la locution *selon moi/nous*. Dans ces cas, le modalisateur est orienté vers le lecteur plutôt que vers le contenu informatif lui-même. L'auteur l'ajoute, commentant son propos pour signaler au lecteur qu'il s'agit d'une opinion personnelle.

68. L'idée de maximalisation de la proximité peut être appliquée aussi, **il me semble**, à certains emplois non spatiaux de *contre*, comme l'emploi "opposition", au sens restreint de "lutte": (frling16)

Les incisives apparaissent rarement dans les articles de médecine ; on n'y trouve qu'une occurrence, contre sept dans les articles de linguistique.

En général, quand l'auteur emploie *sembler* et *paraître*, il signale qu'il est prêt à considérer d'autres idées. Il invite à la discussion plutôt que de feindre qu'il possède la vérité lui-même. Il engage ainsi un dialogue envers les lecteurs, ses collègues.

On retrouve dans une large mesure les mêmes fonctions pragmatiques dans les articles de linguistique et les articles de médecine. Les contextes d'emploi qui illustrent la fonction contenu, comme la présentation d'hypothèses et l'expression d'incertitudes réelles, apparaissent dans les deux disciplines. Les fonctions interpersonnelles sont légèrement plus fréquentes dans les articles de linguistique ; cela est notamment dû au fait que les linguistes emploient de temps en temps *sembler* ou *paraître* pour critiquer avec politesse (ou avec une

politesse apparente) leurs pairs, alors que cet emploi est presque absent chez les médecins-chercheurs<sup>14</sup>. À part cette différence, les principales fonctions pragmatiques de ces verbes sont, chez les médecins-chercheurs comme les linguistes, de contribuer à un ton modeste et poli, de signaler par leur subjectivité inhérente que le locuteur reste ouvert à d'autres points de vue, d'exprimer des hypothèses et des incertitudes réelles, et d'introduire une conclusion tirée sur la base d'un raisonnement interne, lequel peut être décrit ou non dans le texte. Ces fonctions ne sont évidemment pas mutuellement exclusives, elles se chevauchent et s'entremêlent, et les occurrences de *sembler* et *paraître* sont souvent *polypragmatiques*, c'est-à-dire qu'elles peuvent assumer plusieurs fonctions pragmatiques en même temps.

### 7.2.7 Bilan

En résumé, nous avons vu que les linguistes emploient les verbes *sembler* et *paraître* plus fréquemment que les médecins-chercheurs, et que cela vaut surtout pour *paraître*. En ce qui concerne les constructions syntaxiques, *paraître* s'emploie principalement comme verbe attributif, tandis que *sembler* s'utilise dans les trois constructions évoquées. Les médecins-chercheurs préfèrent employer *sembler* comme semi-auxiliaire modal, tandis que les linguistes préfèrent la construction impersonnelle *il semble que*. Les linguistes utilisent souvent *sembler* ou *paraître* pour modaliser des jugements qui incluent un adjectif axiologique, alors que les médecins-chercheurs s'en tiennent à des adjectifs moins explicitement subjectifs. Les linguistes tendent aussi à combiner ces verbes avec un pronom datif à la première personne plus souvent que ne le font les médecins-chercheurs, ce qui porte à croire, à première vue, que les linguistes tendent à s'exprimer d'une façon plus manifeste et que leur présence dans le texte est plus visible. Cependant, la différence dans l'emploi de pronoms datifs, largement due à un seul article, n'est peut-être pas aussi considérable qu'il y paraît. Pour ce qui est des fonctions pragmatiques, celles-ci sont clairement déterminées par le cotexte dans lequel *sembler* ou *paraître* apparaissent : nous avons vu, par exemple, que les adjectifs qui s'associent à *sembler* ou *paraître* influencent l'effet pragmatique de ces verbes. Les exemples ont montré que *sembler* et *paraître* peuvent servir une série de visées : exprimer une incertitude réelle, créer un ton modeste, marquer qu'il s'agit d'une opinion personnelle, introduire une conclusion tirée sur la base d'un raisonnement interne, et atténuer par politesse la critique contre d'autres chercheurs. Courante parmi les linguistes, cette dernière fonction

---

<sup>14</sup> Il convient de noter que les médecins-chercheurs ont une manière différente d'inclure les travaux d'autres chercheurs dans leurs textes. Par exemple, les références bibliographiques sont toujours mises en notes en fin de l'article, et les noms des collègues ne sont pas mentionnés dans le texte même. La critique peut par conséquent prendre d'autres formes dans les articles de médecine.

est très rare dans les articles de médecine. Il n'y a cependant pas d'autre différence nette entre linguistes et médecins-chercheurs en ce qui concerne les fonctions pragmatiques de *sembler* et *paraître*.



### 7.3 Synes et se ut (til/som)

Le verbe *synes* et le verbe à particule *se ut (til/som)* peuvent être considérés comme les équivalents norvégiens des verbes français *sembler* et *paraître*<sup>15</sup>. Par ailleurs, *synes* est le deuxième modalisateur le plus utilisé dans les articles de médecine, et *se ut (til/som)* est le deuxième dans les articles de linguistique (cf. tableau 4, ch. 6). Dans la présente section, nous ferons une étude approfondie de l'emploi de ces marqueurs, similaire à celle effectuée pour *sembler* et *paraître* dans la section précédente. Compte tenu du haut degré de correspondance entre *sembler* et *paraître* d'un côté et *synes* et *se ut (til/som)* de l'autre, nous serons amenée, dans les diverses sous-sections, à formuler des remarques contrastives lorsque cela semble pertinent.

#### 7.3.1 Fonctionnement syntaxique et sémantisme

Notre propos ici n'est pas de donner une analyse exhaustive du fonctionnement syntaxique de *synes* et *se ut (til/som)*, mais plutôt de faire ressortir les parallélismes entre *sembler* et *paraître* en français et *synes* et *se ut (til/som)* en norvégien. En fait, le fonctionnement syntaxique des verbes *synes* et *se ut (til/som)* est assez semblable à celui de *sembler* et de *paraître*. Il est vrai que la syntaxe des marqueurs norvégiens diffère légèrement de celle de leurs correspondants français, et cela est vrai notamment pour *se ut (til/som)*, qui est un verbe à particule, et qui est, en outre, très souvent suivi de la préposition *til* ou de la conjonction de subordination *som*, ce qui rend l'analyse syntaxique plus délicate<sup>16</sup>. Cependant, les ressemblances semblent plus importantes que les différences, et nous prendrons donc comme point de départ les trois constructions syntaxiques identifiées pour *sembler* et *paraître*. Outre la comparaison syntaxique entre les marqueurs français et les marqueurs norvégiens, la présentation vise à identifier les différents types de constructions attestés dans le corpus et leurs éventuels rapports avec des sémantismes particuliers.

#### 1) « Semi-auxiliaires » modaux

*Synes* ou *se ut (til/som)* sont souvent suivis d'un infinitif. *Se ut (til/som)* est dans ces cas toujours suivi de la préposition *til* (cf. l'exemple 3 ci-dessous). L'emploi infinitif est le plus

---

<sup>15</sup> Certes, il y a d'autres traductions possibles de ces verbes français (par exemple les verbes *virke* et *fortone*), mais parmi les diverses alternatives existantes, *synes* et *se ut (til/som)* semblent être les équivalents les plus immédiats.

<sup>16</sup> Il en va de même pour l'analyse automatique : en cherchant par ordinateur les occurrences de *se ut (til/som)*, il a fallu tenir compte du fait que plusieurs éléments peuvent se situer entre le verbe propre (*se*) et les autres éléments constitutifs de ce marqueur (*ut, til, som*), cf. *supra* 6.4.1.2.

souvent personnel (exemples 1 et 3), mais un emploi impersonnel est également possible (exemple 2<sup>17</sup>) :

1. Hans terminologi **synest** difor å vitna om epistemologisk ambivalens (noling13)  
'Sa terminologie **semble** donc témoigner d'une ambivalence épistémologique'
2. I Chomskys cartesianske lingvistikk og underforstått i dei fleste variantar av moderne psykologi - også sosialpsykologi - **synest** det liggja ei førestelling om relasjonen mellom det individuelle intellekt og det kulturelle kollektivet som individet er sosialisert inn i, som strengt disjunktiv. (noling13)  
'Dans la linguistique cartésienne et, de manière sous-jacente, dans la plupart des versions de la psychologie moderne – y compris la psychologie sociale – il **semble** y avoir l'idée que la relation entre l'esprit individuel et la collectivité culturelle dans laquelle l'individu est socialisé est strictement disjonctive.
3. Vaksinen **ser ut til** å redusere antall tilfeller med otitis media med 8 % (17). (nomed09)  
'Le vaccin **semble** réduire le nombre de cas d'otite moyenne de 8 % (17)'.

Contrairement à leurs correspondants français (*sembler* et *paraître*), *synes* et *se ut til* requièrent en principe l'indice de l'infinitif avant le verbe qu'ils modifient. Cela rend difficile d'appeler *synes* et *se ut til* des « auxiliaires modaux », étant donné que les auxiliaires modaux en norvégien se caractérisent entre autres par leur capacité de se combiner avec l'infinitif pur d'un verbe principal (voir Faarlund et al. 2002 : 526). D'un point de vue sémantique par contre, *synes* ou *se ut til* + verbe lexical à l'infinitif est l'équivalent exact de *sembler* ou *paraître* + verbe lexical à l'infinitif : le rôle de *sembler*, de *paraître*, de *synes* ou de *se ut til* est d'affecter le procès exprimé par le verbe d'une modalité épistémique, indiquant ainsi qu'il s'agit d'une impression ou d'une affirmation non vérifiée (voir *supra* 7.2.1). De plus, en ce qui concerne *synes* (*synest* en *nynorsk*), la présence de l'indice de l'infinitif n'est pas obligatoire, comme l'illustre l'exemple 2 ci-dessus.

## 2) Verbes attributifs

*Synes* et *se ut* peuvent se construire avec un attribut du sujet comme complément. Dans cet emploi, *se ut* apparaît soit seul soit avec la conjonction de subordination *som*, mais non avec la préposition *til*. Si l'attribut est un syntagme adjectival, le verbe apparaît sous la forme *se ut* (voir l'exemple 5), et si l'attribut est un syntagme nominal, il apparaît sous la forme *se ut som* (voir l'exemple 6) (Faarlund et al. 2002 : 747). Dans cet usage, *synes* et *se ut* (*som*) se traduisent en français tantôt par *avoir l'air*, tantôt par *sembler / paraître*.

---

<sup>17</sup> En raison de l'inversion causée par le circonstanciel antéposé, le sujet formel *det* suit ici le verbe.

4. Men derved griper han tilbake til en hypotetisk førnordisk form, og det må **synes** meget betenkelig, da det ikke finnes spor av noen slik bruk av \*aina i andre germanske språk.(noling08)  
'Mais par cela il recourt à une forme pré-nordique hypothétique, et cela doit **sembler** très discutabile, étant donné qu'il n'y a aucune trace d'un tel emploi de \*aina dans d'autres langues germaniques.
5. Bildet kan for eksempel vise en mann som **ser** rasende **ut**, og en kvinne med et nøytralt ansiktsuttrykk. (noling11)  
'L'image peut par exemple représenter un homme qui **a l'air** furieux, et une femme avec une expression neutre.'
6. Ved første øyekast kan dette **se ut** som et moteksempel til subjektvegringsanalysen. (noling01)  
'A première vue, cela peut **paraître** un contre-exemple à l'analyse d'orientation anti-sujet.'

Employés comme verbes attributifs, *synes* et *se ut* sont des variantes du verbe copule *være* ('être'), qui ajoutent à la simple idée de relation exprimée par *være* une certaine réticence et une atténuation. Si on remplace *synes* ou *se ut* (*som*) par *være*, on aura une version non-atténuée de l'énoncé. Le rapport entre *synes* / *se ut* et *være* comme verbes attributifs est donc le même que celui entre *sembler* / *paraître* et *être*. Dans l'exemple 5, le verbe *se ut* semble pourtant renvoyer à une simple impression visuelle, et la version correspondante avec *être* ne paraît pas plus catégorique ou plus sûre. Cet exemple est du même type que les exemples 12 et 13 discutés dans la section 7.2 à propos de *sembler* / *paraître* : l'aspect d'apparence l'emporte tellement sur l'aspect de jugement que la valeur épistémico-modale apparaît très affaiblie. Cet exemple, qui est d'ailleurs l'unique exemple de ce type dans notre corpus, n'a donc pas été classé parmi les occurrences épistémico-modales.

Il convient par ailleurs de noter que les occurrences de *se ut* comme verbe attributif se scindent en deux groupes. Outre les occurrences où l'attribut est un adjectif ou un syntagme nominal, on relève certaines occurrences non-épistémiques où l'attribut est un adverbe (par exemple *slik* ('ainsi') ou *korleis* ('comment')). Dans ce type d'emploi, l'idée de réticence et d'atténuation par rapport à *être* est absente. Ces occurrences sont au nombre de sept, et elles se trouvent toutes dans les textes linguistiques :

7. Det felles norsk-engelske vanskelighetshierarkiet **ser** slik **ut** [...]: (noling11)  
'La hiérarchie de difficulté, qui est commune au norvégien et à l'anglais, est composée comme suit [...]:'

8. Vel den leieforma som gjer det lettast å forutseie korleis resten av paradigmet **ser ut** (noling 04)  
'Choisir la forme de base qui rend le plus facile la prédiction du reste de la conjugaison'

Etant donné leur caractère non-épistémique, ces occurrences ne sont pas pertinentes pour la présente étude<sup>18</sup>.

### 3) *Constructions impersonnelles*

Troisièmement, *synes* et *se ut (til/som)* peuvent s'employer dans des constructions impersonnelles où ils sont précédés du pronom neutre *det* ('il') comme sujet formel et suivis d'une proposition subordonnée. *Synes* est typiquement suivi d'une subordonnée introduite par *som* (voir l'exemple 9), alors que le verbe à particule *se ut* est typiquement suivi d'une subordonnée introduite par *at*, et se construit dans ces cas avec la préposition *til* (voir l'exemple 10). *Se ut* peut aussi être suivi d'une proposition introduite par *som* (voir l'exemple 11), mais ces cas sont moins fréquents dans le corpus. Dans tous les cas, la traduction en français sera *il (pronom datif) semble que* ou *il + pronom datif + paraît que*. Dans ces constructions impersonnelles, *synes* et *se ut (til/som)* expriment toujours une valeur épistémique<sup>19</sup>.

9. Det **synes** likevel som om intensivering i disse tilfellene berører andre egenskaper enn selve fargen (noling09)  
'Il **semble** pourtant que l'intensification dans ces cas touche d'autres qualités que la couleur elle-même'
10. Beregning av årlig insidens er usikker, da man ikke har sikre opplysninger om hva som er legegrippens pasientgrunnlag, men det **ser ut til** at insidensen er omtrent som anført av andre (nomed03)  
'Le calcul du taux annuel est incertain, puisqu'on n'a pas de renseignements sûrs concernant la population de malades de ce groupe de médecins, mais il **semble** que le taux soit à peu près équivalent à celui indiqué par d'autres.'
11. Det kan **se ut som** nominativ har vært ledeform i "gammal-vestlandsk", men ikke i "gammal-østlandsk". (noling05)  
'Il peut **sembler** que le nominatif ait été la forme de base dans l'ancien dialecte de l'ouest, mais non dans l'ancien dialecte de l'est.'

---

<sup>18</sup> Il existe aussi un emploi absolu de *se ut* comme verbe attributif, comme dans l'énoncé *hun ser ikke ut* ('elle a une mine épouvantable', littéralement 'elle n'a pas l'air'). Il n'y a aucun exemple d'un tel emploi dans le corpus, qui n'a d'ailleurs aucune valeur épistémico-modale.

<sup>19</sup> Comme c'était le cas pour *sembler* et *paraître*, ces verbes admettent un sujet impersonnel dans d'autres emplois aussi, cf. l'exemple 2 ci-dessus. La catégorie *construction impersonnelle* désigne les constructions du type *det + synes / se ut (til/som) + proposition subordonnée*.

Outre les trois constructions syntaxiques mentionnées ci-dessus, *synes* peut faire partie d'une quatrième construction, celle qui associe sujet personnel + verbe transitif + complément d'objet direct. Le complément d'objet direct apparaît sous la forme d'une complétive introduite par *at*. La conjonction de subordination est d'ailleurs omissible (voir l'exemple 12 : *vi synes (at) det er overraskende...* ) :

12. Vi **synes** det er overraskende at bare to av våre pasienter brukte syrehemmende medikamenter. (nomed20)

'Nous trouvons étonnant qu'uniquement deux de nos malades aient pris des anti-acides.'

Cette construction est restreinte à *synes* dans son sens de verbe d'opinion, c'est-à-dire un emploi non-épistémique, donc non pertinent pour la présente étude. Il y a huit occurrences d'un tel emploi dans le corpus, quatre dans chaque discipline<sup>20</sup>.

Comme nous l'avons vu, les différents sens de *synes* et de *se ut (til/som)* semblent se lier à des constructions syntaxiques précises. Employés comme « semi-auxiliaires » et dans les constructions impersonnelles, *synes* et *se ut (til/som)* vont de pair avec une lecture épistémico-modale. Il en va de même pour l'emploi de *synes* comme verbe attributif, alors que *se ut (som)* peut prendre un sens non-épistémique dans les constructions attributives, particulièrement dans les cas où l'attribut du sujet est un adverbe (cf. les exemples 7-8). En revanche, *synes* employé comme verbe transitif direct et suivi d'une subordonnée complétive va de pair avec la valeur de « verbe d'opinion ». Dans ce sens, *synes* équivaut aux verbes français *penser / trouver* et n'a aucune valeur épistémico-modale.

En ce qui concerne *se ut*, il convient de noter que la locution *se ut til* doit être considérée comme intrinsèquement épistémico-modale<sup>21</sup>, tandis que *se ut* est plus ambigu. Comme nous l'avons vu, cette dernière locution peut prendre des sens autres que le sens épistémico-modal (cf. exemples 5, 7 et 8 et note 18 ci-dessus).

En résumé, c'est nettement la valeur épistémico-modale qui est la valeur principale de *synes* et de *se ut (til/som)* dans les articles de recherche. Sur 61 occurrences de *synes*, 8 seulement sont des occurrences non-épistémiques (il s'agit de *synes* employé comme verbe

---

<sup>20</sup> Notons enfin que *synes* peut être employé dans le sens d'*être visible*, comme dans *undertøyet ditt synes igjennom stoffet* ('ta lingerie est visible à travers le tissu'). Cet emploi, qui n'a d'ailleurs rien de modal ni d'épistémique, n'est pas attesté dans le corpus.

<sup>21</sup> Il ne faut pourtant pas confondre la locution verbale figée avec la séquence *se* (verbe : *voir*) + *ut* (adverbe : *dehors*) + *til* (préposition : *à*), qui, naturellement, a des emplois non-épistémiques : *Fra vinduet sitt kan han se ut til fyret* (*De sa fenêtré, il peut voir le phare*).

d'opinion) et sur 64 occurrences de *se ut*, 9 ont été classées comme non-épistémiques (il s'agit presque sans exception de cas particuliers de *se ut* comme verbe attributif).

### 7.3.2 La fréquence de *synes* et de *se ut (til/som)* dans le corpus

Le lancement de la recherche '*synes.\**'<sup>22</sup> (forme de l'infinitif et du présent) produit 36 attestations dans les articles de linguistique, et la recherche '*syntes.\**' (forme du passé) donne 1 attestation. Quatre de ces attestations (dont l'occurrence du passé) sont des exemples de *synes* comme verbe d'opinion et il reste donc 33 occurrences épistémico-modales à étudier. La fréquence relative du verbe *synes* (toutes valeurs incluses) est de 0,4 dans les articles de linguistique, et les occurrences épistémico-modales de ce verbe constituent une fréquence relative de 0,36 pour mille du nombre total de mots (cf. le tableau 4, section 6.4.2). La répartition des différentes formes du verbe pour les occurrences épistémico-modales est la suivante : *synes* (présent) : 30, *synes* (infinitif) : 3.

Dans les articles de médecine, le lancement des recherches '*synes.\**' et '*syntes.\**' donnent respectivement 22 et 2 attestations. Sur ces 24 occurrences, quatre sont des attestations de *synes* comme verbe d'opinion. Les 20 occurrences épistémico-modales constituent une fréquence relative de 0,46 pour mille, alors que le verbe en général (toutes valeurs incluses) a une fréquence relative de 0,56 (cf. tableau 4, section 6.4.2). La répartition des différentes formes du verbe est la suivante : *synes* (présent) : 19, *syntes* (passé) : 1.

Il ressort de ces résultats (cf. aussi le tableau 4 dans le chapitre 6) que le modalisateur épistémique *synes* s'utilise fréquemment, par les médecins-chercheurs aussi bien que par les linguistes. Sa fréquence relative est un peu plus élevée chez les médecins-chercheurs, mais la différence n'est pas grande. On note aussi que c'est surtout le présent qui est utilisé, les formes du passé étant très rares. Comme mentionné dans 7.2.4, cette tendance peut s'expliquer par le fait que dans un article médical, les sections les plus riches en atténuateurs tendent à être écrites au présent, alors que ce sont les sections les plus pauvres en atténuateurs (notamment la section Méthode) qui tendent à être écrites au passé. Quant aux linguistes, ils tendent à s'en tenir au présent tout au long de l'article.

Pour le verbe à particule *se ut (til/som)*, nous avons repéré 57 attestations dans les articles de linguistique, soit une fréquence relative de 0,63 pour mille. Parmi ces attestations, 48 occurrences, soit 84 %, ont une valeur épistémico-modale. La fréquence relative de *se ut (til/som)* comme modalisateur épistémique est donc de 0,53 dans le corpus linguistique (cf.

---

<sup>22</sup> Le point et l'astérisque permettent d'inclure les occurrences de la forme en *nynorsk* : *synest*.

tableau 4, 6.4.2). Ce verbe à particule est beaucoup moins fréquent dans les articles de médecine où l'on n'a repéré que 7 occurrences, toutes exprimant une valeur épistémico-modale (fréquence relative de 0,16, cf. tableau 4, 6.4.2). Dans notre corpus, *se ut (til/som)* s'utilise essentiellement au présent : parmi les occurrences épistémico-modales, il y a 7 attestations de l'infinitif dans les articles de linguistique et 1 occurrence du passé dans les articles de médecine ; le reste est au présent.

Pour les deux verbes *synes* et *se ut (til/som)*, les formes de l'infinitif sont toutes des exemples où les verbes sont modifiés par un auxiliaire modal. Dans la grande majorité des cas, il s'agit du verbe *kunne* (voir l'exemple 6 ci-dessus), mais on trouve une attestation du verbe *måtte* (voir l'exemple 4 ci-dessus). L'auxiliaire modal renforce la modalité épistémique exprimée par *synes / se ut (til/som)*, autrement dit, il ajoute à l'énoncé un autre élément d'atténuation.

### 7.3.3 Le cotexte immédiat

#### 7.3.3.1 Répartition des constructions syntaxiques

La construction syntaxique la plus utilisée, et de loin, est celle où *synes* ou *se ut (til/som)* précèdent un infinitif, c'est-à-dire qu'ils fonctionnent en quelque sorte comme des semi-auxiliaires. Pour le verbe *synes*, 29 sur 33 attestations épistémico-modales de ce verbe sont de ce type en linguistique, 15 sur 20 en médecine :

13. Det massekulturelle mangfaldet **synest** i det minste **å ha** ein trygg språkleg basis i det ordboksførte standardspråket. (noling02)  
'La diversité culturelle **semble** au moins **avoir** un solide fondement linguistique dans le langage standard qu'on trouve dans les dictionnaires.'
14. En rekke studier **synes å indikere** at så høye oksygenivåer, selv om de varer i bare noen minutter, kan være skadelig (17). (nomed06)  
'Une série d'études **semblent indiquer** que des taux d'oxygène aussi élevés peuvent être nuisibles, même s'ils ne durent que quelques minutes (17).'

Le peu d'occurrences épistémico-modales qui restent sont réparties assez équitablement entre la fonction attributive (voir l'exemple 15) et la construction impersonnelle (voir l'exemple 16). Dans les articles de linguistique, on relève deux attestations de chaque construction, et dans les articles de médecine, trois attestations de *synes* comme verbe attributif et deux attestations de *synes* dans une construction impersonnelle.

15. At *-a* og *-at* skulle være eldre og yngre nøytrumsformer av et (flektbart) pronomen **synes** lite sannsynlig (noling08)

'Que *-a* et *-at* soient des formes plus anciennes et plus jeunes d'un pronom (flexionnel) **semble** peu probable'

16. Vi konkluderer at det **synes** som romluft er like effektivt som 100 % oksygen ved gjenoppliving av nyfødte, men det er nødvendig med flere studier for å klargjøre dette spørsmålet. (nomed06)

'Nous concluons qu'il **semble** qu'une réanimation des nouveaux-nés avec de l'air est aussi efficace qu'une réanimation avec de l'oxygène à 100 %, mais il faut plus d'études pour clarifier cette question.'

En pourcentage, cela nous donne les chiffres suivants:

**Tableau 5** : Répartition des constructions syntaxiques de *synes*

Constr.syntaxique	linguistique	médecine
Semi-auxiliaire	88 %	75 %
Verbe attributif	6 %	15 %
Constr. impersonnelle	6 %	10 %

En ce qui concerne l'emploi épistémico-modal de *se ut* (*til/som*), 29 des 48 attestations chez les linguistes sont des occurrences de *se ut til* + verbe à l'infinitif, c'est-à-dire un emploi de *se ut til* comme semi-auxiliaire (voir l'exemple 17). En médecine, cette construction concerne 6 sur 7 attestations (voir l'exemple 18).

17. Dette eksempelet **ser** altså **ut til** å bekrefte preferanseorganiseringen heller enn å avkrefte den. (noling14)

'Cet exemple **paraît** confirmer le classement par préférence plutôt que de le réfuter.'

18. ...1 mg daglig ser ut til å være en adekvat dosering. (nomed04)

'...1 mg par jour **paraît** être une dose adéquate.'

Dans les textes linguistiques il y a en outre 6 occurrences de *se ut* (*som*) comme verbe attributif :

19. Inntil videre **ser** det altså **ut** som en immuniseringsstrategi å ikke ville bry seg om diakrone data. (noling05)

'Jusqu'à nouvel ordre, cela **semble** une stratégie d'immunisation que de ne pas vouloir tenir compte de données diachroniques.'

En outre, on trouve, dans l'ensemble du corpus, 14 attestations de la construction impersonnelle, 13 dans les articles de linguistique (voir l'exemple 20) et 1 dans les articles de médecine.



20. ..., men det **ser ut til** at betydningen i all hovedsak er knyttet til ett eller flere av følgende domener: (noling06)  
 '..., mais il **semble** que le sens soit essentiellement lié à un ou plusieurs des domaines suivants :'

En pourcentage, nous aurons les chiffres suivants:

**Tableau 6** : Répartition des constructions syntaxiques de *se ut*

Constr.syntaxique	linguistique	médecine
Semi-auxiliaire	60,5 %	86 %
Verbe attributif	12,5 %	0 %
Constr. impersonnelle	27 %	14 %

Pour récapituler, on peut constater que *synes* et *se ut til* s'emploient beaucoup plus souvent devant un infinitif que ne le font leurs correspondants en français (voir la section 7.2.5.1). Bien que les constructions syntaxiques possibles soient dans une large mesure les mêmes pour les deux langues et que le sémantisme des deux paires de marqueurs soit assez similaire en norvégien et en français, il s'avère que les différentes constructions syntaxiques ne s'utilisent pas dans la même mesure dans les deux langues : la fonction « semi-auxiliaire » est beaucoup plus fréquente pour les marqueurs norvégiens, alors que le rôle de verbe attributif est beaucoup plus fréquent pour les marqueurs français. Cependant, le verbe modifié par *synes* / *se ut til* est le plus souvent *være* ('être') et dans beaucoup de ces cas, *synes* / *se ut til* + *være* prennent le rôle d'un verbe attributif. La différence au niveau des constructions syntaxiques pourrait donc s'expliquer par le fait que *sembler* et *paraître* acceptent plus facilement que leurs correspondants norvégiens une construction directe avec un attribut du sujet ; en norvégien le verbe *være* s'interpose plus souvent entre le modalisateur et l'attribut du sujet. Cela vaut en particulier pour *se ut* : la construction *se ut* + attribut du sujet est rare dans les articles de recherche parce qu'elle implique un sens d'apparence assez fort, alors que l'aspect de jugement est moins important. Dans un texte scientifique on dirait donc plutôt *dette ser ut til å være viktig* ('cela paraît être important') que *dette ser viktig ut* ('cela paraît important / cela a l'air important'). Pour ce qui est des constructions impersonnelles, celles-ci sont fréquentes avec *sembler* et *se ut (til/som)*, notamment chez les linguistes, mais relativement rares avec *paraître* et *synes*.

En ce qui concerne le type de verbes et d'attributs combinés avec *synes* / *se ut (til/som)*, il n'y a pas de grandes différences entre les deux disciplines. Comme déjà mentionné, le verbe le plus souvent modifié par *synes* / *se ut til* dans l'emploi infinitif est *være*

(‘être’), et cela vaut pour les deux disciplines. Dans les articles de linguistique, 10 des 29 occurrences de *synes* comme semi-auxiliaire sont combinées avec ce verbe. Dans les articles de médecine, il s’agit de 7 sur 15 occurrences comme semi-auxiliaire. Pour *se ut til*, 16 sur 29 occurrences « semi-auxiliaire » dans le corpus linguistique se combinent avec *være* et 3 sur 6 dans les articles de médecine. Cela veut dire qu’en tout, environ la moitié des occurrences sont suivies du verbe *være*. En ce qui concerne les autres verbes principaux, il n’est pas possible de détecter des tendances spécifiques dans les deux groupes.

Quant aux types d’attributs combinés avec *synes / se ut (til/som)* (directement ou par l’intermédiaire de *være*), on retrouve dans les articles médicaux norvégiens le concept de fréquence (exemples 21-22), mais il y a aussi quelques exemples d’adjectifs marqués d’une plus grande subjectivité (exemple 23) :

21. Totalmengden prednisolon **synes** således å være større i bruddgruppen. (nomed03)  
 ‘La quantité totale de prednisolone **semblait** être plus grande dans le groupe avec fractures.’
22. Kvinner hadde økende frekvens av ulcus perforatum med økende alder, mens det blant menn **så ut til** å være høyest forekomst i aldersgruppen 40 - 49 år. (nomed20)  
 ‘Chez les femmes il y avait une augmentation du taux d’ulcère perforé avec l’âge, alors que parmi les hommes la fréquence **semblait** être la plus élevée dans le groupe d’âge de 40 à 49 ans.’
23. Det **synes** påfallende at vi nå har påvist oesophagusstrikturer som følge av gastroøsofageal refluks hos såpass mange barn over den påfølgende toårsperioden. (nomed11)  
 ‘Il **semble** frappant que nous ayons relevé à présent chez autant d’enfants des strictions de l’œsophage en raison de reflux gastro-œsophagien durant la période de deux ans qui suivait.’

Les linguistes utilisent aussi de temps en temps des attributs à caractère axiologique :

24. ..., og det må **synes** meget betenkelig, da det ikke finnes spor av noen slik bruk av \*aina i andre germanske språk. (noling08)  
 ‘..., et cela doit **sembler** très discutabile, étant donné qu’il n’y a aucune trace d’un tel emploi de \*aina dans d’autres langues germaniques.’
25. Det kan **synes** å være et merkelig fenomen, og må være en etterlevning fra riktig gammel tid. (noling08)  
 ‘Cela peut **sembler** un drôle de phénomène, et cela doit être un vestige des anciens temps.’

Cependant, l’attribut paraît dans plusieurs cas assez « neutre », en ce sens qu’il sert à identifier plutôt qu’à caractériser le sujet :

26. I norsk - i allfall i nynorsk - **ser** den primære leieforma for verb **ut til** å vere **presensforma**. (noling04)

'En norvégien – du moins en *nynorsk* – la principale forme de base pour les verbes **paraît** être **la forme au présent**.'

27. Alt dette **ser ut til** å være **hentydninger til at han ikke snakker ren kristiansundsdialekt lenger**. (noling14)

'Tout cela **paraît** être **une allusion au fait qu'il ne parle plus un dialecte pur**.'

A partir du nombre limité d'exemples que l'on possède, il n'est pas possible de repérer des tendances spécifiques chez les linguistes et les médecins-chercheurs en ce qui concerne l'emploi de types d'attributs.

### 7.3.3.2 L'emploi de modalisateurs accumulés

Nous avons vu dans le chapitre précédent que *sembler* et *paraître* se combinent très rarement avec d'autres modalisateurs épistémiques. Il n'en va pas de même pour leurs correspondants en norvégien, qui se combinent avec d'autres modalisateurs épistémiques dans 14 cas. *Synes* apparaît avec d'autres modalisateurs épistémiques 8 fois ; 5 dans les articles de linguistique et 3 dans les articles de médecine. *Se ut (til/som)* apparaît avec d'autres modalisateurs à 6 reprises, toutes se trouvant dans les articles de linguistique. Dans un de ces cas, *se ut (til/som)* se combine avec deux autres marqueurs épistémico-modaux :

28. Nå **kan vel** det å argumentere mot en 27 år gammal analyse **se ut som** å slå inn ei vidåpen dør (noling05)

'**Je suppose que** le fait d'argumenter contre une analyse qui date d'il y a 27 ans **peut sembler** enfoncer des portes ouvertes'

La collocation la plus fréquente est décidément celle de *synes / se ut (til/som)* + le verbe *kunne*, normalement au présent (*kan*) (exemples 29-30). Cette combinaison représente environ deux tiers des cas. Parmi les autres marqueurs utilisés on peut citer *indikere* ('indiquer'/'suggérer'), *anta* ('supposer'), *sannsynlig* ('probable'), *måtte* ('devoir') et *peke i retning* ('indiquer') (exemples 31-32).

29. Rischels problem **kan synes** å være av rent terminologisk karakter. (noling10)

'Le problème de Rischels **peut sembler** d'un caractère purement terminologique.'

30. Visse generative grammatikere **kan se ut til** å ha gitt bøyingsklasser en rolle som uavhengige diakritika i leksikon. (noling05)

'Certains grammairiens générativistes **peuvent sembler** avoir donné aux conjugaisons un rôle de diacritiques indépendants dans le lexique.'

31. Det **synes** naturlig å **anta** at narkotikadødsfall henger sammen med mengden heroin som er inntatt og kombinasjonen av heroin og andre rusmidler. (nomed08)

'Il **semble** naturel de **supposer** que les surdosages sont associés au dosage d'héroïne pris et à la combinaison d'héroïne avec d'autres stupéfiants.'

32. For *pri* **synes** eksemplet med *sneg* å **peke i denne retning, men jeg har ikke gjort grundige nok undersøkelser til å kunne trekke noen klare konklusjoner**. (noling06)

'Pour *pri*, l'exemple avec *sneg* **semble indiquer** cela, *mais les examens que j'ai effectués ne sont pas suffisamment approfondis pour tirer des conclusions définitives*.'

Dans ce dernier exemple, on note que, outre la cooccurrence avec un autre marqueur spécifique (*peke i retning*), *synes* est suivi d'une tournure plus longue (mise en italiques gras dans l'exemple) qui, dans son ensemble, qualifie le même contenu informatif que *synes*. Il y a encore trois exemples de ce type (cf. par exemple l'extrait 10 ci-dessus), où une tournure plus ou moins longue renforce l'aspect d'incertitude lié à l'information qualifiée par *synes* / *se ut (til/som)*. Par cet ajout d'un élément d'incertitude à l'égard de l'information transmise, ces tournures jouent un rôle semblable à celui d'un deuxième modalisateur. Elles soulignent aussi qu'il s'agit d'une incertitude réelle, et non d'un emploi simplement stylistique.

### 7.3.3.3 Le rôle du datif à la première personne

Bien que l'adjonction d'un syntagme prépositionnel en fonction de pronom datif soit possible, *synes* et *se ut til/som* sont rarement accompagnés d'un tel syntagme. Une telle adjonction est superflue, car l'interprétation par défaut est que la source du jugement est le locuteur. Les syntagmes *for meg* / *for oss* ('à/pour moi' / 'à/pour nous') ne sont donc ajoutés que dans certains cas spécifiques où le locuteur veut mettre en relief l'aspect subjectif et personnel. Il y a une seule occurrence de cet emploi dans le corpus, et elle apparaît dans un cotexte où l'auteur émet des remarques critiques contre un des ses pairs. La mise en relief de l'aspect subjectif sert à atténuer la critique et à la rendre moins offensive.

33. **For meg ser det ut til** at Nedergaard Thomsen her opererer med et idiosynkratisk grammatikalitetsbegrep som leserne dessverre ikke får del i.

'Il **me semble** que Nedergaard Thomsen se sert ici d'un concept idiosyncrasique de la grammaire auquel les lecteurs ne sont malheureusement pas initiés.'

L'ajout d'un tel marqueur de subjectivité ne rend définitivement pas l'énoncé plus sûr (comme c'est souvent le cas pour les expressions correspondantes en français, cf. Nølke 2001), au contraire il semble l'atténuer plus encore, en ce qu'il signale qu'il ne s'agit que d'une opinion personnelle et que d'autres peuvent être d'un avis différent. Dans la mesure où

l'analyse proposée par Nølke est exacte, il semble donc y avoir là une différence notable entre deux marqueurs apparemment équivalents.

#### 7.3.4 Fonctions pragmatiques de *synes* et de *se ut (til/som)*

En suivant les mêmes principes d'analyse que ceux adoptés pour l'analyse de *sembler / paraître*, nous examinerons ici les occurrences de *synes / se ut (til/som)* pour essayer de déterminer quelles sont les fonctions pragmatiques les plus utilisées. Il s'avère que dans les articles de médecine, *synes* et *se ut (til/som)* apparaissent dans presque tous les cas avec une fonction « contenu », et dans la grande majorité de cas, il s'agit de présenter ou de rapporter des résultats issus soit de la recherche des auteurs eux-mêmes, soit de recherches effectuées antérieurement par d'autres chercheurs. Ces résultats ne sont manifestement pas (encore) acceptés par la communauté scientifique comme des faits établis, et ils sont donc présentés d'une manière atténuée :

34. Bilateralitet forekommer i 28 - 50 % av tilfellene, og overlevelsen **synes** utelukkende relatert til spredning utenfor ovariene (24, 25). (nomed18)<sup>23</sup>  
'Il existe une bilatéralité dans 28 – 50 % des cas, et la survie **semble** exclusivement liée à la propagation en dehors des ovaires (24, 25).
35. Vaksinen **ser ut til** å redusere antall tilfeller med otitis media med 8 % (17). (nomed09)  
'Le vaccin **semble** réduire le nombre de cas d'otite moyenne de 8 % (17)'.

On retrouve cet emploi dans les articles de linguistique aussi. Voici deux exemples tirés de conclusions d'articles<sup>24</sup> :

36. Konklusjonen blei at selv om 'bestemthet' ikke alltid gir unik referanse, **synes** det å være slik at det er en tydelig korrelasjon mellom 'bestemthet' og 'spesifisitet' (noling09)  
'La conclusion a été que bien que la 'définitude' ne donne pas toujours une référence unique, il **semble** y avoir une corrélation claire entre 'définitude' et 'spécificité'
37. Det **ser ikke ut til** at noen av de utvalgte faktorene alene er bestemmende for vanskelighetsgraden av en setning. (noling11)  
'Aucun des facteurs sélectionnés ne **semble** pouvoir à lui seul déterminer le niveau de difficulté d'une phrase.'

---

<sup>23</sup> Les chiffres entre parenthèses renvoient aux références bibliographiques d'où sont issus les résultats rapportés.

<sup>24</sup> Alors qu'on parle fréquemment de 'résultats' en médecine, on ne parle pas dans la même mesure de 'résultats' en linguistique. Cela s'explique par la nature des processus de recherche et des objets d'étude investigués (voir Fløttum et al. 2006 : 180-202). Rinck (2006) fait une observation similaire : l'usage du terme 'résultats' varie quantitativement et sémantiquement entre les disciplines et révèle des épistémologies différentes. Nous utilisons ici le terme 'résultats' dans un sens large; il comprend non seulement des résultats quantifiables, mais aussi des observations et des conclusions faites par l'auteur.

Comme déjà mentionné, la présence de tels atténuateurs peut être motivée par une incertitude réelle – les résultats sont peut-être peu clairs et ne permettent pas de tirer de conclusions définitives – ou simplement par un souci stylistique : les atténuateurs font partie des routines scripturales académiques.

Cet emploi de *synes / se ut (til/som)* est fréquent dans les articles de linguistique aussi bien que dans les articles de médecine, mais dans ces premiers, on trouve aussi une série d'autres emplois qui ne sont que peu ou pas représentés dans les articles de médecine. Il s'agit de diverses variantes de la fonction interpersonnelle, dont la plus fréquente est peut-être celle qui consiste à faire preuve de prudence en interprétant les idées d'autres chercheurs. Dans cet usage, les atténuateurs sont des moyens qui permettent à l'auteur de sauver la face vis-à-vis du lecteur s'il s'avère plus tard qu'il avait tort, par exemple qu'il avait mal interprété le travail de l'autre.

38. Av den grunn, og også fordi vår definisjon av sterke verb avviker noe fra den som **synes** å ligge til grunn for BOs definisjon, vil det nøyaktige antall sterke verb være noe lavere her enn det som er angitt for BO. (noling07)

'Pour cette raison, et aussi parce que notre définition des verbes forts diverge quelque peu de celle sur laquelle **semble** se fonder la définition de BO, le nombre exact de verbes forts serait un peu moins élevé que celui indiqué pour BO.'

39. I John Searles teori om språkhandlingar (Searle 1974) er talaren autonom aktør og eineforfattar av ytringa, [...] I Paul Grice sine "maxims of conversation" (Grice 1975), derimot, **synest** konversasjon å bli kartlagd som eit slags samarbeidsprosjekt. (noling13)

'Selon la théorie des actes de langages de John Searle (Searle 1974), le locuteur est acteur autonome et auteur unique de l'énoncé, [...] Dans les « maxims of conversation » de Paul Grice, par contre, la conversation **semble** être décrite comme une sorte de projet en commun.'

Un autre emploi qui est relativement fréquent dans les articles de linguistique, est celui d'adoucir la critique émise contre les travaux d'autres chercheurs. Cet emploi est motivé par de la politesse (les formules polies protègent le chercheur critiqué et permettent d'éviter une attaque trop vive), mais aussi par un désir de se prémunir soi-même contre les conséquences qui peuvent se produire s'il s'avère plus tard que la critique est injustifiée ou mal fondée.

40. Ifølge den sterke utgaven av UA/NI-hypotesen skal UA og NI betraktes som to gjensidig avhengige størrelser. En slik oppfatning **synes** det imidlertid ikke å være empirisk dekning for. (noling10)

'Selon la version forte de l'hypothèse UA/NI, UA et NI doivent être considérés comme deux entités mutuellement dépendantes. Une telle interprétation ne **semble** cependant pas être justifiée par l'expérience.'

41. Clarks formulering: "A presents a signal for B to understand" **synest** difor, reint terminologisk, å avspegla ein interessant ambivalens. (noling13)  
'Ainsi, la formule de Clark: "A presents a signal for B to understand" **semble** refléter, d'un point de vue purement terminologique, une ambivalence intéressante.'

On trouve aussi quelques exemples qui illustrent l'aspect persuasif de l'article de recherche et le souci de l'auteur de se positionner et de promouvoir sa propre recherche :

42. Prediktabilitetskriteriet seier ingenting om kvifor denne forma går igjen som leieform, men dette **ser ut til** å ha ei enkel forklaring, som følgjer av det eg skal kalle Frekvenskriteriet (noling04)  
'Le critère de prévisibilité ne dit rien sur les raisons pour lesquelles cette forme réapparaît comme forme de base, mais cela **paraît** avoir une explication simple qui relève de ce que j'appellerai le critère de fréquence'
43. At -a og -at skulle være eldre og yngre nøytrumsformer av et (flekterbart) pronomen **synes** lite sannsynlig også av den grunn at hverken subjektet eller objektet i norrøne a/at-setninger er spesielt bundet til nom./akk. sg. ntr. Det vet alle norrønfilologer, men jeg fører likevel opp noen bevisende eksempler: (noling08)  
'Que -a et -at soient des formes plus anciennes et plus jeunes d'un pronom (flexionnel) **semble** peu probable également à cause du fait que ni le sujet ni l'objet dans des phrases nordiques en a/at ne sont spécifiquement liés à nom./acc. sing. neutre. Tous les philologues des langues nordiques savent cela, mais j'avance quand même certains exemples qui le démontrent :'

Dans l'exemple 43, l'auteur cherche à convaincre ses lecteurs en maintenant que ce qu'il dit est évident. Malgré les atténuateurs, le ton dans cet exemple est assez pontifiant (cf. *Det vet alle norrønfilologer...* ('Tous les philologues des langues nordiques savent cela...')) et dans un tel cotexte, on pourrait dire que les atténuateurs contribuent à une fausse modestie plutôt que de faire preuve d'une attitude d'ouverture envers d'autres points de vue.

Tout comme *sembler* et *paraître*, *se ut til/som* s'utilise parfois pour rendre compte d'une impression ou d'une opinion possible à laquelle l'auteur ne s'attache pas, et qu'il rejette dans l'argumentation qui suit (*synes* par contre, apparaît rarement dans cet emploi). L'impression initiale s'avère donc, selon l'auteur, partiellement ou entièrement erronée :

44. I eksemplene over **ser** det faktisk **ut til** at det også er subjektsvegring. Hvis subjektsvegringen viser seg å være uavhengig av om DENNE og antesedenten befinner seg innenfor samme helsetning, står LF-flyttings analysen som er skissert ovenfor, mindre trygt. Det vil uvegerlig være mindre overbevisende om det som **ser ut til** å være samme fenomen, må ha to forskjellige forklaringer avhengig av hvor helsetningsgrensen går. Men det viser seg at subjektsvegringen ikke er like sterk utenfor setningsdomenet - alle forekomstene av DENNE nedenfor har subjekt som antesedent (noling01)  
'Dans les exemples ci-dessus, il **semble** en fait aussi y avoir une orientation anti-sujet. Si l'orientation anti-sujet s'avère indépendante de la question de savoir si DENNE et

l'antécédent se trouvent dans la même phrase principale ou non, l'analyse esquissée ci-dessus apparaît moins solide. Elle serait inévitablement moins convaincante si ce qui **paraît** être le même phénomène nécessite deux explications différentes selon la position des limites de la phrase principale. Mais il s'avère que l'orientation anti-sujet n'est pas aussi forte en dehors du domaine de la phrase – toutes les occurrences de DENNE citées ci-dessous ont le sujet comme antécédent '.

Ce type d'emploi intervient souvent dans un contexte où l'auteur rend compte d'un raisonnement. Tel est le cas dans l'exemple 44, qui est une bonne illustration d'un phénomène assez courant dans les articles de linguistique : l'auteur semble inclure le lecteur dans le processus de recherche lui-même en lui permettant de suivre avec lui le raisonnement qui précède la conclusion. Ainsi, l'auteur ne présente pas seulement au lecteur les résultats finaux, mais il l'invite en quelque sorte à prendre part au processus de recherche. C'est ce que Rinck (2006) appelle « la mise en intrigue de la découverte » : le lecteur entre dans la démarche de découverte. Il y a là une différence notable entre le processus de recherche en linguistique et celui en recherche médicale : les médecins-chercheurs rédigent leurs articles après coup, c'est-à-dire une fois que les expériences ou les études sont effectuées et terminées, alors qu'en linguistique, les processus de recherche et de rédaction d'article sont présentés dans une plus grande mesure comme des processus concomitants (voir Fløttum et al. 2007).

### 7.3.5 Bilan

En résumé, l'examen de l'emploi de *synes* et de *se ut (til/som)* a révélé que *synes* est fréquemment employé comme modalisateur épistémique dans les deux disciplines, alors que *se ut (til/som)* apparaît rarement dans les articles de médecine mais fréquemment chez les linguistes. On constate donc un certain parallélisme entre l'emploi de *synes* et *se ut* en norvégien et l'emploi de *sembler* et *paraître* en français. Alors que *sembler* et *synes* s'utilisent fréquemment dans les deux disciplines et figurent parmi les modalisateurs épistémiques les plus fréquents, *paraître* et *se ut (til/som)* sont utilisés relativement souvent par les linguistes, mais ils le sont peu dans les articles de médecine.

En ce qui concerne les constructions syntaxiques, ce sont les emplois comme « semi-auxiliaires » modaux qui dominent largement pour les deux marqueurs (*synes* et *se ut (til/som)*) et dans les deux disciplines. Le verbe modifié est pourtant très souvent *være*, et la construction correspond alors sémantiquement à l'emploi attributif, qui était l'emploi le plus fréquent pour *sembler* et *paraître*. En ce qui concerne la nature des attributs employés, ceux-ci sont parfois de caractère axiologique (pour reprendre la terminologie de Kerbrat-Orrechioni 2002), chez les médecins-chercheurs aussi bien que chez les linguistes. Les exemples sont



pourtant peu nombreux et pour ce qui est du cotexte immédiat, il n'est pas possible de détecter des tendances claires qui distinguent les deux disciplines. Les différences observées dans le corpus français (cf. 7.2.5.1) ne se retrouvent donc pas dans le corpus norvégien, du moins pas de façon aussi nette. Un corpus d'exemples plus grand serait nécessaire pour déterminer dans quelle mesure ces tendances sont effectives.

En ce qui concerne les fonctions pragmatiques, en revanche, une nette différence est repérable : dans les articles de médecine, *synes* et *se ut (til/som)* assument presque sans exception une fonction contenu – et dans la grande majorité des cas il s'agit d'une présentation prudente de résultats – alors que les articles des linguistes font preuve d'emplois beaucoup plus variés. On retrouve dans les articles de linguistique l'emploi consistant à rester nuancé en présentant des résultats et des conclusions, mais également des exemples de fonctions interpersonnelles : *synes* et *se ut (til/som)* servent à modérer la critique contre d'autres chercheurs, à faire preuve de prudence en interprétant leurs idées et à présenter des impressions ou des points de vue envisageables mais non soutenus par l'auteur. Ces types d'emploi, qui illustrent tous la fonction interpersonnelle de l'atténuation, sont absents des articles de médecine de notre corpus. Rappelons qu'une tendance similaire se dessine dans le corpus français (pourtant beaucoup moins nettement) : les linguistes utilisent *sembler* et *paraître* dans une fonction interpersonnelle d'atténuation de critique, alors que cet usage est absent chez les médecins-chercheurs. Cependant, il est clair qu'en ce qui concerne l'aspect pragmatique, les différences entre disciplines se dessinent plus clairement pour l'emploi des marqueurs norvégiens *synes* et *se ut (til/som)* que pour les marqueurs français *sembler* et *paraître*.

#### 7.4 Pouvoir et kunne

L'auxiliaire *pouvoir* / *kunne* est un candidat évident pour des études plus approfondies. Dans notre corpus, *kunne* s'est révélé le principal modalisateur épistémique dans les articles norvégiens, en linguistique aussi bien qu'en médecine (cf. tableau 4, 6.4.2), ce qui corrobore l'observation faite par Breivega (2003 : 156), selon laquelle *kunne* est le marqueur épistémique dominant dans le discours scientifique norvégien. Des études portant sur le discours scientifique anglais montrent que la même observation est valable pour le quasi-équivalent *may*. Plus largement, il semble que ces auxiliaires modaux (*pouvoir*, *kunne*, *may*) représentent traditionnellement les marqueurs épistémico-modaux les plus typiques. La lecture d'études et de présentations de la notion de modalité révèle à quel point ces marqueurs modaux sont considérés comme faisant partie des plus importants : figurant parmi les premiers marqueurs mentionnés, on leur accorde généralement beaucoup de place et on les emploie souvent dans les exemples prototypiques qui servent à illustrer ce qu'est la modalité épistémique (voir par exemple Faarlund et al. 2002 : 579 sqq, Le Querler 1996). Cette idée que *pouvoir* / *kunne* / *may* sont le prototype des marqueurs épistémico-modaux est bien fondée en ce qui concerne le norvégien et l'anglais (cf. les études mentionnées ci-dessus et la présente étude), mais elle est moins évidente pour le français. Il ressort du tableau 3, section 6.4.2, que les articles examinés témoignent d'un emploi limité de la valeur épistémique de *pouvoir*, ce qui corrobore l'observation de Le Querler (2001 : 22) selon laquelle la valeur épistémique de *pouvoir* est relativement rare. La description de *pouvoir* comme marqueur prototypique de la modalité épistémique ne semble donc pas avoir de fondement quantitatif<sup>25</sup>.

Il est intéressant de souligner les différences qui existent entre l'emploi de *pouvoir* épistémique en français et celui de *kunne* épistémique en norvégien. Alors que la valeur épistémique semble constituer une valeur plutôt mineure de *pouvoir*, l'emploi épistémique de *kunne* constitue une valeur importante qui s'utilise assez fréquemment dans les articles norvégiens, notamment dans ceux des médecins-chercheurs. Une différence du même ordre est observable dans le corpus français : les médecins-chercheurs utilisent *pouvoir* épistémique plus que les linguistes. Il existe donc des différences interlangagières aussi bien qu'interdisciplinaires en ce qui concerne l'emploi de *pouvoir* / *kunne* épistémiques. La présente section vise à préciser en quoi elles consistent.

---

<sup>25</sup> Il convient pourtant de noter que les ouvrages français que nous avons consultés portent sur les verbes modaux en général (Dendale et Auwera 2001) ou sur *pouvoir* en particulier (Le Querler 1996) plutôt que sur la modalité épistémique comme catégorie prise dans sa globalité, et cela peut évidemment expliquer l'accent spécifique mis sur *pouvoir* dans ces ouvrages.

La même démarche sera suivie que pour *sembler / paraître* et *synes / se ut (til/som)* : nous commencerons par un bref aperçu du fonctionnement syntaxique de *pouvoir / kunne*, pour en venir ensuite à détailler leur sémantisme. Après avoir rendu compte des critères appliqués pour identifier *pouvoir / kunne* épistémiques, nous passerons à l'analyse comparative des langues et disciplines. D'abord sera abordée la question de la fréquence de *pouvoir / kunne* dans le corpus. Suivront alors quelques observations sur le cotexte immédiat environnant les occurrences épistémiques ainsi qu'un exposé des fonctions pragmatiques véhiculées par *pouvoir / kunne* épistémiques.

#### 7.4.1 Fonctionnement syntaxique de *pouvoir / kunne*<sup>26</sup>

La principale fonction syntaxique de *pouvoir* et de *kunne* est, évidemment, celle d'auxiliaire modal :

1. On **peut** constater que beaucoup se sont effectivement posé la question de cette façon (frling12)
2. På denne måten **kan** de etablere en felles referanseramme og et utgangspunkt for en mer symmetrisk samtale. (noling14)  
'De cette manière, ils **peuvent** établir un cadre commun de référence et un point de départ pour une conversation plus symétrique.'

Mais ils fonctionnent parfois comme des verbes principaux, du moins lorsque ils dénotent une capacité :

3. Han **kan** leksa si, men betyr det at han også etterlever disse reglene når han deltar i samtaler? (noling16)  
'Il **sait** sa leçon, mais est-ce que cela implique qu'il suit ces règles quand il participe à des conversations ?'
4. [...] le corps et les algorithmes **peuvent** beaucoup de choses que la langue ne **peut** pas. (frling12).

La vaste majorité des occurrences relevées dans le corpus correspond à des attestations de *kunne / pouvoir* comme auxiliaires. En tout, il y a six attestations de *pouvoir / kunne* comme verbes principaux - trois dans le corpus Frling (*pouvoir*) et trois dans le corpus Noling (*kunne*). A l'exception d'une occurrence de la séquence *il se peut que*, il s'agit dans tous les cas d'un emploi dénotant la capacité.

---

<sup>26</sup> Notons que les exemples dans cette section permettent d'illustrer les fonctions syntaxiques, et il ne s'agit pas d'exemples de la modalité épistémique.

### 7.4.2 Contenu sémantique

Si le fonctionnement syntaxique de *kunne* / *pouvoir* est relativement simple et rapide à décrire, il est bien plus compliqué de rendre compte du contenu sémantique de ces verbes. Cette difficulté est sans doute due à la grande flexibilité de ces verbes polysémiques : ils possèdent un champ sémantique très étendu et sont susceptibles d'exprimer des valeurs aussi diverses que la permission, la capacité, l'occasion, etc. A priori, *pouvoir* ainsi que *kunne* expriment avant tout la *possibilité*. Il existe cependant toute une gamme de possibilités d'ordre différent, et le cotexte sert souvent à orienter l'interprétation vers une valeur plus précise. Le Querler (1996 : 120) propose de voir le sémantisme de *pouvoir* comme s'organisant autour d'une valeur de base sous-déterminée, celle qu'elle appelle la *possibilité abstraite*. Cette valeur est sous-déterminée en ce sens qu'on doit pouvoir en préciser le sens, pourvu que le cotexte le permette. La valeur de base constitue un noyau de sens commun à toutes les valeurs plus déterminées de *pouvoir*. Le Querler précise que, dans bien des cas, l'interprétation restera à ce niveau sous-déterminé faute d'indices plus précis dans le cotexte. Il est raisonnable de supposer l'existence d'une valeur de base similaire pour *kunne* en norvégien.

On trouvera ci-dessous une liste des différentes valeurs déterminées de *pouvoir*. Cette liste s'inspire notamment des travaux de Le Querler (1996, 2001), qui s'affilient eux-mêmes à des études antérieures sur le sémantisme de *pouvoir*, notamment celles de J.P. Sueur (1975, 1979, 1983). Ici, chaque valeur est mise en rapport avec un des cinq types de modalité identifiés par von Wright (1951) (cf. *supra*, 3.1). Nous distinguerons donc les valeurs suivantes : une valeur déontique (permission), une valeur dynamique (capacité, occasion), une valeur aléthique (possibilité vérifiable ou affirmée), une valeur épistémique (possibilité hypothétique / éventualité) et une valeur existentielle (sporadicité). La valeur de base, le *noyau de sens*, sera simplement appelée *possibilité*. Soit le schéma suivant :

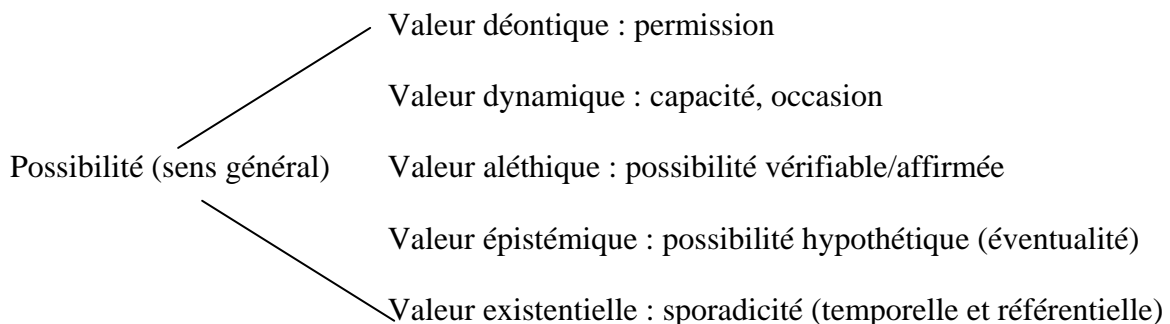


Figure 1 : Valeurs de *pouvoir*

La valeur **déontique** de *pouvoir* dénote la permission<sup>27</sup> :

I *Il peut venir au cinéma, ses parents l'y autorisent. (Il a la permission d'aller au cinéma.)*

La valeur **dynamique** la plus solidement établie est celle de capacité, dénotant des facultés ou des aptitudes :

II a) *Il peut venir à pied, sa jambe est déplâtrée. (Il est capable d'aller à pied.)*

Mais elle peut aussi dénoter l'occasion ou la disponibilité (cf. Bentley 2003 : 64, qui définit la modalité dynamique comme « la modalité de l'aptitude et la disponibilité » :

II b) *Il peut venir, il n'a rien prévu d'autre. (Il a l'occasion de venir.)*

La valeur **aléthique** de *pouvoir* dénote une possibilité théorique vérifiable, une possibilité affirmée :

III *L'équipe française peut encore gagner la coupe du monde. (A ce stade, l'équipe française a la possibilité théorique de gagner la coupe (mais cela suppose qu'elle gagne le prochain match).)*

La valeur **épistémique** est celle d'une possibilité bilatérale, qui exprime une hypothèse :

IV *Il peut pleuvoir en Angleterre en ce moment (mais je ne sais pas s'il pleut effectivement ou non.)*

La valeur **existentielle** exprime la sporadicité, qui peut être temporelle (désignant quelque chose qui arrive de temps en temps),

V a) *Jean peut être odieux. (Jean est odieux de temps en temps)*

ou référentielle, désignant quelque chose qui se produit chez certains, mais non chez d'autres :

V b) *Les Alsaciens peuvent être obèses. (Certains Alsaciens sont obèses, mais pas tous.)*

En répertoriant ces valeurs précises, on risque de donner l'impression qu'il existe des distinctions nettes entre les divers effets de sens et qu'il est toujours possible d'attribuer à une occurrence de *pouvoir* une des valeurs énumérées. Cependant, les choses ne sont évidemment pas si simple. Afin d'illustrer les difficultés que l'on rencontre pour attribuer à *pouvoir* une valeur précise, Le Querler (1996 : 121) donne à titre d'exemple l'énoncé suivant, dont elle envisage quatre lectures possibles :

*On peut orienter avec des examens, des concours, des dossiers ou des entretiens*

---

<sup>27</sup> Nous reprenons ici certains des exemples cités par Le Querler (I, IIa, IV et Va,b) (1996 : 115 et 118). Les exemples IIb et III sont nos propres exemples (forgés) ; l'exemple III s'inspire d'un exemple norvégien élaboré par Kolflaath (correspondance personnelle 01.10.2003).

Les lectures possibles que cite Le Querler sont la sporadicité (*on oriente parfois...*), la capacité (*les professeurs ont les qualités nécessaires d'orienter...*), la permission (*ils sont autorisés à orienter...*) et ce que l'auteur appelle la possibilité matérielle (ibid. : 122) (*on a les moyens d'orienter...*). En dépit de ces interprétations multiples, l'énoncé est tout à fait compréhensible et il n'est pas nécessaire de choisir l'une ou l'autre des valeurs. Plusieurs valeurs sont co-présentes dans le même énoncé, et elles contribuent toutes, à des degrés divers, à sa signification. C'est ce que Fuchs (1996 : 31) appelle l'indétermination interprétative. Dans ce cas, on se contente d'attribuer à *pouvoir* sa valeur de base de possibilité.

On trouve des exemples similaires dans notre corpus ; ainsi en est-il des séquences du type *on peut dire que...* et *det kan hevdes at....*. Dans ces cas, *pouvoir* / *kunne* font partie d'une locution relativement figée dont le rôle est d'introduire un contenu informatif. D'un point de vue sémantique, *pouvoir* et *kunne* n'apportent pas grand chose de plus et sont employés dans un sens assez « pauvre ». Il n'est pas envisageable de rattacher à de telles occurrences un type spécifique de possibilité ; l'interprétation en reste ainsi à un certain niveau de sous-détermination (voir Le Querler 1996 : 120), ce qui ne pose aucunement problème pour la compréhension de l'énoncé<sup>28</sup>.

Par ailleurs, Le Querler répertorie, outre les valeurs traditionnellement reconnues, une série d'effets de sens discursifs qu'elle ne considère pas comme des valeurs indépendantes de *pouvoir*, mais comme « la résultante de la prise en compte de l'énoncé dans son ensemble, voire même d'une partie plus large du discours » (Le Querler 1996 : 124). Ces effets de sens comprennent, entre autres, la concession (*Il peut venir, je ne lui adresserai pas la parole*) et la délibération (*Je me demande comment il a pu venir*).

La liste des valeurs répertoriées ci-dessus ne se veut donc pas une liste exhaustive des différents effets de sens que *pouvoir* est susceptible de produire. Elle identifie certains domaines de sens, qui peuvent aisément être illustrés, mais il n'en reste pas moins que, dans des énoncés attestés, *pouvoir* peut susciter des effets de sens qui ne se laissent pas nécessairement rattacher à l'une ou l'autre des valeurs spécifiques et, avant tout, il peut être utilisé, comme nous l'avons vu, dans des cotextes où plusieurs valeurs sont compatibles et co-existantes, sans qu'un choix s'impose entre les différentes valeurs (cf. Fuchs 1996 : 31).

---

<sup>28</sup> Bien que *pouvoir* et *kunne* n'aient pas de valeur épistémique dans ces exemples, ce type de tournures fonctionne néanmoins comme un élément atténuateur, dans la mesure où le locuteur ne présente pas ce qui suit comme un fait établi, mais comme un point de vue parmi d'autres. Il n'assume même pas la responsabilité de ce point de vue, il l'attribue à quelqu'un d'indéfini (cf. *infra* 7.4.5)

Si cette liste de valeurs est élaborée pour le verbe français *pouvoir*, elle semble aussi valoir pour *kunne* en norvégien. Dans Faarlund et al. (2002 : 581), seulement trois interprétations possibles de *kunne* sont évoquées, la valeur épistémique : *Kåre kan være syk* ('*Kåre peut être malade*'), la valeur déontique : *Kåre kan reise hjem* ('*Kåre peut rentrer*') et la valeur dite dynamique : *Kåre kan svømme* ('*Kåre sait nager*'). Cependant, les valeurs de sporadicité et de possibilité affirmée existent aussi pour *kunne* en norvégien : *Bergensere kan være svært pratsomme* ('*Les Berguenois peuvent être très bavards*'), *Brann kan fortsatt vinne cupen* ('*Brann peut toujours gagner la coupe*'). En ce qui concerne la valeur dynamique, Faarlund et al. (2002 : 581) en contestent le statut modal, en maintenant que dans cet emploi, *kunne* fonctionne comme un verbe ordinaire avec un sens lexical plutôt que comme un auxiliaire modal avec un sens grammatical. Cette valeur ne relève donc pas, selon eux, du domaine de la modalité. Il est vrai que le statut modal de la valeur dynamique est moins évident que celui des autres valeurs. En témoigne par exemple le fait que c'est la valeur dynamique que l'on trouve dans les rares cas où *pouvoir* / *kunne* fonctionnent comme verbe lexical (cf. les exemples 3-4 ci-dessus). Dans la tradition anglo-américaine, il est pourtant courant d'associer la valeur de capacité (*ability*) à la modalité dynamique (voir Palmer 2001 : 10 et Vihla 2000 : 210) et donc de considérer la valeur dynamique comme une valeur modale sur le même pied que les autres valeurs. Rappelons enfin que *kunne* au sens dit dynamique se traduit en français tantôt par *pouvoir* à valeur de capacité (*han kan løfte selv de tyngste koffertter* = '*il peut soulever les valises les plus lourdes*'), tantôt par *savoir* (*han kan svømme* = '*il sait nager*'). *Kunne* au sens dynamique n'équivaut donc pas entièrement à *pouvoir* de capacité.

Pour résumer, *kunne* et *pouvoir* sont susceptibles d'exprimer un éventail de valeurs différentes, mais il est souvent difficile (et inutile) de rattacher une occurrence spécifique à une valeur précise. Comme la présente étude porte sur *pouvoir* et *kunne* en tant que modalisateurs épistémiques, une tâche importante a été d'identifier les occurrences épistémiques parmi toutes les occurrences repérées. A cause des frontières floues entre les différents domaines de sens, les nombreux cas d'indétermination et le nombre élevé de valeurs, ce travail de classification a été beaucoup plus compliqué pour les verbes *pouvoir* et *kunne* que pour un verbe tel que *synes*, dont les différents sens sont faciles à distinguer les uns des autres. S'il est aisé de fournir des exemples prototypiques de *pouvoir* épistémique, la tâche consistant à attribuer aux exemples authentiques telle ou telle valeur s'avère beaucoup plus ardue. Pour faciliter et systématiser cette classification, en cherchant à limiter la part de subjectivité et d'arbitraire, nous avons eu recours à certains critères visant à définir *pouvoir*

épistémique. Ces critères doivent être mis en relation avec les critères d'identification des modalisateurs épistémiques en général (voir section 6.4.1.3), mais ils concernent de manière plus spécifique ce seul marqueur et nous les présenterons donc ci-dessous.

### 7.4.3 *Pouvoir / kunne* épistémiques – critères de classification

On peut avoir recours à des critères syntaxiques aussi bien que sémantiques pour identifier les occurrences épistémiques de *pouvoir / kunne*. Pour ce qui est des critères syntaxiques, on peut tout d'abord constater qu'une interprétation épistémique ne peut se faire qu'à condition que le verbe soit utilisé comme auxiliaire, ou bien dans la séquence figée *il se peut que*. Dans les rares cas où *pouvoir / kunne* fonctionnent comme verbe principal, une lecture épistémique est exclue. Une lecture épistémique est également exclue quand *pouvoir / kunne* sont affectés par la négation (cf. Sueur 1979 : 107 ; Le Querler 1996 : 129 ; Kronning 2005), ce qui ne doit pas surprendre étant donné que la négation de *pouvoir / kunne* amène à une suppression de la bilatéralité qui caractérise la possibilité épistémique<sup>29</sup>. Ainsi, les occurrences négatives de *pouvoir / kunne* ne peuvent pas être épistémiques, et les occurrences épistémiques n'acceptent pas l'adjonction d'une négation sans qu'une altération de la valeur exprimée par *pouvoir / kunne* s'impose. Si l'on ajoute à l'énoncé *il peut pleuvoir en Angleterre en ce moment* une négation portant sur *pouvoir* (*il ne peut pas pleuvoir en Angleterre en ce moment*), cela entraînera un passage de l'interprétation épistémique à une interprétation radicale. L'énoncé n'exprime plus une hypothèse, une éventualité ou une incertitude, mais la simple négation d'une possibilité. L'adjonction d'une négation portant sur *pouvoir / kunne* épistémique oriente donc autrement l'interprétation du modal. En revanche, les autres valeurs – à l'exception de la valeur sporadique, qui exprime aussi une possibilité bilatérale, – acceptent la négation sans altération de la valeur exprimée à l'origine. Ainsi, les occurrences de *pouvoir* dans *il peut aller au cinéma* et *il ne peut pas aller au cinéma* expriment toutes deux une valeur déontique, celles dans *il peut parler l'anglais* et *il ne peut pas parler l'anglais* une valeur dynamique et celles dans *l'équipe française peut encore gagner* et *l'équipe française ne peut plus gagner* expriment une valeur aléthique (simple affirmation/négation d'une possibilité vérifiable).

Pour ce qui est des critères sémantiques, Le Querler (1996, 2001) met en évidence qu'un des traits qui caractérisent *pouvoir* épistémique est sa portée sémantiquement extra-prédicative, c'est-à-dire qu'il porte sur l'énoncé dans sa globalité, et non uniquement sur la relation sujet-verbe, comme c'est le cas pour les autres valeurs principales de *pouvoir* (sauf

---

<sup>29</sup> Sauf si l'on a une forme interro-négative : *ne-peut-il pas...* ?



pour la valeur sporadique). Le Querler (2001 : 25-26) illustre cette opposition de la manière suivante :

portée extra-prédicative : possibilité – [sujet - verbe], par exemple

*Elle peut avoir mal compris*, paraphrasable par *Peut-être qu'elle a mal compris*

portée intra-prédicative : [sujet - possibilité - verbe], par exemple

*Il peut prendre la voiture*, paraphrasable par *Il a l'autorisation de prendre la voiture*.

Si l'extra-prédicativité sémantique constitue un critère nécessaire, elle n'est pas un critère suffisant. Outre la valeur de sporadicité, certains des effets de sens discursifs de *pouvoir* ont une portée extra-prédicative aussi (Le Querler 1996 : 127-129, 2001 : 25-26).

Une deuxième caractéristique de la valeur épistémique qu'évoque Le Querler est que *pouvoir* épistémique se glose par *il se peut que* ou par des adverbes épistémiques tels que *probablement*, *peut-être* (Le Querler 1996). Ainsi, *elle peut avoir mal compris* se paraphrase par *peut-être qu'elle a mal compris*, *il se peut qu'elle ait mal compris*, etc. Cette caractéristique est en effet liée à la première étant donné que la formule *il se peut que* et les adverbes épistémiques sont toujours sémantiquement extra-prédicatifs. Bien qu'une expression comme *il se peut que* et des adverbes comme *probablement* et *peut-être* soient à considérer comme intrinsèquement épistémiques, la classification est rendue difficile par le fait que d'autres valeurs de *pouvoir* se laissent parfois paraphraser de la même manière (cf. exemple 5 ci-dessous), peut-être non sans un léger changement de sens, mais qui est parfois à peine saisissable.

Or, ce qui distingue la valeur épistémique de toute autre valeur (y compris la valeur de sporadicité) est qu'elle exprime *une possibilité présumée ou probable* (Faarlund et al. 2002 : 593), une possibilité imaginable qui n'est pas nécessairement valable. Cela nous amène au dernier critère, peut-être le plus important : il est possible d'ajouter, après un énoncé contenant *pouvoir* épistémique, une tournure comme *je n'en sais rien* ou *mais ce n'est pas sûr*.

Pour résumer, nous avons donc, outre les critères syntaxiques, deux critères sémantiques selon lesquels s'est opérée la classification de *pouvoir* épistémique :

- la portée de *pouvoir* est sémantiquement extra-prédicative, ce qui implique que l'énoncé se laisse gloser par *il se peut que* ou par *probablement*, *peut-être* etc.

- On peut ajouter à l'énoncé une tournure telle que *je n'en sais rien* ou *mais ce n'est pas sûr*.

Dans les cas où l'énoncé avec *pouvoir* remplit ces critères, *pouvoir* a été classifié comme épistémique.

Même si l'on peut s'appuyer sur des critères assez concrets comme ceux évoqués ci-dessus, il ne faut pas oublier le rôle du cotexte. Dans bien des cas, ce rôle est décisif quant à la détermination du sens précis de *pouvoir* / *kunne*. Par exemple, par son cotexte le plus immédiat, *peut* dans

*il peut s'agir d'infections bactériennes* (frmed07),

est susceptible de recevoir une interprétation épistémique (une glose par *peut-être* est possible, ainsi que l'adjonction d'une tournure telle que *je n'en sais rien*), mais si l'on prend en compte un cotexte plus large, on finit par attribuer à cette occurrence de *pouvoir* une valeur sporadique :

5. Chez d'autres, il **peut** s'agir d'infections bactériennes gastrointestinales à *Campylobacter* ou plus souvent parasitaires à *Giardia lamblia*... (frmed07)

Il est en effet possible de remplacer *peut* par *peut-être*, mais une glose par *parfois* est plus appropriée<sup>30</sup>. L'adjonction d'une tournure d'incertitude ne semble pas appropriée.

Il résulte de ce qui précède qu'il peut être particulièrement difficile de distinguer la valeur épistémique et la valeur sporadique. Cette difficulté réside dans le fait que la sporadicité est un intermédiaire entre l'aléthique et l'épistémique (cf. *supra* 3.1.2.3) et comme nous l'avons vu, ces deux dernières catégories ne sont pas non plus toujours facilement distinguables (cf. section 3.1). Compte tenu du manque de distinctions nettes entre les différentes catégories et les différentes valeurs, un certain degré d'arbitrarité et de subjectivité dans la classification est inévitable. Toutefois, les critères présentés ci-dessus permettent de minimiser autant que possible leur influence.

#### 7.4.4 La fréquence de *pouvoir* / *kunne* dans le corpus

Pour des raisons pratiques, la recherche des attestations de *pouvoir* / *kunne* dans le corpus a été restreinte aux formes du présent de l'indicatif. En ce qui concerne *pouvoir* nous avons

---

<sup>30</sup> De manière plus générale, on peut dire que *pouvoir* épistémique se glose par *peut-être* et *pouvoir* sporadique par *parfois*.

aussi effectué une analyse des formes du conditionnel, étant donné que le conditionnel de *pouvoir* s'est avéré le deuxième modalisateur épistémique le plus fréquent dans le corpus exploratoire.

Avant de présenter les fréquences de *pouvoir* / *kunne* épistémique, nous allons considérer les fréquences de ces verbes en général, toutes valeurs incluses. Cela est important afin de pouvoir déterminer dans quelle mesure ces auxiliaires modaux sont utilisés dans leur acception épistémique. Étant donné que *pouvoir* et *kunne* sont souvent considérés comme des marqueurs prototypiques de la possibilité épistémique (cela vaut du moins pour l'auxiliaire norvégien), cette question est importante et la réponse peut suggérer des hypothèses intéressantes sur les différences qui existent entre ces verbes généralement considérés comme équivalents.

Nous reviendrons plus loin aux résultats pour les formes conditionnelles. Regardons d'abord les résultats concernant l'emploi du présent. Ces résultats peuvent être représentés schématiquement comme suit<sup>31</sup> :

**Tableau 7** : L'emploi de *pouvoir* et de *kunne* au présent (toutes valeurs incluses)

	français			norvégien			total		
	mots	occ. <i>peut+</i>	f/1000	mots	occ. <i>kan</i>	f/1000	mots	occ. <i>peut+</i>	f/1000
linguistique	68727	360	5,2	90579	747	8,3	159306	1107	6,9
médecine	61180	158	2,6	43234	228	5,3	104414	386	3,7
total	129907	518	4,0	133813	975	7,3	263720	1493	5,7

Tout d'abord, il convient de noter que sur les 518 occurrences de *pouvoir* au présent, les formes à la troisième personne sont largement majoritaires avec 394 occurrences de *peut* et 111 occurrences de *peuvent*. Il y a 13 occurrences de *pouvons* (9 dans le corpus Frling et 4 dans le corpus Frmed), et aucune occurrence de *pouvez* ou de *peux*<sup>32</sup>. Sur les 975 occurrences de *kan* dans le corpus norvégien, 125 attestations apparaissent à la première personne (117 avec *vi* ('nous') et 8 avec *jeg* ('je')), et à l'exception de 2, toutes relèvent du corpus

<sup>31</sup> Nous utilisons la notation *peut+* pour désigner toutes les formes du présent de *pouvoir*, c'est-à-dire *peux*, *peut*, *pouvons*, *pouvez* et *peuvent*. Les verbes norvégiens ne se conjuguant pas en personne, la forme *kan* est la seule forme du présent de *kunne*. Elle s'utilise à toutes les personnes et remplace alors *peux*, *peut*, *pouvons*, *pouvez* et *peuvent* en français. L'ensemble des formes du conditionnel en français (*pourrais*, *pourrait*, *pourrions*, *pourriez*, *pourraient*) seront désignés par *pourrait+* (cf. tableau 9).

<sup>32</sup> L'absence des formes à la deuxième personne s'explique facilement par le fait que l'auteur d'un article de recherche ne s'adresse que très rarement aux destinataires d'une façon directe. L'absence de la séquence *je peux* et le peu d'occurrences de *nous pouvons* indiquent que ce modal est rarement employé avec la première personne, ce qui ne doit pas surprendre vu la fréquence généralement basse de pronoms personnels à la première personne dans les articles français (cf. Fløttum et al. 2006 : 73-81, 263-264).

linguistique. Les linguistes norvégiens utilisent donc ce verbe avec la première personne plus souvent que les autres groupes. Or, comme nous allons le voir, on ne retrouve pas la même observation en ce qui concerne les occurrences épistémiques.

Les chiffres obtenus montrent que *kan* est manifestement plus fréquent dans les articles norvégiens que *peut+* ne l'est dans les articles français. Dans l'ensemble, il y a 975 occurrences de *kan*, soit une fréquence relative de 7,3 pour mille, alors qu'il n'y a que 518 occurrences des formes correspondantes de *pouvoir*, soit une fréquence relative de 4,0. La première chose qui frappe est donc que l'emploi de *kan* par les chercheurs norvégiens dépasse presque deux fois l'emploi de *peut+* par les chercheurs français.

Le second trait frappant est la distribution inégale des formes *kan* et *peut+* entre les articles de linguistique et de médecine. *Kan* ou *peut+* s'emploient beaucoup plus fréquemment par les linguistes que par les médecins-chercheurs : comme il ressort du tableau 7, il y a 747 occurrences de *kan* dans le corpus Noling, soit une fréquence relative de 8,3 pour mille, contre 228 dans le corpus Nomed, soit une fréquence relative de 5,3 pour mille. Quant au corpus français, il y a 360 occurrences de *peut+* dans les articles de linguistique, soit une fréquence relative de 5,2 pour mille contre 158 dans les articles de médecine, soit une fréquence relative de 2,6 pour mille. La fréquence relative de *kan* et de *peut+* dans le corpus linguistique est donc de 6,9 pour mille, tandis qu'elle est de 3,7 pour mille dans les articles de médecine.

En résumé, nous pouvons donc conclure que la fréquence de *peut+* / *kan* varie beaucoup d'une discipline à l'autre et d'une langue à l'autre. Les chercheurs norvégiens emploient *kan* plus fréquemment que leurs collègues français n'emploient *peut+*, et les linguistes emploient *kan* / *peut+* dans une plus grande mesure que ne le font les médecins-chercheurs. Les variations interdisciplinaire et interlangagière semblent être aussi importantes l'une que l'autre. Il y a dans les deux langues une différence remarquable entre les disciplines, et réciproquement, il y a une différence remarquable entre les langues pour chaque discipline.

Passons maintenant à la fréquence de *pouvoir* / *kunne* épistémiques. La sélection des attestations s'est faite, comme nous l'avons vu, en fonction des critères présentés en 7.4.3. Le tableau 8 résume les résultats de la classification des formes *peut+* et *kan* en occurrences épistémiques et occurrences non-épistémiques. Les occurrences épistémiques apparaissent toutes à la troisième personne, à l'exception d'une occurrence épistémique à la première personne dans le corpus Noling (*jeg kan*, cf. exemple 46 ci-dessous).

**Tableau 8** : L'emploi de *pouvoir* / *kunne* épistémiques

	Français			Norvégien			Total		
	<i>peut+</i>	<i>peut+</i> ép.	%	<i>kan</i>	<i>kan</i> ép.	%	<i>peut+</i> / <i>kan</i>	occ.ép.	%
linguistique	360	4	1,1	747	62	8,3	1107	66	6
médecine	158	8	5,1	228	45	20	386	53	14
total	518	12	2,3	975	107	11	1493	119	8

Il ressort de ces résultats que l'emploi épistémique de *peut+* est très rare dans les articles de recherche. Cela vaut pour les deux disciplines, mais surtout pour la linguistique. Chez les linguistes, 4 des 360 occurrences ont été classifiées comme épistémiques, soit un pourcentage d'uniquement 1,1 % du nombre total d'occurrences de *peut+*. Chez les médecins-chercheurs, 8 des 158 occurrences ont été étiquetées épistémiques, ce qui donne un pourcentage de 5,1 %. Dans l'ensemble, 2,3 % seulement des occurrences de *peut+* ont une valeur épistémique. Voici quelques exemples :

6. Remarquons que certains phénomènes **peuvent** déjà être expliqués à la lueur de ce que nous avons dit précédemment. (frling13)
7. Enfin, la méthode simplifiée d'utiliser les intervalles de confiance à 95 % du SSC **peut** avoir occasionné une sous-estimation des intervalles de confiance réels. (frmed05)
8. Elle témoigne d'une vulnérabilité de la personnalité qui virtuellement **peut** être à l'origine de l'apparition de telles conduites (frmed13)

Si la valeur épistémique est très rare avec *peut+*, il en va autrement pour le verbe modal norvégien *kunne*. Environ 11 % des occurrences de *kan* expriment un sens épistémique. Dans les articles de linguistique, 8,3 % des occurrences de *kan* ont été étiquetées épistémiques, contre 20 % dans les articles de médecine. Exemples :

9. Parallellen til eksemplet i (14) **kan** tyde på at det er tilfelle (noling01)  
'Le parallèle de l'exemple (14) **peut** indiquer que cela est le cas'
10. Tar man på den annen side eddadiktenes og skaldekvadenes skildring av konkrete enkeltsituasjoner og personskildringer i betraktning, blir det forståelig at det der var lite behov for den slags generelle utsagn. Dette **kan** være grunnen til at de brukes så lite i slike dikt (noling08)  
'Si l'on prend en compte, par contre, la description que donnent la poésie des Eddas et les lais scaldiques des situations et personnes concrètes, on comprendra qu'il n'y avait de grands besoins d'énoncés généraux de ce type. Cela **peut** être la raison pour laquelle ils sont si peu utilisés dans de tels poèmes'

11. Forholdet mellom volum og kvalitet, som også er blitt påpekt tidligere for brokkirurgi (4), **kan** her være av betydning. (nomed05)  
'Le rapport entre quantité et qualité [...] **peut** avoir une importance.'
12. Hos en pasient med revmatoid artritt **kan** prednisolonbehandlingen ha forsinket diagnosen. (nomed03)  
'Chez un malade atteint d'arthrite rhumatoïdale, le traitement par la prednisolone **peut** avoir retardé le diagnostic.'

Outre la frappante variation interlangagière, il y a donc dans le corpus norvégien une différence interdisciplinaire considérable. Cette différence se retrouve dans le corpus français, mais de manière moins nette puisque le nombre d'occurrences de *pouvoir* épistémique est moins élevé que celui de *kunne* épistémique. Cependant, dans l'ensemble, les occurrences épistémiques sont beaucoup plus nombreuses dans les articles de médecine que dans les articles de linguistique.

Pour résumer, on remarque qu'au total, environ 8 % des occurrences de *kan* / *peut*+ revêtent un sens épistémique (119 sur 1493). La grande majorité de ces occurrences relèvent du corpus norvégien (107 occurrences), ce qui nous amène à conclure que l'emploi épistémique de *kan* en norvégien est beaucoup plus fréquent que le même emploi de *peut*+ en français. Les résultats indiquent aussi une variation interdisciplinaire importante. Dans les deux langues, ce sont les médecins-chercheurs qui se servent le plus de ces auxiliaires modaux dans leur valeur épistémique.

Comparée à l'emploi extensif de *kan* et de *may* épistémiques (cf. les tableaux 4 et 5, section 6.4.2), la quasi-absence de *peut*+ épistémique dans le corpus français est étonnante. La valeur épistémique de *pouvoir*, qui est le modal épistémique par excellence et dont les correspondants dans d'autres langues sont traditionnellement considérés comme faisant partie des modalisateurs épistémiques les plus importants, est-elle vraiment si rare ? Peut-être que la valeur épistémique de ce verbe est plus fréquente avec d'autres formes que le présent, par exemple avec le conditionnel, comme l'a suggéré l'analyse exploratoire ? Cette supposition semble fondée, étant donné que dans les formes conditionnelles de ce verbe se combinent deux modalisateurs épistémiques potentiels : l'un est évidemment le verbe modal *pouvoir* et l'autre est le tiroir du conditionnel, qui dans certains de ses emplois fonctionne comme un marqueur grammatical de la modalité épistémique<sup>33</sup>. Il y a donc un double marquage épistémique intrinsèque aux formes conditionnelles de *pouvoir* épistémique.

---

<sup>33</sup> Dans notre corpus, il semble y avoir peu d'exemples où le tiroir du conditionnel est l'unique marque de la modalité épistémique (cette observation se base sur l'étude du corpus exploratoire uniquement), mais le

Le tableau 9 résume les fréquences de *pourrait*+ dans le corpus<sup>34</sup>.

**Tableau 9** : L'emploi de *pourrait*+

	Nombre de mots	<i>Pourrait</i> +		<i>Pourrait</i> + ép.		% d'occ.ép.
		Occ.	fr.rel	Occ.	fr.rel	
linguistique	68727	52	0,76	4	0,06	7,7
médecine	61180	32	0,52	21	0,34	66
total	129907	84	0,65	25	0,19	30

Rappelons que le choix d'étudier le marqueur *pourrait* est motivée par la fréquence de *pourrait*+ épistémique dans le corpus exploratoire, et non par les critères de pertinence contrastive ou de caractère prototypique, comme c'était le cas pour *peut*+ (voir la section 6.3). Cela signale déjà que la fréquence relative de *pourrait*+ épistémique doit être plus élevée que celle de *peut*+ épistémique, et c'est effectivement le cas. Comme il ressort du tableau 9, il y a en tout 84<sup>35</sup> occurrences de *pourrait*+, dont 25 (soit un pourcentage de 30%) ont été classées comme épistémiques. Rappelons que pour le présent *peut*+, uniquement 2,3 % des occurrences ont été classées comme épistémiques. Cela aurait pu nous inciter à conclure que la valeur épistémique est liée au mode du conditionnel – ou du moins à des temps ou des modes autres que le présent – mais il est intéressant de noter que la question de temps ou de mode n'est pertinente que pour l'une des disciplines, à savoir la médecine. Pour les articles de linguistique, le changement de mode importe peu : la valeur épistémique de *pouvoir* n'est quasiment pas utilisée, ni avec le présent, ni avec le conditionnel. La fréquence relative de *pourrait*+ épistémique y est de 0,06 pour mille, ce qui correspond à la fréquence de *peut*+ épistémique (voir tableau 3, section 6.4.2). Pour les linguistes, la valeur épistémique du modal *pouvoir* n'est donc simplement pas une valeur pertinente. Il en va autrement pour les médecins-chercheurs. 66 % des occurrences de *pourrait*+ expriment une valeur épistémique, c'est-à-dire que la valeur épistémique est la première valeur de *pourrait*+ dans les articles de médecine. On observe donc un contraste net avec les articles de linguistique, où cette valeur ne constitue que 7,7 % des attestations. Pour les auteurs des articles de médecine, *pourrait*+ est le deuxième modalisateur épistémique, devancé uniquement par le semi-auxiliaire *sembler*

---

conditionnel sert cependant parfois à renforcer la valeur épistémique d'un verbe déjà épistémique, comme c'est le cas dans *pourrait*+ épistémique.

<sup>34</sup> Comme il a déjà été mentionné, les formes *pourrais*, *pourrait*, *pourrions*, *pourriez* et *pourraient* seront désormais dénotées par la forme de base *pourrait*+

<sup>35</sup> Sur ces 84 occurrences, il y en a 83 à la troisième personne (67 occurrences de *pourrait*, 16 de *pourraient*) et une à la première personne (*pourrions*). Autrement dit, il en va de même pour les formes au conditionnel que pour celles au présent : elles sont presque exclusivement utilisées à la troisième personne.

(voir tableau 3, 6.4.2). Dans les articles de médecine, la valeur épistémique est donc très fréquente avec les formes du conditionnel, alors qu'elle est rare avec les formes du présent. Les linguistes, de leur côté, ne mobilisent pas la valeur épistémique, y compris pour les formes au conditionnel.

L'emploi extensif de *pourrait*+ épistémique dans les articles de médecine, qui détermine la différence frappante entre disciplines, peut s'expliquer par une influence de l'anglais. Étant donné que l'anglais est la *lingua franca* de la recherche médicale, les médecins-chercheurs francophones lisent probablement beaucoup de littérature médicale en anglais. Or, comme il ressort du tableau 5, section 6.4.2, *may* et *might* sont des marqueurs importants de la modalité épistémique dans le discours médical anglais. L'usage épistémique de ces verbes totalise respectivement 73 % (76 sur 104 occurrences) et 85 % (22 sur 26 occurrences) des occurrences. Le pourcentage est donc encore plus élevé que pour *pourrait*+ dans les articles français, ce qui ne doit pas surprendre, vu que l'anglais dispose du verbe *can* comme une alternative non-épistémique de *may*, alors qu'une telle alternative n'existe pas en français. Cependant, que ce soit en anglais ou en français, ces chiffres suggèrent que dans le discours médical, ces auxiliaires modaux s'utilisent avant tout comme des marqueurs épistémico-modaux, tandis qu'en linguistique, leur usage est beaucoup plus varié. À cet égard, les médecins-chercheurs francophones adoptent une manière de s'exprimer qui ressemble plus à celle de leurs collègues anglophones qu'à celle de linguistes francophones, et le facteur disciplinaire l'emporte alors sur le facteur de la langue.

#### **7.4.5 Le cotexte immédiat**

Cette section traitera principalement du cotexte immédiat des occurrences *épistémiques*. Or, comme les linguistes francophones n'utilisent quasiment pas *pouvoir* dans son acception épistémique, il pourrait être intéressant d'examiner les occurrences non-épistémiques aussi, notamment celles de *pourrait*+ afin de pouvoir décrire d'une façon un peu plus précise les différences d'emploi de ce marqueur qui se sont manifestées entre les deux groupes (linguistes et médecins-chercheurs). Un bref examen de ces occurrences révèle que les linguistes emploient *pourrait*+ surtout avec le pronom *on*. En fait, plus de la moitié des occurrences de *pourrait*+ se combine avec *on* pronom sujet. Cela les distingue des chercheurs en médecine qui en général ont un emploi relativement limité de *on* (Fløttum et al. 2006)), et dans les articles de médecine, on ne trouve qu'un seul exemple où *pourrait* a le pronom *on* comme sujet. Chez les linguistes, en revanche, la combinaison *on pourrait* semble constituer une séquence quasiment figée. Le contenu sémantique de *pourrait* semble alors relativement



« pauvre » : il est difficile d'en préciser le sens à l'aide d'une valeur déterminée. Il n'est pas épistémique, mais la séquence *on pourrait* + verbe principal peut tout de même avoir un effet atténuant, dans la mesure où elle sert à diluer l'identité de l'instance responsable du point de vue transmis (voir aussi Grossmann et Rinck 2004 : 41). En ayant recours au pronom *on*, l'auteur peut donc présenter un point de vue envisageable, sans transmettre ses propres jugements concernant la validité du point de vue évoqué :

13. Ainsi, toute une série de contraintes syntaxiques semblent liées à ce qu'on pourrait appeler une "allergie" à la préposition. (frling04)<sup>36</sup>
14. On pourrait objecter que cela est dû à des facteurs extra-linguistiques et non syntaxiques. (frling06)

Alors que les médecins-chercheurs font preuve d'un emploi extensif de *pourrait* épistémique, les linguistes font donc preuve d'un emploi extensif de la séquence *on pourrait* + verbe principal, séquence qui est susceptible d'avoir un effet atténuant ou déresponsabilisant sur le point de vue transmis.

Passons maintenant aux occurrences épistémiques de *peut+*, *pourrait+* et *kan* et aux éléments qui constituent leur cotexte immédiat. Compte tenu de la basse fréquence de *peut+* épistémique dans le corpus, nous traiterons ensemble *peut+* et *pourrait+*. Nous commencerons par quelques observations sur les types de verbes et de sujets entourant les occurrences épistémiques, et examinerons ensuite dans quelle mesure *pouvoir* et *kunne* épistémiques sont combinés avec d'autres modalisateurs épistémiques.

Comme c'était le cas avec *synes* et *se ut (til/som)*, c'est *være* qui est le verbe principal se combinant le plus souvent avec *kan* épistémique. Dans le corpus linguistique, un tiers des occurrences épistémiques de *kan* modifient *være*. Dans les articles de médecine, on en dénombre environ 40 %. Dans le corpus français, *peut* / *pourrait+* épistémiques se combinent presque autant de fois avec le verbe *s'expliquer* / *expliquer* qu'avec *être*, ce qui souligne le rôle que joue ce marqueur dans l'expression d'explications possibles (voir ci-dessous). Le cotexte dans le corpus norvégien révèle aussi l'importance de ce rôle : un examen des sujets combinés avec *kan* ainsi que des attributs attachés à *være* comme verbe principal montre que ceux-ci se constituent très souvent de syntagmes comportant les lexèmes *forklaring* ('explication'), *årsak* ('cause') et *grunn* ('raison'), et cela vaut pour les deux disciplines<sup>37</sup> :

---

<sup>36</sup> Dans cet exemple, les guillemets soulignent le fait que l'auteur ne souscrit pas entièrement à l'appellation qu'il vient de proposer.

<sup>37</sup> Il convient cependant de noter que dans le corpus de linguistique, un grand nombre de ces occurrences relèvent d'un même article (noling17).

15. Hvis DENNE ikke kan ha subjektet som antesedent, **kan en mulig forklaring** ligge i den tematiske rollen til subjektet. (noling01)

'Si DENNE ne peut pas avoir le sujet comme antécédent, **une explication possible peut** résider dans le rôle thématique du sujet.'

16. Dette kan indikere at oksygen hemmer respirasjonen, og det **kan være årsaken til** at apgarskåre ved ett minutt var lavere i oksygengruppen enn i romluftgruppen. (nomed06)

'Cela peut indiquer que l'oxygène entrave la respiration, et cela **peut être la cause** de ce que le score Apgar à 1 minute était plus bas dans le groupe « oxygène pur » que dans le groupe « air ».'

Par ailleurs, on peut remarquer que *pouvoir / kunne* épistémiques se combinent rarement avec un sujet animé, quoique cela arrive plus souvent dans les articles médicaux que dans les articles linguistiques. Le sujet réfère dans ces cas à un groupe de malades et l'auxiliaire sert à exprimer une hypothèse ou une supposition concernant ce groupe :

17. Årsaken til fenomenet er ikke kjent, men tallene indikerer at **kvinner kan ha** andre strategier for sprøytedeling enn menn. (nomed08)

'La cause du phénomène n'est pas connue, mais les chiffres indiquent que **les femmes peuvent avoir** d'autres stratégies que les hommes pour partager des seringues.'

*Være* est donc en définitive le verbe principal le plus récurrent avec *kan* épistémique. Dans les cas où *kan* modifie un verbe autre que *være*, il s'agit relativement souvent d'un verbe épistémique comme *tyde på*, *synes*, *virke*. Cette observation, qui vaut pour les articles de linguistique et dans une moindre mesure pour les articles de médecine, suggère que *kan* se combine souvent avec d'autres modalisateurs épistémiques, ce qui crée une accumulation d'atténuateurs qui se rapportent au même contenu informatif. Dans les articles de linguistique, *kan* épistémique se combine avec d'autres modalisateurs épistémiques dans environ 42 % des cas. Les modalisateurs les plus récurrents sont *se ut (til/som)*, *mulig* et *tyde på*. Le taux est à peu près le même dans les articles de médecine : 40 % des occurrences de *kan* épistémique apparaissent avec d'autres modalisateurs épistémiques, les plus récurrents étant *mulig*, *tyde på* et *indikere*. Dans beaucoup de cas, c'est le marqueur grammatical *kan* qui souligne et renforce l'incertitude exprimée par le marqueur lexical. *Kan* est donc facultatif d'un point de vue syntaxique, et son rôle est d'atténuer plus encore l'énoncé :

18. Frekvenstallene spriker sterkt for de ulike lokalisasjoner, noe som **kan tyde på** metodologiske vanskeligheter. (nomed19)

'Les chiffres diffèrent considérablement d'une localisation à l'autre, ce qui **peut indiquer** des difficultés méthodologiques.'

19. Dette **kan indikere** at det en viss læringskurve for scintigrafiundersøkelsen.(nomed02)  
'Cela **peut indiquer** qu'il y a une certaine courbe d'apprentissage pour l'examen de scintigraphie.'

20. En kooperativ fullføring blir introdusert utenfor et potensielt turskiftepunkt og **kan** dermed **se ut til** å bryte med turtakingsreglene.(noling15)  
'Un accomplissement coopératif est introduit en dehors d'un moment potentiel pour un échange de tour de parole et, par conséquent, **peut sembler** rompre avec les règles de la prise de tour de paroles.'

Dans d'autres cas, le marqueur grammatical (*kan*) et le marqueur lexical se renforcent mutuellement, et d'un point de vue syntaxique, l'un ou l'autre suffit. Parfois le marqueur lexical sert à mettre en relief la valeur épistémique de l'auxiliaire, qui, sans le support du marqueur lexical, apparaîtrait peut-être moins univoque et se prêterait plus facilement à d'autres lectures (cf. l'exemple 21):

21. Derfor **kan** preoperativ scintigrafi **muligens** føre til redusert antall reoperasjoner (13) (nomed02).  
'Pour cette raison, une scintigraphie préopératoire **peut peut-être** mener à une réduction du nombre de réopérations (13).'

22. **Det er mulig at** pasienter med denne problematikken tidligere **kan** ha vært oversett når striktoren har vært av moderat grad.(nomed11)  
'**Il est possible qu'**antérieurement, des malades atteints de cette problématique **puissent** avoir échappé à l'attention lorsque la striction a été d'un degré modéré.'

Dans le corpus français, l'accumulation de modalisateurs épistémiques n'est pas aussi fréquente, mais on en trouve néanmoins quelques exemples. Dans les articles de linguistique, *peut+* et *pourrait+* épistémiques sont combinés avec d'autres marqueurs épistémico-modaux à 3 reprises. Dans les articles de médecine, il y a 5 exemples où *peut+* ou *pourrait+* apparaissent avec d'autres marqueurs épistémiques.

23. Ce choix d'un corpus poétique procède, dans une certaine mesure, du hasard (en ce sens que nous avons exploré ledit corpus pour d'autres études). Mais nous **pensons** qu'il **pourrait** bien se révéler crucial pour le genre d'enquête que nous voulons mener (frling17)

24. Des analyses effectuées sur la cohorte APROCO ont **suggéré** que des difficultés d'adhésion au traitement **pourraient** expliquer cette discordance entre réponse immunologique et réponse virologique chez certains malades [9].(frmed10)

L'exemple 23 est un exemple où l'autre modalisateur sert à expliciter la source de l'évaluation épistémique exprimée (*nous pensons*). En vue de déterminer dans quelle mesure la source de l'évaluation exprimée par *kunne* / *pouvoir* épistémiques est rendue explicite par

les choix lexicaux de l'auteur, nous avons, en examinant le cotexte de *kunne / pouvoir* épistémiques, cherché aussi les compléments adverbiaux d'opinion tels que *selon nous, j'en mener* etc., outre les modalisateurs épistémiques à proprement parler. Or, il s'avère que la source est très rarement explicité, ce qui ne doit peut-être pas surprendre, car l'interprétation par défaut est naturellement que c'est l'auteur lui-même qui est responsable de l'évaluation, de sorte qu'il est inutile de le préciser. Dans les très rares cas où l'auteur ou les auteurs eux-mêmes sont explicitement indiqués comme étant la source de l'évaluation épistémique, cela est fait à l'aide d'un autre modalisateur épistémique (voir l'exemple 23 ci-dessus et l'exemple 37 ci-dessous).

#### 7.4.6 Fonctions pragmatiques de *pouvoir / kunne* épistémiques

Passons maintenant au fonctionnement pragmatique de *pouvoir / kunne* épistémiques. Soulignons tout d'abord et encore une fois qu'en pratique, les différents emplois se chevauchent souvent. Ils constituent des catégories théoriques, mais dans les exemples authentiques, il n'y a pas de frontières nettes entre elles. Nous verrons plusieurs exemples de ces chevauchements dans ce qui suit.

Dans les articles de médecine, *pouvoir* épistémique s'emploie surtout pour signaler des limites méthodologiques (voir l'exemple 25) – c'est-à-dire pour évoquer d'éventuelles sources d'erreur – et pour avancer des hypothèses (voir l'exemple 26) ou pour proposer des explications possibles (voir l'exemple 27). Autrement dit, *pouvoir* épistémique s'utilise dans la grande majorité des cas avec une fonction contenu :

25. Notre étude comporte un certain nombre de limites. L'auto-déclaration ou la non-déclaration **peuvent** avoir biaisé les résultats. L'utilisation de données auto-déclarées **peut** avoir faussé les estimations relatives à la clientèle des programmes de dépistage (frmed05)
26. La diminution des surdoses mortelles **pourrait** être liée à une intégration active des usagers de drogues dans le réseau de soins. (frmed19)
27. Nous n'avons pas observé de différence d'âge entre ST et SP comme d'autres auteurs [10]; ceci **pourrait** s'expliquer par l'âge jeune de notre population. (frmed15)

L'emploi qui consiste à proposer des explications possibles est particulièrement fréquent avec *pourrait*. L'explication proposée peut être l'unique explication évoquée, comme dans l'exemple 28:

28. On observe une interaction des variables de l'âge et du sexe en ce qui concerne le séjour moyen : à partir de l'âge de 5 ans, le séjour moyen des filles et des femmes est plus long que celui des garçons et des hommes, alors qu'il n'y a guère de différence entre les deux groupes avant cet âge. Cette variation **pourrait** s'expliquer par un lien entre le séjour moyen à l'hôpital et la présence d'une comorbidité, plus fréquente chez les femmes plus âgées. (frmed02)

ou elle peut constituer une alternative parmi une série d'explications envisagées (exemples 29-30) :

29. Les mécanismes à l'origine de ces manifestations sont peu clairs : il **pourrait** s'agir d'anomalies des cellules T helper [...]; il **pourrait** également s'agir d'une conséquence d'anomalies génétiques du chromosome 6... (frmed07)

Dans l'exemple 30 ci-dessous, les auteurs cherchent à expliquer l'observation selon laquelle « les hospitalisations pour l'asthme sont plus nombreuses à l'automne et moins nombreuses à l'été ». Plusieurs explications possibles sont évoquées (marquées dans l'exemple par des italiques) ; la première est présentée comme plus difficilement contestable que les suivantes, qui sont présentées comme plus hypothétiques.

30. Tel que mentionné précédemment [...] les hospitalisations pour l'asthme sont plus nombreuses à l'automne et moins nombreuses à l'été. Cette baisse pendant la saison estivale *s'explique en partie par les congés et les fermetures de lits d'hôpitaux causées par la réduction du personnel de santé.* La hausse des hospitalisations à l'automne **peut être associée à l'augmentation des infections virales qui accompagne la rentrée scolaire [...] et à la saison des allergies à l'herbe à poux (de la mi-août à la fin septembre), [...]** car les personnes asthmatiques peuvent tolérer la réaction pendant un certain temps avant de nécessiter un séjour à l'hôpital. *Une autre cause possible est l'augmentation du temps passé à l'intérieur, ce qui augmente l'exposition aux allergènes tels que les acariens et les animaux domestiques.* (frmed02, mes italiques)

L'emploi de *kan* dans les articles médicaux norvégiens est tout à fait comparable : les fonctions les plus récurrentes sont celles de proposer des explications possibles (exemples 31-34), signaler des limites méthodologiques (exemples 34-35) et émettre des hypothèses (exemples 36-37) :

31. Intakt serotoninfunksjon i hypothalamus spiller en viktig rolle for metthetsfølelsen når man spiser, og det er mulig at risperidon, som en følge av dets blokade av 5-HT<sub>2A</sub>-reseptorer, uønsket **kan** ha skrudd av denne metthetsbryteren. Alternativt **kan** D<sub>2</sub>-reseptorblokade ha spilt en viktig rolle (17, 18). (nomed04)  
'Le sentiment d'être rassasié dépend de la fonction intacte de la sérotonine dans l'hypothalamus, et il est possible que la rispéridone, suite à son blocage des récepteurs 5-HT<sub>2A</sub>, **puisse** avoir éteint ce commutateur de satiété (17,18).'

32. Årsaken til den høye letaliteten blant pasienter som ble operert med Billroth II eller Billroth I (fire av seks pasienter døde), er ikke kjent. En sykere pasientgruppe med høyere alder, lengre operasjonstid og operatører med mindre trening i ventrikkelkirurgi **kan** være mulige forklaringer. (nomed20)  
 'La cause du taux de léthalité élevé chez les malades opérés selon Billroth II ou Billroth I (quatre sur six malades sont morts), n'est pas connue. Un groupe de malades plus malades et plus âgés, une durée d'opération plus longue et des opérateurs moins exercés dans la chirurgie ventriculaire **peuvent** être des explications possibles.'
33. Uoverensstemmelsene mellom studiene kan forklares med det lave antall pasienter som er undersøkt, samt den relativt korte oppfølgingstiden. En annen grunn **kan** være at den histopatologiske diagnosen av en borderlinetumor er vanskelig. (nomed18)  
 'Les désaccords entre les études s'expliquent par le nombre limité de malades examinés ainsi que par un suivi relativement court. Une autre raison **peut** être que le diagnostic histopathologique d'une tumeur borderline est délicat.'
34. I vår undersøkelse er det lite samsvar mellom patologisk refluksindeks og gjentatte apnéanfall (sju av 21). En forklaring **kan** være at utvelgelsen av barna **kan** ha vært noe liberal. (nomed07)  
 'Dans notre étude, il y a peu de concordance entre un index de reflux pathologique et des attaques d'apnée récidivantes (sept sur 21). Une explication **peut** être que la sélection des enfants **peut** avoir été un peu souple.'
35. Før 1970 var sannsynligvis en del borderlinetumorer rapportert som høyt differensierte karsinomer, og forbedringen i prognosen fra 1950-årene **kan** derfor være noe undervurdert her. (nomed18)  
 'Avant 1970, une partie des tumeurs borderlines ont probablement été enregistrées comme des carcinomes hautement différenciés, et pour cela, l'amélioration du pronostic des années 1950 **peut** être un peu sous-estimée ici.'
36. Dette **kan** tyde på at risikoen for å utvikle strikturer på grunn av udiagnostisert eller underbehandlet refluks, henholdsvis øsofagitt er større i denne gruppen. (nomed11)  
 'Cela **peut** indiquer que le risque de développer des strictions à cause d'un reflux non diagnostiqué ou insuffisamment traité, respectivement l'œsophagite, est plus élevé dans ce groupe.'
37. Vi antar at en femte kjertel **kan** være årsak til hennes moderate hyperkalsemi, men pasienten er ennå ikke reoperert. (nomed16)  
 'Nous supposons qu'une cinquième glande **peut** être la cause de son hypercalcémie modérée, mais la patiente n'est pas encore réopérée.'

On observe aussi que dans les articles médicaux, *pourrait+* ou *kan* épistémiques sont dans bien des cas précédés ou suivis d'une remarque sur les connaissances limitées de la médecine (remarque mise en italiques gras dans les exemples ci-dessous. Voir aussi les exemples 17, 29 et 32). On signale donc une lacune dans les connaissances de la science médicale en général,

ce qui montre que les modalisateurs épistémiques utilisés dénotent dans ces cas une incertitude réelle (cf. *supra* 6.3) :

38. **Le mécanisme à l'origine du développement de la granulomatosose n'est pas connu.** Dans une étude portant sur 17 malades, 16 présentaient des proliférations anormales des cellules T en présence de différents mitogènes [3]. Une dysrégulation de la fonction des cellules T ou de l'activation des macrophages **pourraient** être impliquées dans la formation des granulomes [3]. (frmed07)
39. **Den patofysiologiske mekanismen er uavklart**, samtidig som det ikke er påvist noen tidsmessig sammenheng mellom refluksepisoder og apnéanfall (11, 12). Hos enkelte pasienter finnes imidlertid klart sammenfall i tid, og mekanismen **kan** være en obstruktiv apné med larynxspasme sekundært til reflux (13). (nomed07)  
'**Le mécanisme physiopathologique n'en est pas connu**, [...] le mécanisme **peut** être une apnée obstructive avec spasmes laryngés secondaire au reflux (13).

Outre ces emplois qui illustrent la fonction contenu des modalisateurs épistémiques, on trouve dans les articles médicaux quelques rares exemples où *kunne / pouvoir* entrent dans un passage plus argumentatif ou dialogique/polémique et où les auteurs cherchent à aller au devant d'éventuelles remarques potentiellement critiques. C'est ce que l'on observe dans les exemples 40 et 41 ; nous avons là un exemple de la fonction interpersonnelle, qui illustre l'aspect interactif aussi bien que l'aspect persuasif : les auteurs s'imaginent des réactions potentielles de la part des lecteurs (*det kan også hevdes at, le coût annuel [...] peut sembler élevé*) et ils y répondent par des subordonnées introduites par le connecteur concessif *mais* (*men i denne studien..., mais cette estimation...*) :

40. En mulig forklaring på noe kortere behandlingstid kan være at pasienter behandlet i allmennpraksis har et mildere sykdomsforløp enn sykehuspopulasjoner. Det kan også hevdes at dette **kan** skyldes overdiagnostisering, men i denne studien er feildiagnoser og usikre diagnoser forsøkt luket ut før inklusjon i behandlingsdelen. (nomed03)  
'Une explication possible de la durée plus courte du traitement **peut** être que les patients traités par les généralistes ont une évolution plus bénigne de la maladie que les populations hospitalisées. On peut aussi maintenir que cela **peut** être dû à un surdiagnostic, mais dans la présente étude, on a essayé d'exclure les diagnostics erronés et les diagnostics incertains avant inclusion dans le traitement.'
41. Le coût annuel des hospitalisations pour l'asthme, chiffré entre 18 et 21 millions de dollars, **peut** sembler élevé, mais cette estimation est sans doute de beaucoup inférieure au coût réel, étant donné que l'analyse ne tient pas compte des consultations d'urgence ni du remboursement des médecins. (frmed02)

Chez les linguistes, on retrouve dans une large mesure les mêmes emplois. Celui qui consiste à proposer des explications possibles est particulièrement fréquent, mais il convient de noter

qu'environ la moitié des exemples relèvent d'un même article (l'article Noling17). Dans l'extrait de Noling17 ci-dessous, il y a quatre occurrences de *kan* qui remplissent cette fonction. Ce passage abonde d'ailleurs en modalisateurs épistémiques : outre les occurrences épistémiques de *kan*, on y trouve *sannsynligvis* ('probablement'), *lite sannsynlig* ('peu probable'), *det er f.eks. mulig at...* ('il est par exemple possible que...').

42. Det er sannsynligvis flere årsaker til at noen pasienter er upresise og vage når de svarer på spørsmål om røyking. En viktig årsak **kan** være, som jeg har diskutert ovenfor, at leger som stiller upresise og åpne spørsmål om dette sensitive temaet, åpner for at pasienter kan gi upresise svar. En annen årsak **kan** være at pasientene ikke husker hvor mye, hvor lenge osv. de har røykt. Dette finner jeg lite sannsynlig. En tredje årsak **kan** være at noen pasienter og leger har forskjellig forståelse av hva det vil si å være røyker. Det er f.eks. mulig at noen røykere ikke regner seg som røykere dersom de bare røyker av og til (jf. Foss et al 1998). En fjerde årsak **kan** være at pasientene ikke vil svare presist.(noling17)  
'Il existe probablement plusieurs causes aux réponses vagues et imprécises que donnent certains malades à des questions concernant leur habitude de fumer. Une cause importante **peut** être [...] que les questions imprécises et ouvertes que posent certains médecins sur ce sujet délicat amènent à des réponses imprécises de la part du malade. Une deuxième cause **peut** être que les malades ne se souviennent pas du nombre de cigarettes fumées, du nombre d'années pendant lesquelles ils fument etc. Je trouve cela peu probable. Une troisième cause **peut** être que certains malades et médecins ont des points de vue divergents sur ce qu'est un fumeur. Il est par exemple possible que certains fumeurs ne se considèrent pas comme tels s'ils ne fument que de temps en temps (cf. Foss et al. 1998). Une quatrième cause **peut** être que les malades ne veulent pas répondre d'une façon précise.'

On trouve également plusieurs exemples des deux autres emplois : celui consistant à avancer des hypothèses (voir exemple 43 ci-dessous) et celui consistant à signaler des limites méthodologiques (exemples 44-46). Pour ce dernier emploi, il convient cependant de noter que la plupart des exemples relèvent de Noling07, qui est un article expérimental. On peut donc supposer que cet emploi est une caractéristique du sous-genre de l'article expérimental plutôt que de l'article de recherche en général ou de l'article médical en particulier.

Dans l'exemple 43, l'auteur présente d'abord une hypothèse à laquelle il ne souscrit pas. Après avoir argumenté contre cette hypothèse, il présente sa propre hypothèse. On voit donc qu'un emploi qui illustre en principe la fonction contenu (le fait d'avancer des hypothèses) entre dans un cotexte qui illustre les aspects persuasif et interactif de l'article scientifique :

43. Selv om suffikset *-gi* [...] skiller seg fra partiklene *-a* og *-at* når det gjelder hvilket setningsledd de negerer, kan man kanskje tenke seg at selve overføringen av den negerende betydning fra *ne* til partikkelen *-gi* **kan** ha skjedd på omtrent samme måte som ved *-a* og *-at*. [...] Det er imidlertid påfallende at det ikke fins et eneste belegg med bevart



*ne* pluss *-gi* i rent negerende betydning, mens overgangstypen *ne* pluss *-a*, *-at* er sikkert belagt. Det **kan** tyde på en prinsipiell forskjell mellom de to negasjonstypene [...] (noliong08)

‘Bien que le suffixe *-gi* [...] se distingue des particules *-a* et *-at* quant au complément de phrase affecté par la négation, on peut peut-être imaginer que la transmission même du sens négatif de *ne* à la particule *-gi* **peut** s’être déroulée d’à peu près la même manière que celle de *-a* et *-at*. [...] Il est cependant remarquable qu’il n’existe pas un seul exemple contenant *ne* et *-gi* dans un sens purement négatif, alors qu’il existe des exemples confirmés de la forme de transition *ne* et *-a*, *-at*. Cela **peut** indiquer qu’il y a une différence fondamentale des deux types de négation.’

On peut trouver la même coexistence d’aspects dans des exemples où sont évoquées des limites méthodologiques. Après avoir signalé les limites éventuelles de l’étude, l’auteur cherchera souvent à minimiser leur importance, afin de se défendre contre la critique potentielle que ces limites peuvent impliquer. On glisse donc d’une fonction contenu qui consiste à signaler des limites méthodologiques vers un emploi interpersonnel qui consiste à anticiper des commentaires critiques. Ainsi en est-il de l’exemple ci-dessous où l’auteur, après avoir évoqué une éventuelle source d’erreur, conclut (suite à une argumentation) qu’elle a sans doute été finalement sans conséquence. On pourrait donc dire que la fonction pragmatique est d’anticiper – et secondairement se défendre contre – une critique, aussi bien que d’évoquer les limites de l’étude. *Kan*, avec l’adjectif *mulig* (‘possible’), signalent une possible limite méthodologique et anticipent une potentielle objection. *Skulle* (‘devrait’) est ici employé avec une valeur épistémico-modale et assume une fonction persuasive : l’auteur se défend contre l’objection envisagée :

44. En mulig feilkilde i teksttesten er at informanten **kan** ha misforstått faktorer i tekstene eller i komparasjonssetningen. Men siden Juan har nådd et relativt avansert nivå i norsk, og siden det er lagt vinn på å konstruere entydige setninger med det vi antar er kjente leksemer og med enkel syntaks (Uri 1997a), skulle ikke dette representere noe stort problem. (noling20)

‘Une possible source d’erreur dans le test textuel est que l’informateur **peut** avoir mal compris certains éléments dans les textes ou dans la phrase comparative. Mais comme Juan a atteint un niveau relativement avancé en norvégien, et comme on s’est efforcé de construire des phrases univoques avec une syntaxe simple et avec des lexèmes supposés connus (Uri 1997a), cela ne devrait pas poser un grand problème.’

Les exemples 45 et 46 illustrent le même phénomène :

45. Enkelte feil **kan** skyldes utvalget av verb i testen. Ettersom testen fokuserer på form **kan** den gi flere feil og feil av bestemte typer. Det **kan** f.eks. være slik at de verbene fra den store svake klassen som rimer på sterke verb, også har en fonologi som passer med den lille svake klassen. På den annen side var det ikke bare verb som rimte med sterke, som

ble overgeneralisert fra V-klassen til K-klassen. I ethvert fall mener vi at det feilmønsteret vi får i en slik test, kan si noe interessant om hvordan prosesseringen foregår. (noling07)  
'Certaines des erreurs **peuvent** être causées par la sélection de verbes dans le test. Etant donné que le test met l'accent sur les formes, il peut susciter un plus grand nombre d'erreurs et des erreurs d'un certain type. Par exemple, ceux d'entre les verbes de la grande classe de verbes réguliers qui riment avec des verbes irréguliers peuvent avoir en même temps une phonologie qui concorde avec la petite classe de verbes réguliers. De l'autre côté, il n'y avait pas seulement les verbes qui rimaient avec des verbes irréguliers qui ont été surgénéralisés de la classe V à la classe K. En tout cas, nous sommes de l'avis que les tendances d'erreurs qui se manifestent dans un tel test peuvent nous dire quelque chose d'intéressant sur la manière dont le traitement se déroule.'

46. Det er også mulig at jeg **kan** ha oversett enkelte belegg i denne store stoffmengde. Allikevel tror jeg at den bearbejdede tabell III gir et nokså riktig bilde [...] (noling08)  
'Il est également possible que je **puisse** avoir laissé échapper certains exemples dans cette grande quantité de données. Tout de même, je crois que la version remaniée du tableau III fournit une image assez correcte [...]

Dans ces exemples, on voit comment les trois aspects de l'article de recherche peuvent être co-présents dans le même passage, et comment les différentes fonctions pragmatiques des modalisateurs épistémiques s'entremêlent.

L'anticipation d'éventuelles remarques critiques, emploi illustrant l'aspect dialogique et interactif de l'article de recherche, est plus récurrent chez les linguistes que chez les médecins-chercheurs, mais reste pourtant faible par rapport à d'autres emplois, comme celui consistant à proposer des explications possibles. Un autre emploi dont on trouve quelques exemples chez les linguistes, mais qui néanmoins reste assez restreint, est celui consistant à faire preuve de prudence en interprétant les idées ou le travail d'autres chercheurs :

47. Rischels problem **kan** synes å være av rent terminologisk karakter.(noling10)  
'Le problème de Rischel **peut** sembler être d'un caractère purement terminologique.'

En conclusion, on pourrait donc dire que *pouvoir* et *kunne* épistémiques s'emploient principalement pour proposer des explications possibles et pour avancer des hypothèses. Dans les articles médicaux, ils s'utilisent en outre très souvent pour signaler des limites méthodologiques. Cet emploi est également présent, mais dans une bien moindre mesure, dans les articles de linguistique. Cette différence disciplinaire peut s'expliquer par le fait que l'emploi consistant à signaler des limites méthodologiques est typique des articles expérimentaux et ceux-ci sont peu nombreux dans le corpus de linguistique alors que la plupart des articles médicaux sont expérimentaux. L'emploi qui consiste à anticiper des remarques critiques ainsi que celui qui consiste à faire preuve de prudence en se prononçant

sur les idées d'autres chercheurs sont assez faiblement représentés par rapport aux autres emplois mentionnés, mais on en trouve quelques exemples, principalement dans les articles de linguistique.

#### 7.4.7 Bilan

Dans ce qui précède, nous avons vu qu'il existe des différences importantes entre les deux disciplines ainsi qu'entre les deux langues dans l'emploi des auxiliaires modaux *pouvoir / kunne*. Pour ce qui est de l'aspect quantitatif, il a été observé que *kan* et *peut+* s'utilisent plus souvent par les linguistes que par les médecins-chercheurs. Au niveau des langues, *kan* s'emploie plus dans le corpus norvégien que *peut+* dans le corpus français. Or, ces chiffres, qui valent pour l'emploi de *kan / peut+* en général, ne se vérifient pas au niveau de l'emploi épistémique de ces auxiliaires modaux. Il s'avère que la valeur épistémique est beaucoup plus fréquente dans les articles de médecine que dans les articles de linguistique. Cela vaut pour les deux langues, mais la différence est particulièrement nette dans le corpus français, étant donné que les linguistes francophones n'utilisent quasiment pas cet auxiliaire dans son acception épistémique. Quant à la différence entre langues, *kan* épistémique est beaucoup plus récurrent que *peut+* épistémique, mais *pourrait+* épistémique apparaît fréquemment dans la partie médicale du corpus français. La haute fréquence de *pourrait+* épistémique dans le corpus médical s'explique probablement par une influence de l'anglais.

Un examen du cotexte immédiat des attestations épistémiques de *kan / peut+ / pourrait+* a montré que la combinaison avec d'autres marqueurs épistémico-modaux est très fréquente dans le corpus norvégien, mais plus rare dans le corpus français. Cette observation corrobore donc celle de Breivega (2003), selon laquelle *kan* épistémique apparaît très souvent avec des marqueurs épistémiques lexicaux. Bien que la combinaison de marqueurs épistémiques ne soit pas très fréquente dans le corpus français, il ne faut pas oublier que dans les formes du conditionnel (*pourrait+*), deux modalisateurs épistémiques coexistent, à savoir l'auxiliaire modal propre et le tiroir du conditionnel.

Le cotexte indique également que *pouvoir / kunne* épistémiques s'utilisent souvent pour signaler des explications possibles, et cette observation est confirmée par l'analyse des fonctions pragmatiques ; cet usage est fréquent dans les deux disciplines et dans les deux langues. Un autre usage récurrent dans tous les sous-corpus est la présentation d'hypothèses. Ces deux usages coïncident certes souvent. Chez les médecins-chercheurs, les explications ou les hypothèses qualifiées à l'aide de *pouvoir / kunne* épistémiques sont en outre souvent suivies ou précédées d'une remarque sur les connaissances limitées de la médecine en

général, ce qui souligne qu'on a affaire à une véritable incertitude. Outre ces deux usages, *pouvoir* / *kunne* épistémique s'utilisent très souvent pour signaler des limites méthodologiques, mais cet emploi se restreint dans une large mesure aux articles médicaux. Il y a là une nette différence entre les deux disciplines, mais cette différence relève probablement plutôt d'une différence entre différents sous-genres que d'une différence effective entre les deux disciplines investiguées. Le fait de signaler des limites méthodologiques semble être une caractéristique des articles expérimentaux, et comme ce type d'article est plus fréquent en médecine qu'en linguistique, la différence de genre a des répercussions sur les différences entre disciplines. Enfin, nous avons observé que les fonctions interpersonnelles (anticiper des remarques critiques, faire preuve de prudence en se prononçant sur les idées d'autres chercheurs) sont plus fréquentes dans les articles de linguistique que dans les articles médicaux. Il n'y a en revanche pas de différences notables entre les deux langues à cet égard.

## 7.5 Bilan

Dans ce chapitre, l'accent a été mis sur l'aspect pragmatique des modalisateurs épistémiques. Les occurrences de six marqueurs centraux ont été étudiées dans leurs cotextes respectifs afin d'en préciser les effets discursifs et de déterminer le rôle qu'elles jouent dans le texte. L'étude de ces marqueurs, qui se correspondent dans les deux langues, a révélé qu'au niveau pragmatico-fonctionnel, le variable de la langue a peu d'importance : on retrouve les mêmes fonctions pragmatiques dans le corpus français et le corpus norvégien. Cette observation est attendue, étant donné que nous sommes partie de la modalité épistémique en tant que catégorie sémantico-pragmatique dont le rôle est supposé être le même à travers différentes langues : c'est elle qui constitue le *tertium comparationis* de la présente étude. On pourrait cependant imaginer certaines spécificités linguistiques ; il serait par exemple envisageable que les langues utilisent à des degrés divers la fonction contenu et la fonction interpersonnelle, mais rien n'indique dans notre corpus de telles différences. En ce qui concerne le cotexte immédiat, cependant, une différence nette entre les deux langues se manifeste : les modalisateurs épistémiques dans le corpus norvégien apparaissent plus souvent en combinaison entre eux que ceux du corpus français. Cela peut être un indice du fait que les norvégiens utilisent plus l'atténuation que les francophones (une accumulation d'atténuateurs rend l'énoncé moins sûr), mais cela peut simplement aussi s'expliquer par l'utilisation de combinaisons figées. L'étude du cotexte a par ailleurs révélé que la source du jugement épistémique est rarement indiquée de manière explicite. Les marqueurs étudiés se présentent donc le plus souvent sous forme d'*expressions évaluatives objectivisées* (Lundquist 1990 : 135), c'est-à-dire des expressions dont la marque de la première personne semble être effacée. L'emploi de *sembler* dans le corpus Frling fait figure d'exception : dans ce corpus, *sembler* est souvent précédé d'un pronom datif (le plus souvent *me*), mais la plupart des exemples relèvent d'un même article, ce qui interdit de tirer des conclusions sur ce sujet.

Si la langue a peu d'influence sur l'usage fonctionnel des modalisateurs épistémiques, la discipline, en revanche, joue un rôle important. On retrouve dans une large mesure les mêmes fonctions chez les linguistes et les médecins-chercheurs, mais ces fonctions ne sont pas d'importance égale. Nous avons vu que dans les articles de médecine, les modalisateurs épistémiques examinés s'utilisent principalement avec une fonction contenu. Ils apparaissent dans la présentation de résultats et de conclusions, ils servent à proposer des hypothèses et des explications possibles et à signaler les limites de l'étude et leurs effets potentiels. Dans les articles de linguistique, on retrouve les mêmes usages, mais cette dernière fonction (signaler

les limites de l'étude) est beaucoup plus rare chez les linguistes, ce qui s'explique probablement par le nombre limité d'articles expérimentaux dans la partie linguistique du corpus. Outre ces usages, on trouve dans les articles de linguistique d'autres emplois qui sont plus explicitement interpersonnels ou persuasifs et qui sont rares dans les articles de médecine : il s'agit principalement d'usages motivés, partiellement ou entièrement, par la politesse, consistant à faire preuve de prudence quand on interprète ce que d'autres ont dit et à affaiblir la critique émise contre d'autres chercheurs. En outre, nous avons vu plusieurs exemples de modalisateurs épistémiques apparaissant dans des cotextes explicitement argumentatifs, où il est clair que l'auteur prend position et argumente en faveur de ses propres points de vue, parfois en s'opposant ouvertement à d'autres.

De ces observations on peut conclure qu'en ce qui concerne l'emploi de modalisateurs épistémiques, les aspects interactif et persuasif de l'article de recherche sont plus saillants dans les articles de linguistique que dans les articles de médecine, c'est-à-dire qu'ils se manifestent linguistiquement d'une manière plus explicite. Cependant, cette observation ne vaut pas pour tous les phénomènes étudiés dans le projet KIAP, du moins pas en ce qui concerne l'aspect interactif. En ce qui concerne les références bibliographiques, par exemple, l'aspect interactif est le plus saillant dans la discipline de médecine, en ce sens que le nombre de références est beaucoup plus élevé dans cette discipline que dans les autres (Fløttum 2003c ; Fløttum et al. 2006 : 219-221). D'autres voix se manifestent donc dans la discipline de médecine aussi bien que dans la discipline de linguistique, mais les linguistes entrent probablement plus facilement en dialogue voire en opposition avec ces autres voix dans leurs textes. Le rôle que les voix acquièrent semble donc quelque peu différent dans les deux disciplines.

Etant donné que l'aspect persuasif est plus saillant dans les articles de linguistique, on peut conclure que les linguistes promeuvent leur propre recherche d'une manière plus directe et explicite que ne le font les médecins-chercheurs. Cependant, l'importance et le nombre élevé de modalisateurs assumant une fonction contenu indiquent que l'atténuation n'est pas prioritairement, dans aucune des disciplines, une stratégie explicite de persuasion, mais qu'il s'agit souvent plutôt du souhait d'être précis quant au degré de certitude avec lequel on avance quelque chose. Nous avons vu que dans bien des cas, le facteur qui justifie le modalisateur est une incertitude réelle de la part des auteurs (cf. les phrases d'incertitude et les remarques sur les connaissances limitées, qui apparaissent principalement dans les articles médicaux), et que, par conséquent, l'emploi d'atténuateurs ne peut pas toujours être expliqué en faisant référence à la notion de politesse, ou au désir de persuader le lecteur (voir aussi

Lewin 2005). Or, l'emploi de modalisateurs épistémiques contribue à l'image d'un chercheur honnête et sérieux qui a pris en compte tous les facteurs potentiellement en jeu et toutes les sources d'erreurs envisageables. Cela a en retour un effet sur la recherche de l'auteur et permet de la promouvoir : elle apparaît comme plus sérieuse et plus crédible. Le fait de mettre quelques marques d'incertitude a donc pour effet d'augmenter la fiabilité de ce qui est dit. Par ailleurs, il est légitime, en linguistique, d'argumenter explicitement en faveur de son propre point de vue et à l'encontre d'autres. Cela peut s'expliquer par la nature des objets d'étude et des méthodes appliquées dans les deux disciplines : en médecine, on peut dans une plus grande mesure s'appuyer sur des résultats directement mesurables, alors qu'en linguistique le développement d'un raisonnement et, par conséquent, l'élaboration d'une argumentation explicite, semblent plus importants.

## 7.6 Quatre études de cas

Nous avons jusqu'ici étudié les marques de modalisation épistémique dans un contexte relativement restreint. Quelques études de cas permettront de mieux illustrer comment un auteur peut se servir de modalisateurs épistémiques tout au long d'un texte, dans le but d'informer, de convaincre ou d'inclure d'autres points de vue. Dans la présente section nous allons regarder de plus près quatre articles, un de chacun des sous-corpus de nos analyses qualitatives (Noling, Nomed, Friling, Frmed). Les quatre articles sélectionnés ne sont pas à considérer comme des exemples types de leurs groupes respectifs – chercher des articles “typiques” aurait été une tâche sans doute impossible et de surcroît peu utile. Comme le but ici est de mieux illustrer l'aspect pragmatico-fonctionnel des modalisateurs épistémiques, nous avons choisi des articles qui contiennent un nombre relativement élevé de ces marqueurs. La fréquence relative n'est pas nécessairement élevée – cela dépend de la longueur de l'article – mais les articles contiennent tous un nombre non négligeable d'attestations. Si nous n'avons pas volontairement recherché des articles typiques, les articles n'en sont pas moins représentatifs de leur groupe au sens où, à bien des égards, ils illustrent des tendances valables pour leurs groupes respectifs.

### 7.6.1 L'article *Noling 05*

Notre première étude de cas est l'article Noling05, qui est un article de linguistique norvégienne, écrit par un auteur masculin et publié en 1996 dans *Norsk Lingvistisk Tidsskrift*. Avec 35 attestations des modalisateurs épistémiques retenus dans la présente étude, l'article est celui du corpus Noling qui comporte le nombre le plus élevé de ces marqueurs. La fréquence relative (qui est de 4,2 pour mille) n'est cependant pas parmi les plus importantes (cf. appendice 2), mais le nombre élevé d'attestations fait de cet article un bon objet pour une étude de cas.

L'article porte sur la morphologie verbale en norvégien et le concept de « formes de base » morphologiques. L'objectif de l'article, tel qu'il est formulé dans le résumé, est d'aller à l'encontre d'une analyse antérieurement proposée, selon laquelle la forme du passé doit être considérée comme la forme de base d'une classe spécifique de verbes norvégiens, plus précisément, d'une classe appelée « *telja*-klassen ». L'auteur argue en faveur d'une analyse où la forme du présent est à considérer non seulement comme la forme de base de cette classe particulière, mais aussi comme la forme de base des verbes norvégiens en général.



On voit immédiatement que l'objectif prête à un ton critique et argumentatif. L'auteur explicite dans le résumé déjà qu'il va contester une théorie et en promouvoir une autre. La *discussion* a donc une position centrale dans l'article, ce qui peut expliquer le nombre élevé de modalisateurs épistémiques.

L'article s'ouvre par une présentation et une discussion générale du concept « forme de base » en morphologie (section 1). Dans la deuxième section, l'auteur passe à la classe verbale dont il est question. Cette section comprend une présentation de l'analyse contestée, avec une sous-section intitulée « objections contre Telemann (1969) et Kiefer (1970) » (Telemann et Kiefer étant ceux qui défendent l'analyse contestée). L'auteur y propose aussi une analyse alternative et en présente les avantages et les effets. Dans la section 3, ensuite, l'auteur propose quelques parallèles entre l'analyse de verbes et l'analyse de noms, avant de conclure l'article (section 4). Sa conclusion principale est que le concept de « formes de base morphologiques » est utile, mais qu'il faut l'ancrer dans l'usage du langage et non pas considérer les formes de base morphologiques comme des formes grammaticales sémantiquement vides. Il s'ensuit qu'il est préférable selon lui de considérer le présent (plutôt que le passé) comme la forme de base morphologique des verbes appartenant à la classe spécifique en question.

Dans ce qui suit, nous allons étudier de plus près les différentes sections de l'article et voir comment les modalisateurs épistémiques contribuent à la structure argumentative<sup>38</sup> du texte.

Dans l'introduction, l'auteur présente le cadre théorique de l'article, qui est celui de la grammaire cognitive. En l'opposant à la grammaire générative, il présente et défend cette théorie :

1. Innafor Kognitiv Grammatikk er ikke bøyingsklasser noe annet enn generaliseringer over likhet på uttrykkssida. Visse generative grammatikere **kan se ut til** å ha gitt bøyingsklasser en rolle som uavhengige diakritika i leksikon. [...] Innafor Kognitiv Grammatikk ville en altså hevde at bøyingsklasser ikke er reint grammatiske enheter, men uttrykk for generalisering over likhet. **Det virker rimelig å tru** at talerne kan observere likhet mellom for eksempel *skreik*, *beit* og *dreiv*.

'[...] Certains grammairiens générativistes **peuvent sembler** avoir donné aux conjugaisons un rôle de diacritiques indépendants dans le lexique. [...] Dans la grammaire cognitive, on dirait que les conjugaisons ne sont pas simplement des unités grammaticales, mais des manifestations d'une généralisation sur les analogies. **Il semble raisonnable de croire** que les locuteurs peuvent observer une analogie entre par exemple *skreik*, *beit* et *dreiv*.'

---

<sup>38</sup> Par *structure argumentative* nous entendons ici simplement la manière dont les idées sont exposées et développées au cours du texte, sans nous référer à une théorie particulière. La visée de l'argumentation et de convaincre.

Notons qu'en se prononçant sur les idées des grammairiens générativistes, l'auteur s'exprime avec précaution : *kan se ut til...* Comme nous l'avons vu, cela correspond à un usage important des modalisateurs épistémiques : le fait de critiquer et d'interpréter le travail d'autres chercheurs comporte toujours un risque, mais ce risque de « perdre la face » se réduit considérablement par l'emploi de quelques marqueurs d'atténuation.

Bien que la grammaire cognitive constitue le fondement théorique de l'auteur, il présente aussi ses doutes concernant certains aspects de cette théorie. A travers des énoncés tels que 2, l'auteur signale qu'il reste ouvert à d'éventuelles modifications qui pourraient résulter de recherches futures et évite ainsi de se présenter comme trop catégorique :

2. **Kanskje** vil innholdskravet i lengden vise seg å være vel sterkt; **det gjenstår å se**. Inntil videre er det **en sterk hypotese**.  
'**Peut-être que** la demande de contenu s'avérera trop forte à la longue ; **cela reste à voir**.  
Entre-temps, cela reste **une hypothèse forte**.'

Après l'introduction, l'auteur passe à une discussion du concept *forme de base* en morphologie. Dans l'extrait 3, il présente le fondement de ce concept :

3. Mange teoretikere **antar** altså at ei eller flere former i paradigmet har ei særstilling, men de bruker ikke alle denne **antakelsen** på akkurat samme måten.  
'Beaucoup de théoriciens **supposent** qu'une ou plusieurs formes dans le paradigme flexionnel ont une position particulière, mais ils n'utilisent pas tous cette **supposition** de la même manière.'

Le fondement du concept est donc présenté comme une supposition, mais la valeur épistémico-modale des formes *antar* et *antakelsen* semble ici relativement faible. Plutôt que de marquer une incertitude par rapport à la proposition qualifiée, on pourrait dire qu'elles marquent simplement un point de départ théorique. Le verbe *anta* et le nom *antakelsen* admettent donc les mêmes emplois que *assume* et *assumption* en anglais (*supra* 6.4.1.3) ; ils s'emploient parfois comme de véritables marqueurs d'incertitude, et dans d'autres cas l'aspect d'incertitude paraît moins important.

L'auteur continue ensuite en présentant deux manières différentes d'utiliser le concept de forme de base en morphologie :

- 4a. For det første går det an å se ei form i paradigmet som synkront grunnleggende [...] Slik har en **nok** tenkt i mange tradisjonelle grammatiske beskrivelser av skandinavisk. [...] Dette er **nok** den vanligste måten å bruke begrepet ledeform på.

'Premièrement, on peut considérer comme synchroniquement fondamentale une des formes dans le paradigme flexionnel [...] C'est ce qu'on a **probablement** pensé dans beaucoup de descriptions grammaticales traditionnelles des langues scandinaves. [...] Cela est **probablement** l'acception la plus fréquente du concept.'

- b) En annen bruksmåte står Bybee (1985) for. [...] Bybee setter altså likhetstegn mellom ledeform1 og ledeform2. **Det er ikke uten videre gitt** at dette er noen fornuftig manøver. 'Une autre acception du terme est celle de Bybee. Bybee assimile la forme de base 1 à la forme de base 2. **Il ne va pas de soi** que cette manœuvre soit raisonnable.'
- c) Etter mitt syn er det rimelig å gjøre som Bybee, altså å ekvivalere ledeform1 og ledeform2, [...] [D]et [er] **rimelig å tru** at talerne har en viss følelse for frekvens, og Bybee knytter altså begrepet ledeform til frekvens. 'A mon avis, il est raisonnable de faire comme Bybee [...] **Il est raisonnable de croire** que les locuteurs ont un certain sens de la fréquence, ...'

Voilà encore un exemple de ces cas où un auteur, en se prononçant sur les idées d'autres – nommés ou non – ajoute un élément d'atténuation (*slik har en nok tenkt...* ('c'est ce qu'on a **probablement** pensé...')). L'autre attestation de *nok* sert à qualifier le contenu informatif comme probable (exemple 4a). A première vue, l'auteur paraît être quelque peu sceptique sur la seconde acception du concept de forme de base (exemple 4b), mais il s'avère plus loin que ce scepticisme apparent fait partie d'une argumentation en faveur de cette acception (cf. 4c). L'auteur n'hésite pas à évoquer les points faibles d'un point de vue qu'il soutient lui-même, ce qui peut en fait augmenter sa crédibilité comme chercheur, dans la mesure où il signale ainsi le caractère réfléchi de son propos.

Dans l'extrait 5, l'auteur anticipe une conclusion en faveur de laquelle il argumentera plus loin. A l'aide de l'adverbe *trulig*, cette conclusion est présentée de façon atténuée, conformément aux conventions du discours scientifique. L'adverbe *nok* dans l'énoncé qui suit illustre bien la standardisation du phénomène de l'atténuation dans le discours scientifique : il est évident pour tout locuteur natif du norvégien que la proposition qualifiée par *nok* est vraie ; l'ajout d'un atténuateur dans ce contexte est donc superflu et simplement stylistique.

5. I regelen er det **trulig** ubestemt form ental som er ledeform i det norske substantivparadigmet (jf nærmere diskusjon nedafor). Men ved substantivet *forelder* forekommer **nok** flertalsforma mye oftere enn entalsforma, så her er helst *foreldre* ledeform.

'Ordinairement, c'est **vraisemblablement** la forme indéfinie au singulier qui est la forme de base dans le paradigme flexionnel des noms en norvégien [...] Mais en ce qui concerne le nom *forelder*, la forme plurielle apparaît **probablement** beaucoup plus souvent que la forme singulière, [...]

L'auteur continue à discuter les différentes manières de concevoir la notion de 'formes de base morphologiques', en maintenant le ton argumentatif :

6. En kan også tenke seg at ledeforma er den eneste representanten for ordet i leksikon, **men det er ikke gitt**. For eksempel opererer Carstairs-McCarthy (1991) med primær ledeform, sekundær ledeform osv. Iallfall i store bøyingsparadigmer **kan det virke** rimelig å regne med mer enn ei ledeform. [...] **Det virker også rimelig å anta** at vi husker best det vi oftest hører. Frekvens er likevel bare relevant innafor noen morfologiske modeller (se Enger 1994, Endresen 1995). For eksempel sier Aronoff (1994:37) at frekvens er teoretisk irrelevant, og innafor modeller som holder språkssystem og språkbruk skarpt atskilt, er Aronoffs syn **trulig** det eneste mulige.

'On pourrait s'imaginer que la forme de base est le seul représentant du mot dans le lexique, mais **cela n'est pas évident**. [...] Du moins pour les grands paradigmes flexionnels, **il peut sembler** raisonnable de compter plus qu'une forme de base. [...] **Il semble raisonnable aussi de supposer** qu'on se souvient le mieux de ce qu'on entend souvent. [...] au sein de modèles qui distinguent clairement langue et parole, le point de vue d'Aronoff est **probablement** le seul possible.'

Dans l'extrait 7, l'auteur cherche à expliquer pourquoi c'est l'infinitif, plutôt que le présent, qui a traditionnellement été considéré comme la forme de base morphologique des verbes norvégiens. Il présente et discute des explications possibles:

7. Ei tenkelig forklaring **kunne være** at den grammatiske tradisjonen har vært sterkt preget av trading, men den **hypotesen** kan **neppe** forklare alt. Derimot **kan** en viktig del av forklaringa være at en har sett tonemet i presens som forutsigelig, 'Une explication conceivable **pourrait être**..., mais cette **hypothèse** n'explique pas tout. En revanche, une partie essentielle de l'explication **peut être**...'

Avant de passer à la classe verbale en question, l'auteur émet des hypothèses concernant certains exemples qu'il trouve particulièrement difficiles à expliquer :

- 8 a. **Trulig** er framlydskonsonantismen her tatt inn fra infinitiven.

'Le son initial dans cet exemple est **probablement** tiré de l'infinitif.'

- b. Et eksempel er vestlandsk *fure* mot østlandsk *furu*. Det **kan se ut som** nominativ har vært ledeform i "gammal-vestlandsk", men ikke i "gammal-østlandsk". Det er likevel ikke opplagt hvorfor det **skulle være** tilfellet. **Det er vanskelig å se at det skulle ligge** frekvensskilnader til grunn.

'Un exemple est la forme *fure*, utilisée dans l'Ouest, versus la forme *furu*, utilisée dans l'Est. **Il peut sembler** que le nominatif ait été la forme de base dans l'ancien dialecte de l'Ouest mais non dans l'ancien dialecte de l'Est. Il n'est pourtant pas évident de voir pourquoi cela **serait** le cas. **Il est difficile de s'imaginer** que des différences de fréquence en **soit** la cause.'

Il en conclut que les formes de base ne peuvent pas tout expliquer, mais que cela n'est pas nécessairement une faiblesse :

9. Etter mi mening er det dessuten ikke nødvendigvis slik at en teori (eller modell) må gi oss alle svarene. **Kanskje** burde en heller se det som en fordel om en tilnæringsmåte hjelper oss til å se interessante spørsmål.

'A mon avis, il n'est d'ailleurs pas nécessaire qu'une théorie (ou un modèle) doive nous fournir toutes les réponses. **Peut-être** qu'on devrait plutôt considérer comme un avantage qu'une approche théorique nous aide à poser des questions intéressantes.'

Par l'adverbe *kanskje* l'auteur semble ici s'adresser au lecteur, et il s'agit donc d'une fonction interpersonnelle. On pourrait facilement imaginer un point d'interrogation après la dernière phrase, ce qui rendrait encore plus clair l'aspect interactif.

Après avoir présenté différentes manières de concevoir le concept de formes de base, l'auteur opte pour un point de vue inspiré par la grammaire cognitive. Dans une telle optique, les formes de base ont un contenu sémantique et elles sont ancrées dans l'usage du langage. Il oppose ce point de vue à un point de vue générativiste où les formes de base sont considérées comme des formes grammaticales sémantiquement vides et où, selon l'auteur, les aspects de prévisibilité et d'économie ont joué un rôle trop important dans la détermination des formes de base.

Cette prise de position théorique est suivie de la partie de l'article où l'auteur présente (section 2.1) et critique (section 2.2) l'analyse antérieurement proposée (une analyse d'inspiration générativiste) et où il présente sa propre analyse (section 2.3). La première section, où il présente l'analyse qu'il va critiquer, est assez neutre, de caractère informatif et contient peu d'atténuateurs. La section suivante, où il avance ses remarques critiques, est, comme on pourrait s'y attendre, beaucoup plus riche en atténuateurs. L'auteur commence par justifier son intérêt pour cette analyse qui date d'assez longtemps. Ainsi, il va au devant de la critique et il interagit avec le lecteur en ce qu'il tente de devancer les remarques de ce dernier :

10. Nå **kan vel** det å argumentere mot en 27 år gammal analyse **se ut som** å slå inn ei vidåpen dør, ikke minst ettersom Telemans og Kiefers analyse **vel** ikke har blitt allment akseptert innafor nordstikken. Men bortsett fra at analysen deres ikke er blitt grundig drøftet tidligere og fortjener å diskuteres seriøst, er analysen på mange vis typisk for sitt paradigme.

'**Je suppose que** le fait d'argumenter contre une analyse qui date d'il y a 27 ans **peut sembler** enfoncer des portes ouvertes, surtout étant donné que l'analyse de Telemans et Kiefer n'a pas, **je suppose**, été généralement admise au sein des recherches sur les langues nordiques. Mais à part le fait que leur analyse n'a pas été profondément discutée antérieurement et qu'elle mérite d'être sérieusement discutée, elle est, à bien des égards, une analyse typique de son paradigme.'

Cependant, la plupart des atténuateurs dans cette section servent à atténuer la critique émise contre d'autres chercheurs ou à faire preuve de prudence en se prononçant sur leurs idées. Ces emplois sont clairement liés à la politesse : la présence de quelques atténuateurs bien choisis a pour effet de rendre la critique moins dure, et l'auteur protège ainsi la face des chercheurs critiqués. En même temps, les atténuateurs protègent l'auteur lui-même, par exemple s'il s'avère plus tard qu'il avait tort. Voici quelques exemples :

11. Strategien er helt parallell til Teleman og Kiefer, enda Anderson i si tilnærming til morfologien **vel** har fjernet seg mer fra Chomsky & Halle (1968) enn mange andre forskere innafor det generative paradigmet.

'La stratégie est entièrement parallèle à celle de Teleman et Kiefer, bien qu'Anderson, dans son approche de la morphologie, se soit **sans doute** éloignée plus de Chomsky & Halle (1968) que beaucoup d'autres chercheurs appartenant au paradigme générativiste.'

12. Analysen til Teleman og Kiefer er psykolingvistisk sett lite tiltalende. Vi veit at barn lærer preteritum relativt seint, seinere enn infinitiv og presens. Da **virker det tvilsomt** at preteritumsstammen **skulle være** grunnleggende.

'L'analyse de Teleman og Kiefer est peu convaincante d'un point de vue psycholinguistique. On sait que les enfants acquièrent le **prétérit** relativement tard, plus tard que l'infinitif et le présent. **Il semble donc douteux** que la racine du **prétérit soit** la forme de base.'

13. Jensen (1990:158) setter da opp de underliggende representasjonene som gir størst mulig prediktabilitet. Det er **vel** nettopp et eksempel på tendensen til å se allomorfi som fonologi.

'Jensen (1990:158) énumère les représentations sous-jacentes qui donnent le plus grand degré de prévisibilité. Il doit bien y avoir là justement un exemple de la tendance à voir l'allomorphie comme de la phonologie.'

14. Katamba (1993:31) er **vel** en representant for denne tendensen når han skiller mellom regelbundet fonologisk veksling på den ene sida og morfologisk suppletivisme på den andre. Skillet er for enkelt.

'**Je suppose que** Katamba (1993:31), en distinguant d'un côté l'altération phonologique gouvernée par des règles et de l'autre côté le suppletivisme morphologique, est un représentant de cette tendance. La distinction est trop simple.'

15. I sin kritikk av Chomsky & Halles fonologiske modell taler Kiparsky (1982) om diakritisk merking i leksikon forkledd som fonologisk regel. Analysen til Teleman og Kiefer **ser ut til** å være et eksempel på dette.

'Dans sa critique du modèle phonologique de Chomsky & Halle, Kiparsky (1982) parle de marquage diacritique dans le lexique camouflé en règle phonologique. L'analyse de Teleman et Kiefer **paraît** en être un exemple.'

16. Teleman og Kiefer sier preteritum er ledeform – mer presist at preteritumsstammen er grunnleggende – for telja-klassen, mens de – i tråd med mer tradisjonelle analyser - sier at infinitivsstammen er grunnleggende for de øvrige klassene. Men dette **virker** lite tiltalende.

'Teleman et Kiefer maintiennent que le prétérit est la forme de base – plus précisément, que la racine du prétérit est la forme de base – de la classe verbale « telja », mais aussi – en accord avec des analyses plus traditionnelles – que la racine de l'infinitif est la forme de base des autres classes verbales. Mais cela **semble** peu convaincant.

17. Inntil videre **ser det altså ut som** en immuniseringsstrategi å ikke ville bry seg om diakrone data.

'Jusqu'à nouvel ordre, cela **semble** une stratégie d'immunisation que de ne pas vouloir tenir compte de données diachroniques.'

18. Nå er ikke Andersons "amorfe morfologi" emne for denne artikkelen, men en sentral forutsetning innafor den modellen **ser ut til** å være at "dobbel merking" er utelukket

'La morphologie amorphe d'Anderson n'est pas le sujet du présent article, mais un présupposé central de ce modèle **semble** être que le « marquage double » est exclu'

19. Etersom Telemans og Kiefers analyse **neppe** er holdbar, **kan** parallellen **tyde på** at det også er noe i vegen med Andersons analyse av tyske verb.

'Etant donné que l'analyse de Teleman et Kiefer ne tient **probablement pas** la route, le parallèle **peut indiquer** qu'il y a aussi quelque chose qui cloche dans l'analyse d'Anderson sur les verbes allemands.

Malgré les atténuateurs, l'auteur est plutôt direct et la critique reste assez virulante, comme il ressort de l'exemple 20 :

20. Altså må Telemans og Kiefers analyse avvises.

'Par conséquent, l'analyse de Teleman et Kiefer doit être rejetée.'

Il ne s'agit donc pas pour lui de prendre des gants avec ses adversaires. Cependant, par l'emploi d'atténuateurs comme ceux cités dans les exemples 11-19, l'auteur émet ses critiques d'une manière qui répond aux exigences d'un style scientifique.

En présentant sa propre analyse (2.3), l'auteur prend soin de devancer des remarques critiques et d'y répondre en avance. Ce faisant, il conteste (prudemment par *neppe*) le fondement de la critique envisagée :

21. Mange vil da også mene at det gir en klønete beskrivelse å ikke beskrive vokalvekslinga *a-e* i *mann-menn* og *telja-talde* ved samme regel. Dette argumentet hviler likevel på diskutabile forutsetninger [...] Forutsetningene er **neppe** holdbare. Iallfall bør omlyden i norsk i dag **neppe** betraktes som en fonologisk regel.

'Beaucoup diraient que cela donne une description peu élégante que de ne pas décrire l'inflexion vocalique *a-e* dans *mann-menn* et dans *telja-talde* par la même règle phonologique. Cet argument repose pourtant sur des prémisses discutables [...] Les prémisses ne tiennent **probablement pas**. Du moins, l'inflexion vocalique dans le norvégien d'aujourd'hui ne doit **probablement pas** être considérée comme une règle phonologique.'

*Neppe* marque un relativement haut degré de certitude, mais il s'agit d'une certitude par rapport à la *fausseté*, et non la vérité de la proposition qualifiée (cf. *supra*, 3.1.2.1). Cette particule modale fonctionne ainsi comme une variante atténuée de *ikke* (*ne...pas*).

Dans la troisième partie, l'auteur passe des verbes aux noms, tout en proposant des parallèles entre le cas des noms et le cas des verbes. Son analyse sur les noms sert à étayer le point de vue cognitiviste sur le concept de formes de base, pour lequel il a opté lors de la première section. A nouveau, il présente deux analyses différentes et avance des arguments pour et contre. On reconnaît le ton argumentatif et polémique de la première partie :

22. Likevel fins det gode argumenter mot analyse B. Et argument er at ubestemt form i regelen er mer frekvent enn bestemt. Da **virker det mindre sannsynlig** at bestemt form er ledeform.

'Il existe pourtant de bons arguments contre l'analyse B. Un de ces arguments est que la forme indéfinie est en général plus fréquente que la forme définie. **Il semble donc moins probable** que la forme définie soit la forme de base.'

Son argumentation aboutit à une conclusion qui va en faveur de l'une des analyses et qui implique une critique de l'autre :

23. Kort sagt **tyder** altså både språkhistorisk evidens og evidens fra barnespråk **på** at det i regelen er ubestemt form som er grunnleggende. Dette taler for analyse U. Dette **indikerer** at omsyn til prediktabilitet og økonomi ikke aleine er nok. Eksempelet **indikerer** også at Wurzels tilnærming til Kennformen har sine svake sider.

'En bref, les données diachroniques aussi bien que les données du langage enfantin **suggèrent** qu'ordinairement, c'est la forme indéfinie qui est de base. Cela plaide en faveur de l'analyse U. Cela **indique** qu'il ne suffit pas de tenir compte des phénomènes de prévisibilité et d'économie seuls. L'exemple **indique** aussi que l'approche de Wurzel à l'égard de *Kennformen* a ses côtés faibles.'

Dans la quatrième et dernière partie, qui est celle de la conclusion, les atténuateurs sont utilisés afin de résumer avec prudence les conclusions tirées dans l'article. On note donc une conformité aux conventions du discours scientifique et au principe évoqué par Myers (1989 : 13) selon lequel « *all statements of new knowledge must be hedged* ». Les conclusions d'un article scientifique sont rarement présentées sans atténuation.

24. Det **virker** rimeligere å ta presens som ledeform både i telja-klassen spesielt og i norsk verbalmorfologi generelt. Det **virker** også rimelig å ta ubestemt form ental som ledeform ved norske substantiv.

'Il **semble** plus raisonnable de considérer le présent comme la forme de base de la classe « telja » en particulier aussi bien que dans la morphologie verbale norvégienne en général. De plus, il **semble** raisonnable de considérer la forme indéfinie du singulier comme la forme de base des noms norvégiens.



25. Det **kan tyde på** at visse aspekter av Naturlig Morfologi trenger å revurderes.  
'Cela **peut indiquer** que certains aspects de la morphologie naturelle demandent une remise en question.'
26. Ledeformer er en mulig måte å gjøre dette på (men **neppe** den eneste).  
'Les formes de base constituent une alternative possible (mais ce n'est **probablement pas** la seule).
27. **Trulig** er den ideen som er innbakt i "The single reference form requirement" gunstig å ha med seg videre.  
'Il sera **vraisemblablement** utile de prendre en compte l'idée inhérente à "The single reference form requirement".'

En résumé, on peut dire que dans la première partie, l'auteur présente et discute les fondements théoriques et envisage les arguments en faveur ou à l'encontre des différents points de vue. La présence de modalisateurs épistémiques est liée à cette argumentation. Une fois les fondements théoriques posés, l'auteur passe, dans la deuxième partie, à la classe de verbes sur laquelle s'appliquent les analyses proposées. Les modalisateurs servent alors surtout à atténuer la critique émise contre d'autres chercheurs et à faire preuve de prudence en interprétant des textes écrits par d'autres linguistes. Dans la troisième partie, qui est très brève comparée aux deux précédentes, l'auteur porte son attention sur les noms et poursuit son argumentation par rapport aux différentes manières de concevoir les formes de base morphologiques. Dans la quatrième et dernière partie, les marqueurs épistémico-modaux sont liés à la présentation nuancée des conclusions.

Les marqueurs épistémico-modaux sont répartis de manière relativement équitable entre les quatre sections : pour chacune d'elles, le nombre de modalisateurs se situe autour d'un nombre moyen de 5 à 7 occurrences par page. Leurs fonctions varient cependant d'une section à l'autre : dans la deuxième section, c'est la fonction interpersonnelle qui domine : comme nous l'avons vu, la plupart des modalisateurs dans cette partie du texte servent à chercher à convaincre le lecteur (voir par exemple l'extrait 12) ou à créer l'impression que l'auteur s'adresse directement au lecteur. La fréquence importante de la particule modale *vel* dans cette partie en est un exemple : cette particule sert souvent à inciter une réponse de la part de l'interlocuteur (Uri 1999). A l'oral, les énoncés contenant *vel* sont souvent suivis d'un "n'est-ce pas?" La particule marque une certaine incertitude, mais elle vise en même temps à faire vérifier l'information par l'interlocuteur. Par l'emploi de *vel*, l'auteur signale qu'il présuppose que l'information transmise est connue et que le lecteur est d'accord (cf. Uri

1999). Voilà peut-être la raison pour laquelle l'auteur ne se donne pas vraiment la peine d'étayer les assertions qualifiées par *vel*.

Dans la section 4, c'est la fonction contenu qui domine : les modalisateurs dans cette partie sont utilisés afin de préciser le statut de vérité des propositions qualifiées. Ils indiquent le degré de certitude que l'auteur souhaite attacher au contenu informatif. Dans les autres sections (1 et 3), les fonctions sont plus mixtes : on y trouve des exemples des deux catégories.

Pour finir, quelques remarques s'imposent sur le choix de modalisateurs qu'a fait l'auteur de Noling05. Le lecteur alerte aura peut-être remarqué que l'auteur de Noling05 utilise plusieurs marqueurs qui ne font pas partie des marqueurs retenus pour l'analyse quantitative. Il s'agit notamment des particules modales *vel*, *nok* et *neppe* ainsi que de *virke*, verbe polysémique qui est souvent un synonyme de *synes / se ut til/som* et que l'auteur emploie tantôt seul, tantôt en combinaison avec d'autres modalisateurs épistémiques. En fait, ce verbe – dont il n'y avait qu'une occurrence épistémico-modale dans le corpus exploratoire – est employé dans un sens épistémico-modal 12 fois dans Noling05, ce qui en fait le modalisateur épistémique le plus fréquent dans cet article. *Virke* est suivi de *kan* (10 occurrences épistémico-modales), *se ut* (8) et *vel* (8). Ensuite suivent *nok* (6), *neppe* (5), *anta* (4), *trulig* (4), *kanskje* (3), *tyde på* (3) et une petite série d'autres marqueurs qui sont employés une ou deux fois. Il ressort de cette liste que l'auteur paraît avoir une préférence particulière pour les particules modales : outre *vel*, *nok* et *neppe*, il utilise aussi *jo* et *nemlig* ; cela étant, comme celles-ci marquent la certitude plutôt que l'incertitude, nous les laissons de côté pour le moment. Bien que *vel*, *nok* et *neppe* fassent partie des modalisateurs épistémiques les plus fréquents dans Noling05, cela ne semble pas refléter les choix lexicaux des linguistes norvégiens en général. Il y a en tout 18 occurrences de *neppe* dans le corpus Noling, dont cinq relèvent de Noling05 (*neppe* étant monosémique, ces 18 occurrences expriment toutes une modalité épistémique). Il y a 16 occurrences épistémico-modales de *vel*, dont la moitié se trouve dans Noling05, et il y a 14 occurrences épistémico-modales de *nok*, dont six sont tirées de Noling05. Autrement dit, l'article Noling05 contient presque 40 % des occurrences des particules modales marquant une incertitude. Bien que le nombre d'attestations puisse sembler élevé, l'emploi de particules modales semble être lié aux préférences individuelles de certains auteurs et elles ne sont généralement pas fréquentes. En ce qui concerne la partie médicale du corpus, les marqueurs de ce type sont quasi-absents et leur fréquence dans le corpus total est donc relativement basse.

Il en va de même pour le verbe *virke* : la fréquence de ce verbe dans Noling05 pourrait indiquer qu'on a laissé échapper un marqueur important lors de la sélection de marqueurs. Or, une recherche automatique de ce verbe dans le corpus – évidemment suivie d'une analyse désambiguïsante des attestations – révèle qu'il n'y a que cinq autres occurrences épistémico-modales dans le corpus Noling et une dans le corpus Nomed.

On peut en conclure que le choix de modalisateurs qu'a fait l'auteur de Noling05 n'est pas tout à fait représentatif des linguistes norvégiens en général. L'auteur semble favoriser un style quelque peu oral ou informel : les particules modales sont souvent considérées comme appartenant plutôt au langage oral, et *virke* est une variante moins formelle de *synes*. *Se ut* s'utilise dans les deux registres, mais il ressort des tableaux dans le chapitre 6.4 que les médecins-chercheurs préfèrent s'en tenir à la variante la plus formelle, *synes*. Le fait que Noling05 contient plusieurs marqueurs qui ne sont pas inclus dans notre sélection n'est donc pas nécessairement un indice d'un choix inapproprié ou insuffisant de marqueurs. En revanche, les modalisateurs privilégiés par l'auteur de Noling05 le distingue de ses collègues. Ce constat illustre l'importante variation individuelle qui peut exister au sein d'un même groupe et nous rappelle l'importance de prendre celle-ci en compte dans nos observations.

### 7.6.2 L'article *Nomed08*

L'article *Nomed08* traite des différences de sexe chez les héroïnomanes à Oslo. Il est écrit par un auteur féminin et a été publié dans *Tidsskrift for den norske lægeforening* en 2000. L'article contient 21 attestations des marqueurs épistémico-modaux retenus, soit une fréquence relative de 8,9 pour mille. Cette fréquence est très élevée par rapport aux autres articles du corpus *Nomed*, dans lequel la moyenne est de 3,3 (cf. tableau 4, chapitre 6). Cet article est celui du corpus *Nomed* qui est le plus riche en atténuateurs : avec sa fréquence relative de 8,9 il se situe en troisième position dans le corpus total, uniquement devancé par deux articles anglais. La fréquence élevée de modalisateurs épistémiques dans l'article *Nomed08* fait de lui un texte bien approprié pour examiner de plus près les différentes fonctions pragmatiques de ces marqueurs.

L'article *Nomed08* se distingue de la majorité des articles médicaux par le fait qu'il est écrit par un seul auteur : les articles en recherche médicale sont généralement rédigés par plusieurs auteurs. Il est bien possible que cela puisse expliquer le nombre élevé d'atténuateurs : un auteur qui écrit seul doit sans doute être encore plus prudent et modeste qu'un groupe d'auteurs – cela paraît raisonnable surtout si l'on prend en considération le fait que la rédaction en groupe est devenue la norme en recherche médicale, et cela à tel point que le lecteur serait presque sceptique envers les articles signés par un seul individu (Gilhus 2003 : 19). Il convient pourtant de noter que cela vaut en particulier pour les articles publiés dans les revues internationales les plus prestigieuses, et qu'il est plus accepté de rompre avec cette tendance générale dans une revue nationale telle que *Tidsskrift for den norske lægeforening*.

L'article s'ouvre par un résumé qui présente brièvement l'objectif – qui est d'examiner des différences de sexe entre les héroïnomanes en termes de santé et de type d'abus –, la méthode et les résultats. L'introduction ne reprend pas l'objectif, mais présente brièvement quelques différences de sexe déjà connues – établies par des études antérieures. Viennent ensuite les sections obligatoires sur les matériels et méthodes, les résultats et la discussion. L'auteur clot l'article par une brève conclusion. Le résultat principal est de montrer que les toxicomanes féminines s'injectent de l'héroïne plus souvent et à des doses plus élevées que les toxicomanes masculins, sans que cela semble avoir d'effet ni sur le taux de mortalité ni sur le taux de maladies transmises par l'utilisation de seringues.

Dans l'introduction, les atténuateurs utilisés sont liés à la présentation de résultats issus d'études antérieures. Ces résultats sont introduits de façon atténuée probablement parce

qu'ils ne sont pas (encore) considérés comme confirmés ou « vrais » par la communauté scientifique (cf. *supra* 6.5).

1. Kvinner **ser** blant annet **ut til** å utvikle avhengighet raskere enn menn (7) og søke behandling tidligere enn menn (8). Skandinaviske studier **kan tyde på** at kvinner har noe bedre nytte av behandling enn menn (9).  
'Les femmes **semblent** développer une dépendance à l'égard de la drogue plus vite que les hommes (7) et demander un traitement plus tôt que les hommes (8). Des études scandinaves **peuvent indiquer** que les femmes profitent un peu plus que les hommes du traitement (9).'

Normalement, les atténuateurs sont plus ou moins absents de la section qui présente les matériels et la méthode (Salager-Meyer 1994, Varttala 2001). Cependant, dans la section Matériels et Méthode de Nomed08, on relève plusieurs attestations d'atténuation. Les atténuateurs sont liés à l'évaluation que l'auteur fait de la méthode pour laquelle elle a opté : une telle évaluation de la méthode appliquée est quasi-obligatoire, mais elle apparaît le plus souvent dans la section Discussion plutôt que dans la section Matériels et Méthode. Sur ce point, l'article Nomed08 se distingue du prototype de l'article en recherche médicale. Certains des atténuateurs qui apparaissent dans cette section servent à défendre la représentativité de l'échantillon étudié.

2. Representativiteten til de første 900 intervjuene er blitt diskutert i mer detalj i Bretteville-Jensen & Sutton (10), men summarisk kan man si at det er fem indikasjoner på at et forholdsvis representativt utvalg av injeksjonsmisbrukere i Oslo er kommet med:  
- Sprøytebussen delte i 1997 ut mer enn 1,5 millioner sprøyter og hadde mer enn 103 000 enkeltbesøk (11). Dette **indikerer** at en stor andel av misbrukerne i Oslo-området benytter seg av Sprøytebussens tjenester, da **det er anslått** at det finnes ca. 5 000 injeksjonsmisbrukere i Oslo (5).  
'[...] en bref, on peut dire qu'il y a cinq indications pour dire qu'un échantillon relativement représentatif des toxicomanes intraveineux à Oslo a été pris en compte :  
- Le bus d'échange de seringues a distribué en 1997 plus de 1,5 millions de seringues et a reçu plus de 103 000 visites (11). Cela **indique** qu'une grande partie des toxicomanes à Oslo utilisent les services du bus d'échange de seringues, étant donné que le nombre de toxicomanes intraveineux à Oslo est **estimé à 5000** (5).'

D'autres atténuateurs sont liés à la présentation de limites méthodologiques, et ils visent souvent à convaincre les lecteurs que les limites ne sont pas trop importantes. C'est ici que l'on trouve les exemples les plus clairs de la fonction interpersonnelle des modalisateurs dans cet article. Dans l'exemple 3, la première instance de *trolig* et la deuxième instance de *sannsynligvis* remplissent une fonction persuasive plutôt qu'informatrice, alors que c'est l'inverse pour la première instance de *sannsynligvis* et la deuxième de *trolig* :

3. - Misbrukere med et høyfrekvent injeksjonsmisbruk vil **sannsynligvis** besøke Sprøytebussen oftere enn misbrukere som injiserer sjeldnere. Tyngre misbrukere var imidlertid noe mer tilbøyelig til ikke å ønske å delta i studien, slik at en overrepresentasjon av de førstnevnte **trolig** ikke er et stort problem.

[...]

Da intervjuene er anonyme, er **trolig** enkelte personer blitt intervjuet mer enn én gang i løpet av undersøkelsesperioden. Imidlertid har vi **sannsynligvis** greid å unngå at samme person er blitt intervjuet flere ganger innenfor den enkelte kvartalsvise innsamling, ...

‘- Les toxicomanes qui s’injectent très fréquemment vont **probablement** visiter le bus d’échange de seringues plus souvent que ceux qui s’injectent plus rarement. Les toxicomanes sévères étaient pourtant quelque peu plus enclins à refuser de participer à l’étude, et une surreprésentation des premiers n’est donc **vraisemblablement** pas un grand problème.

[...]

A cause de l’anonymat, certaines personnes ont **vraisemblablement** été interviewées plus d’une fois. Cependant, nous avons **probablement** réussi à éviter qu’une même personne soit interviewée plusieurs fois dans la même collecte trimestrielle,...

La section Résultats est dépourvue d’atténuateurs (ce qui concorde avec la tendance générale pour les articles en médecine), mais on en trouve en revanche une multitude dans la partie Discussion. Il s’agit d’occurrences orientées vers le contenu plutôt que vers le lecteur. Dans l’extrait 4, le rapport apparemment évident qui existe entre le nombre de surdoses mortelles d’une part et la quantité d’héroïne injectée ainsi que la combinaison de l’héroïne avec d’autres drogues d’autre part est présenté de façon atténuée, probablement parce que les résultats de l’étude ne confirment pas l’existence d’un tel rapport. La tournure *det synes naturlig å anta* (‘il semble naturel de supposer’) est ici employée pour transmettre une hypothèse qui après coup s’est avérée en partie erronée ou peu pertinente :

4. Det **synes** naturlig å **anta** at narkotikadødsfall henger sammen med mengden heroin som er inntatt og kombinasjonen av heroin og andre rusmidler.

‘Il **semble** naturel de **supposer** que les surdoses mortelles sont associés au dosage d’héroïne pris et à la combinaison d’héroïne avec d’autres stupéfiants.’

Cependant, le taux de surdoses mortelles est plus bas parmi les toxicomanes féminines que parmi les toxicomanes masculins. Comme l’hypothèse initiale s’est avérée incapable d’expliquer ces différences de sexe, l’auteur cherche à trouver d’autres explications. Les atténuateurs viennent accompagner les explications possibles :

5. Hva kan forklare den lavere dødeligheten blant kvinnelige intravenøse stoffmisbrukere?  
[...] Analyser foretatt ved Statens rettsstoksikologiske institutt viser dessuten at det er en økning i andelen narkotikadødsfall der man finner alkohol i blodet til de avdøde (16). Dette gjelder menn, og som tabell 3 viser, oppgir også langt flere menn enn kvinner bruk av alkohol i foregående måned. Menn oppgir også større forbruk av cannabis. Det **kan**

med andre ord **synes** som kombinasjonen av heroin og alkohol/hasj er mer risikopreget enn det man tidligere har antatt. Dersom dødelighetsrisikoen øker med alder og misbrukslengde, **kan** imidlertid den relativt lave kvinneandelen blant eldre misbrukere, slik tabell 1 illustrerer, bidra til å forklare forskjellene.

'[...] Autrement dit, il **peut sembler** que la combinaison de l'héroïne avec l'alcool/le cannabis est plus risquée que ce que l'on supposait auparavant. Si le risque de mortalité augmente avec l'âge et la durée de l'abus, la proportion relativement faible de femmes parmi les toxicomanes âgés, comme l'illustre le tableau 1, **peut** pourtant contribuer à une explication des différences.'

L'auteur cherche aussi à expliquer pourquoi les femmes sont atteintes moins fréquemment que les hommes de maladies transmissibles comme l'hépatite A, B et C. Encore une fois, des atténuateurs (*skulle anta* ('aurait pensé')) sont utilisés pour introduire une hypothèse qui s'avère erronée :

6. Man **skulle anta** at risikoen for denne type sykdommer, som i hovedsak smitter via sprøytedeling, var mer utbredt blant kvinnelige enn mannlige sprøytenarkomane, med tanke på at kvinner injiserer oftere. Opplysninger fra Statens institutt for folkehelse **tyder på** at dette ikke er tilfellet.

'On **aurait pensé** que le risque de ce type de maladies, qui se transmettent surtout par le partage de seringues, était plus répandu parmi les toxicomanes féminines que parmi les toxicomanes masculins, étant donné que les femmes s'injectent plus souvent. Des renseignements de la part de l'Institut national de la santé publique **suggèrent** que cela n'est pas le cas.'

L'hypothèse n'est pas totalement réfutée, car l'emploi de *tyde på* indique que l'auteur n'exclut pas la possibilité que l'information de l'Institut national de la santé publique reste à vérifier. En cherchant à expliquer le phénomène, l'auteur souligne que la cause en est inconnue ; elle signale donc un manque de connaissances qui vaut pour le domaine d'étude en général et pas seulement pour elle. Il s'agit ici d'une incertitude réelle et la fonction des modalisateurs est avant tout du type « contenu » :

7. Tallene **tyder** med andre ord **på** at kvinner har en lavere risiko for hepatittsmitte, til tross for høyere frekvens av sprøytebruk. *Årsaken til fenomenet er ikke kjent*, men tallene **indikerer** at kvinner **kan** ha andre strategier for sprøytedeling enn menn.

'Autrement dit, les chiffres **suggèrent** que les femmes ont un risque plus faible d'être infectées par l'hépatite, malgré la fréquence plus élevée d'injections. *La cause de ce phénomène n'est pas connue*, mais les chiffres **indiquent** que les femmes **pourraient** avoir d'autres stratégies que les hommes pour partager des seringues.'

Dans la conclusion, qui est citée ci-dessous dans son ensemble, les modalisateurs employés servent à atténuer les nouvelles assertions (les résultats principaux de l'étude) ou à proposer des explications possibles. La conclusion est brève, mais contient cinq modalisateurs épistémiques, ce qui donne une proportion de 42,7 pour mille.

8. Dataene presentert i denne artikkelen **tyder på** at norske kvinnelige sprøytenarkomane injiserer oftere enn menn og at de injiserer mer heroin per injeksjon. Årsakene til dette kan være biologiske (kvinner trenger **kan hende** mer heroin enn menn for å oppnå samme rusnivå) eller sosiale (kvinner **kan** for eksempel ha mer å "medisinere" bort - de **kan** være mer stigmatisert, ha større psykiske belastninger, flere problemer i forhold til barn, partnere eller foreldre osv.). Uansett årsak **ser ikke den høyere injeksjonsfrekvensen og det større forbruket ut til** å resultere i en høyere risiko for narkotikadødsfall eller sykdommer som vanligvis smitter via sprøytedeling (hepatitt A eller B). Derimot har kvinnelige sprøytenarkomane økt risiko for HIV og seksuelt overførbare sykdommer.
- 'Les données présentées dans cet article **suggèrent** que les toxicomanes féminines norvégiennes s'injectent plus souvent que les hommes et qu'elles s'injectent des doses plus grandes. Les causes de cela peuvent être d'ordre biologique (les femmes ont **peut-être** besoin d'un plus grand dosage d'héroïne que les hommes pour obtenir les mêmes effets) ou social (les femmes **peuvent** par exemple avoir plus de problèmes qu'elles doivent soulager avec des médicaments – elles **peuvent être** plus stigmatisées, avoir des épreuves psychiques plus sévères, plus de problèmes liés aux enfants, compagnons, parents etc.). Quelle que soit la cause, la fréquence plus élevée d'injections et les dosages plus grands ne **paraissent** pas avoir pour conséquence un risque plus élevé de surdoses mortelles ou de maladies qui se transmettent normalement par le partage de seringues (hépatite A ou B). En revanche, les toxicomanes féminines ont un risque plus élevé de VIH et de maladies sexuellement transmissible.'

En résumé, on peut dire que les modalisateurs épistémiques de Nomed08 assument, dans la grande majorité de cas, une fonction contenu plutôt qu'une fonction interpersonnelle. Ils sont utilisés principalement pour présenter les résultats et les conclusions d'une manière atténuée, en conformité avec les conventions qui règnent dans le monde de la science, ou pour présenter des explications possibles d'un phénomène. Certains sont utilisés pour signaler des limites méthodologiques, et ils acquièrent alors parfois une fonction interpersonnelle consistant à convaincre le lecteur que les limites sont, au fond, peu importantes.

Contrairement à Noling05, Nomed08 contient très peu de modalisateurs épistémiques qui ne font pas partie des marqueurs retenus pour l'analyse quantitative. Le marqueur le plus utilisé est *kan* (6 occurrences), suivi de *tyde på* (4), *anta* (3) puis de *indikere*, *sannsynligvis*, *se ut*, *synes*, *trolig* (avec 2 occurrences chacun). Tous ces marqueurs, sauf *indikere*, sont inclus dans l'analyse quantitative. Outre les deux occurrences de *indikere*, il n'y a que deux occurrences d'autres marqueurs non inclus : *regne med* (1) et *anslått* (1). En ce qui concerne le choix de marqueurs, Nomed08 est donc typique de son groupe, et reflète les résultats de l'analyse exploratoire.



### 7.6.3 L'article *Frling15*

Frling 15 est un article de linguistique française écrit par un auteur masculin et publié en 2002 dans *Travaux de Linguistique*. L'article contient 17 attestations des marqueurs sélectionnés, soit une fréquence relative de 4,1 pour mille. Cette fréquence est assez élevée en comparaison des autres articles du groupe Frling, dans lequel la fréquence moyenne est de 1,9 pour mille.

L'article porte sur la sémantique des prépositions, et l'idée principale de l'auteur est que l'approche traditionnelle de la sémantique des prépositions, qui tend à assimiler le système des prépositions avec un schéma spatial, est insuffisante pour expliquer leur sens. Selon lui, il faut prendre en compte non seulement les valeurs spatio-configurationnelles, mais aussi des valeurs aspectuelles, subjectives et qualitatives et les inscrire dans la signification de base des prépositions.

Notons que, tout comme pour Noling05, l'objectif de cet article est de mettre en discussion une théorie traditionnelle et d'en proposer un remaniement ou du moins d'importants élargissements. Comme nous l'avons vu, un tel objectif va de pair avec un ton polémique, qui, à son tour, implique l'usage d'atténuateurs, et cela pourrait expliquer pourquoi Frling15 est relativement riche à ce niveau. Il ne faut pas oublier cependant que la mise en débat de travaux antérieurs et d'analyses traditionnelles est très courante dans les articles de linguistique.

Après le résumé et une introduction qui présentent brièvement l'objectif et l'idée principale, l'auteur passe à une présentation systématique des critiques qu'il adresse à l'encontre de l'analyse traditionnelle (section 1). Dans la section 2, il propose quelques modifications de la théorie traditionnelle et présente sa propre approche. Il argumente en faveur de ce point de vue en mettant en avant quelques observations sur une série de prépositions dites « spatiales ». Dans la dernière section (la conclusion), il propose les conclusions tirées à partir des observations faites dans la section précédente.

Presque la moitié des modalisateurs épistémiques de l'article Frling15 s'accumulent dans la section sur la préposition *dans* (on revient à cette section plus loin). Le reste des marqueurs se distribue dans le texte de manière plus éparse. A l'exception de la section qui porte sur *dans*, il n'y a donc pas de partie spécifique aux modalisateurs, contrairement à ce qui est souvent le cas dans les articles en recherche médicale avec la partie Discussion. Il est cependant quelque peu étonnant que la première section, où l'auteur émet des remarques critiques contre la théorie traditionnelle, soit dépourvue de marqueurs épistémico-modaux. La critique est présentée de façon assertive (cf. *oublie, se heurte au fait évident*) :

1. Sur un plan cognitif, la référence à l'espace vu oublie les enseignements de la théorie de la Gestalt (ou encore de la phénoménologie de la perception) qui souligne à l'inverse l'idée que les objets sont des produits de l'activité subjective, ou encore des corrélats liés à l'exploration par le sujet d'un "espace" d'abord autocentrique et qualitatif.

[...]

Sur un plan linguistique, l'assimilation du système des prépositions avec un schéma spatial se heurte au fait évident que les prépositions ne sont qu'un véhicule parmi bien d'autres des repérages spatiaux, statiques et dynamiques.

Dans la brève introduction, citée ici dans son ensemble (exemple 2), on trouve en revanche deux marqueurs épistémico-modaux, tous deux liés à la mise en discussion de la théorie traditionnelle de la sémantique prépositionnelle. Avec le premier, qui est une occurrence de *paraître*, une distance est créée entre les apparences et la réalité (ce qui est une fonction typique des verbes *sembler* et *paraître* et leurs équivalents, cf. *supra* 7.2.6 et Kibbee 1995). Le deuxième sert à atténuer la critique contre la théorie mise en débat ; une réponse affirmative à la question posée aurait rendu la critique bien plus virulente que la réponse hésitante exprimée par *peut-être*.

2. J'entends revenir ici brièvement sur l'idée déjà exposée dans des travaux antérieurs (Cadiot 1997, 1999, Visetti & Cadiot 2000, Cadiot & Visetti, 20011) selon laquelle l'approche de la sémantique prépositionnelle qui alloue un statut privilégié à un sens spatial-configurationnel, volontiers identifié avec le sens littéral ou premier ainsi opposé aux sens dérivés ou seconds (e.g. récemment Lindstromberg 1997), n'est pas aussi convaincante qu'il y **paraît**. **Peut-être** doit on aller jusqu'à dire que c'est son évidence "externe", psychologique, qui la condamne comme explication interne.

Ces deux exemples sont d'une nature assez argumentative et la fonction interpersonnelle semble ici l'emporter sur la fonction informative. Il en va de même pour l'exemple 3, qui est tiré de la deuxième section où l'auteur présente une alternative à l'approche des prépositions. Dans cet exemple spécifique, l'auteur cherche à justifier son approche. Ce faisant, il fait preuve d'une certaine modestie (cf. notamment la première partie de la phrase), mais on observe une certaine harmonie entre d'un côté les marqueurs de modestie (marqués en gras) et de l'autre côté des marqueurs qui impliquent que ce que l'auteur a à dire est important et approprié (marqués en italiques) :

3. **Sans prétendre ni bien sûr que cette option soit vaine, ni que notre effort pour l'élargir soit entièrement neuf** (je pense par exemple aux travaux de C. Vandeloise qui mettent l'accent sur la dimension plus fonctionnelle des repérages spatiaux), **il me paraît essentiel** d'aller plus loin en *insistant sur le fait que* les emplois des prépositions spatiales sont conditionnés par des valeurs *tout autres* que configurationnelles:

L'emploi de l'adjectif *essentiel* et les autres marqueurs qui signalent la pertinence contrebalancent donc l'impression de modestie produite par *paraître* et la première partie de la phrase (cf. *supra* 7.2.5.1).

L'emploi de *paraître* combiné avec le pronom personnel de la première personne au singulier (ou bien au pluriel, cf. la seconde occurrence de *paraître* dans l'exemple suivant) aide l'auteur à se positionner et à se démarquer par rapport à la tradition. En même temps, ces éléments linguistiques font apparaître les idées de l'auteur comme une alternative parmi d'autres : l'auteur montre qu'il reste ouvert à d'autres opinions.

4. La sémantique prépositionnelle **me paraît** d'abord déterminée par des valeurs renvoyant à la dépendance, au contrôle, à l'appropriation réciproque, à l'anticipation et à l'attente, même s'il leur arrive aussi de s'effacer au profit de critères configurationnels, voire strictement géométriques. Loin de considérer, comme on le fait le plus souvent, les valeurs aspectuelles, subjectives et qualitatives comme des suppléments que la reconstruction linguistique devrait dériver dans un deuxième temps, il **nous paraît** qu'il faut les inscrire au cœur des "motifs" les plus originels attribués aux prépositions.

Après avoir présenté son point de vue de façon générale, l'auteur passe à une discussion de la sémantique de certaines prépositions spécifiques. Les observations lors de cette analyse sont utilisées pour étayer le point de vue défendu par l'auteur. Dans l'exemple ci-dessous, qui relève de la section sur la préposition *sur*, le modalisateur épistémique *paraître* et le marqueur sporadique *souvent* servent à atténuer la critique qui reste toutefois assez dure (cf. les termes *invalidier*, *fragiliser* et *réclamer des remaniements*) :

5. A nos yeux, ces exemples n'invalident pas seulement les explications en termes de spatialité ou de physicalité. Ils fragilisent également les explications en termes de schématisation topologique 'abstrait', qui **paraissent** souvent artificielles, et réclament en tout état de cause des remaniements qui font douter de leur qualité d'invariant.

Après avoir avancé cette critique, l'auteur conclut prudemment en faveur de sa propre théorie :

6. Ainsi l'instruction topologique, même purement configurationnelle et déspatialisée (c'est-à-dire déliée de l'espace perçu), **paraît** céder le pas devant un principe plus ouvert et plus riche de définition-délimitation de deux 'segments' ou 'phases' par le biais de leur 'mise en contact'.

Comme déjà noté, c'est la section qui porte sur *dans* qui contient le plus de modalisateurs épistémiques. Dans cette section, l'auteur discute la notion d'inclusion traditionnellement

associée à la préposition *dans*. Il commence par formuler une assertion assez évidente et le modalisateur *sans doute* paraît ici être utilisé pour des raisons stylistiques :

7. Que l'inclusion fasse partie du motif de *dans* n'étonnera **sans doute** personne.

Dans l'exemple suivant, l'effet atténuant de l'occurrence de *sans doute* est beaucoup plus net. Ici, l'auteur continue en problématisant la liaison *dans-inclusion* :

8. En même temps, de nombreux auteurs [...] soulignent que l'inclusion n'est pas la seule valeur fondamentale en cause. *Dans* entretient une affinité immédiate avec la dépendance, le contrôle, la causalité: [...] Dans le cas des emplois spatiaux, on doit invoquer une relation contenant/contenu, qui intègre précisément inclusion et dépendance. [...] C'est **sans doute** pour cette raison (la prégnance de la notion de dépendance à construire, ou encore de dépendance pour un état de phase) que la mise en jeu de l'inclusion obéit à des contraintes spécifiques.

Il tire ensuite quelques conclusions à partir de son raisonnement, en commençant par *il semble donc que...* (cf. *supra* 7.2.6)

9. Il **semble** donc que la relation d'inclusion portée par *dans* soit entièrement un effet de cette dynamique de mise en phase de la relation de dépendance. [...] L'inclusion n'en est donc qu'une dimension, **sans doute** plus éclipétique qu'on ne croit, comme l'**indiquent** les exemples suivants où l'on observe une certaine difficulté à stabiliser à chaque fois un espace bien défini qui supporte le schème topologique de l'inclusion.

L'auteur clôt la section sur *dans* en défendant son propre point de vue sur les prépositions :

10. Ces quelques observations sur ces "prépositions spatiales" d'excellence que sont SUR, SOUS, CONTRE, DANS suffisent **sans doute**, non à démontrer, mais au moins à faire comprendre notre propos.

La section intitulée Conclusion est dépourvue de modalisateurs épistémiques, mais comme nous l'avons vu, les idées principales de l'article ont été présentées de façon atténuée ailleurs dans l'article. Le fait que la conclusion soit sans marqueurs épistémico-modaux ne veut donc pas dire que les conclusions (les « résultats ») sont présentées sans atténuation.

Malgré la proportion relativement élevée de modalisateurs épistémiques dans Frling15, l'auteur n'apparaît pas particulièrement modeste ou prudent, du moins pas de notre point de vue norvégien. Cela peut s'expliquer par l'objectif de l'article qui est d'argumenter contre une autre théorie. Il en résulte que les remarques critiques ont nécessairement une place centrale dans le texte, et la présence de quelques modalisateurs épistémiques (que ceux-ci

soient directement liés à la critique ou non) permet d'harmoniser le ton de l'article. Pour la plupart des occurrences, c'est la fonction interpersonnelle qui domine : beaucoup d'entre elles signalent un positionnement de l'auteur (cf. par exemple les exemples 3 et 4) et contribuent à la persuasion du lecteur (cf. par exemple l'exemple 10).

En ce qui concerne les marqueurs choisis, l'auteur de Frling15 utilise principalement *paraître* (6 attestations), *sans doute* (5) et *sembler* (3). À part une occurrence de l'adjectif *vraisemblable* et une occurrence du subjonctif *soit*, avec la locution *il semble que*, tous les modalisateurs épistémiques font partie de ceux inclus dans l'analyse quantitative. On ne relève donc pas de préférences stylistiques particulières dans le choix de marqueurs que fait l'auteur de Frling15.

#### 7.6.4 L'article *Frmed11*

L'article *Frmed11* est rédigé par un groupe de trois médecins-chercheurs et a été publié en 2001 dans les *Annales de Médecine Interne*. Il a été retenu pour une étude de cas en raison de son nombre élevé de modalisateurs épistémiques : l'article contient 27 occurrences des marqueurs sélectionnés, et la fréquence relative est de 3.7 pour mille (cf. appendice 2). Dans le sous-corpus *Frmed*, il n'y a que deux articles qui le devancent en fréquence relative et aucun en fréquence absolue. L'article est assez long en comparaison des autres articles du même sous-corpus.

L'article *Frmed11* traite d'une maladie génétique appelée *l'adrénoleucodystrophie*. Après un résumé relativement long, l'objectif de l'article est énoncé vers la fin de l'introduction. Les auteurs visent à donner une description clinique de la maladie ainsi qu'une présentation de la thérapeutique envisageable. L'article est donc avant tout de caractère descriptif, et se distingue en cela du prototype de l'article en recherche médicale, qui est l'article expérimental. L'article *Frmed11* ne suit pas même la structure quasi-obligatoire du type IMRED, ce qui s'explique sans doute par le fait que ce plan vaut avant tout pour les articles expérimentaux et n'est pas nécessairement un plan approprié pour un article de synthèse tel que *Frmed11*.

Suite à un bref rappel historique, qui ne contient aucun modalisateur épistémique, les auteurs procèdent à la description clinique de la maladie. Les modalisateurs épistémiques abondent dans cette section (qui est d'ailleurs sous-divisée en un grand nombre de sous-sections plus ou moins longues), et il s'agit presque sans exception d'occurrences qui remplissent une fonction contenu : ils sont utilisés pour émettre des hypothèses, pour rapporter des résultats ou des conclusions d'études antérieures – résultats et conclusions qui ne sont probablement pas complètement confirmés, d'où l'atténuation – et pour suggérer des explications possibles. Ces usages se chevauchent souvent, comme le montre l'extrait 1, où la première proposition (*L'adrénomyélonéuropathie représenterait...*) rapporte un résultat d'une étude antérieure et la deuxième proposition (*mais cette proportion pourrait être...*) présente une supposition qui pourrait relever de l'étude citée, mais aussi des auteurs eux-mêmes ; aussi, les exemples donnés ci-dessous – qui ne constituent qu'une sélection parmi les multiples occurrences relevées – ne seront pas regroupés selon leur emploi spécifique, mais énumérés par ordre chronologique, selon leur ordre d'apparition :

1. L'adrénomyélonéuropathie représenterait environ 13 % des causes d'insuffisance surrénale primitive chez l'adulte [17], mais cette proportion **pourrait être** d'environ 20 % à 30 % dans la population masculine des pays développés où la tuberculose est rare. Une étude a

montré que les tests de stimulation classiques (test au synacthène) étaient insuffisants pour identifier les malades présentant une atteinte surrénale partielle compensée ; l'élévation plasmatique de l'ACTH **semble** représenter le marqueur le plus précoce du dysfonctionnement cortico-surrénalien [18].

2. L'impuissance a pu être exceptionnellement décrite comme le symptôme initial de l'adrénomyélongueuropathie [15], [19]. Par ailleurs, la fertilité **semble** diminuer à partir de la 3e décennie.
3. Les lésions sont les mêmes dans l'adrénoleucodystrophie et dans l'adrénomyélongueuropathie, mais le cortex surrénalien est plus atrophique dans l'adrénomyélongueuropathie que dans l'adrénoleucodystrophie, **probablement** parce que l'insuffisance surrénale y évolue depuis plus longtemps [6], [26].
4. Les cellules inflammatoires infiltrent le pourtour de la lésion en s'étendant vers son centre, ce qui **suggère** un phénomène secondaire plutôt que primitif.
5. Ces observations **suggèrent** que les mécanismes auto-immuns jouent un rôle dans la pathogénie des lésions dans la forme cérébrale de l'adrénoleucodystrophie.
6. Chez les malades atteints d'adrénoleucodystrophie comme chez les sujets normaux, le cholestérol estérifié par les AGTLC est un mauvais substrat pour les hydrolases. La proportion importante d'esters de cholestérol avec AGTLC dans l'adrénoleucodystrophie limite **probablement** leur capacité à servir de précurseurs pour les stéroïdes surrénaliens [19]
7. [L]e (ou les) autre(s) facteur(s) intervien (nen) t **probablement** en modulant la sévérité de la réponse inflammatoire. Powers et al. **pensent que** cette inflammation est secondaire à des mécanismes immunologiques médiés par les cytokines [33]. En effet, l'immunoréactivité révèle des taux élevés de TNF dans les macrophages et les astrocytes de la zone active des lésions. On **suppose que** la libération des lipides contenant les AGTLC constitue le facteur déclenchant de la cascade.

Dans un grand nombre de cas, la proposition modalisée est précédée ou suivie d'une remarque sur les connaissances limitées de la médecine sur cette maladie. Plutôt que d'évoquer un manque de connaissance personnel, les auteurs de Frmed11 signalent donc une lacune dans les connaissances de la science médicale en général, lacune touchant toute la profession. Les maintes occurrences du pronom *on* réfèrent ici à la communauté médicale (cf. les exemples 9, 10 et 11), et il en va de même pour les passifs (cf. les exemples 10-14). Les modalisateurs épistémiques remplissent clairement une fonction contenu, et traduisent même une véritable incertitude. Le fait qu'il s'agit d'une véritable incertitude est rendu très claire par les choix lexicaux (*pas d'explication claire, on ne peut déterminer précisément, on ne sait pas encore exactement* etc.). La présence des marqueurs épistémico-modaux dans de tels cotextes n'a donc rien à voir avec la politesse ou la modestie, et elle n'est définitivement pas

une stratégie de détour (cf. Foullioux et Tejedor de Felipe 2004 : 114) ; au contraire, ces marqueurs sont des éléments linguistiques employés pour des raisons d'honnêteté et de précision du propos :

8. *Il n'y a pas d'explication claire* à cette différence de répartition des phénotypes entre les populations française et américaine, d'une part, et hollandaise, d'autre part ; il existe **probablement** un dépistage plus précoce des porteurs sains aux Pays-Bas par rapport aux deux autres pays [10], [15], [22], [23].
9. Dans l'adrénoleucodystrophie, *on ne peut déterminer précisément laquelle des 2 voies prédomine. Il semble toutefois que ce soit* la synthèse endogène qui prime selon les essais thérapeutiques (voir : § "Traitement").
10. Par ailleurs, *on ne sait pas encore exactement de quelle façon* une anomalie de l'ALDP peut occasionner un dysfonctionnement de l'activité de la lignocéroyl CoA ligase dans le peroxyosome. **On suppose que** l'ALDP transporte la lignocéroyl CoA ligase dans le peroxyosome, ou est requise d'une façon ou d'une autre pour son fonctionnement. Elle **pourrait aussi être** nécessaire au transport des AGTLC ou à celui de leur dérivé Co-enzyme A dans le peroxyosome [29], [30]. Enfin, le gène codant pour l'enzyme lignocéroyl CoA ligase *n'est pas connu*.
11. Puisque l'accumulation des AGTLC est la principale anomalie biochimique dans l'adrénoleucodystrophie, **on suppose que** c'est cette accumulation qui est responsable de la pathologie. *Il n'en existe pas actuellement de preuve directe*, et le mécanisme de l'altération des tissus *n'est pas encore bien établi* en particulier en ce qui concerne le tissu nerveux.
12. La pathogénie de l'atteinte des cellules du testicule est *mal connue*, mais elle est **probablement** similaire à celle proposée dans l'insuffisance surrénale,
13. La pathogénie des lésions du système nerveux est complexe et encore *mal connue* ; **il est probable que** l'augmentation des taux d'AGTLC altère la structure membranaire et le fonctionnement des cellules nerveuses.
14. Le mécanisme de l'augmentation de la bioactivité du TNFalpha *n'est pas encore connu*. D'autres auteurs **proposent** l'intervention de mécanismes pathogéniques auto-immuns pour expliquer ce phénomène inflammatoire [5]. En conclusion, alors qu'il **semble** évident que les AGTLC jouent un rôle dans la pathogénie de l'adrénoleucodystrophie, que le gène de l'adrénoleucodystrophie a été récemment identifié, et que l'étude de la protéine de membrane du peroxyosome pour laquelle il code permet de grands progrès dans la compréhension des mécanismes physiopathologiques, l'accumulation des AGTLC ne peut à lui seul tout expliquer.

La section qui porte sur les traitements n'est pas aussi riche en atténuateurs, mais ce sont les mêmes usages qui reviennent. Dans l'exemple suivant, le choix du verbe *suggérer* sert à présenter les résultats d'une étude de manière atténuée. Cela est particulièrement



important, étant donné qu'il s'agit des résultats préliminaires d'une étude toujours en cours. Ces résultats sont toutefois corroborés par ceux d'une autre étude :

15. Dans une étude en cours incluant 135 malades atteints d'adrénomyélonéuropathie, les résultats préliminaires **suggèrent** que le traitement ne ralentit pas la progression de la maladie, malgré la normalisation des taux plasmatiques, ce qui correspond également aux résultats de l'étude effectuée par Aubourg et al. [38]

Dans l'exemple 16, les auteurs tirent une conclusion (introduite par *il semble que*) à partir des résultats issus d'une étude antérieure. Ils ne considèrent cependant pas les résultats comme confirmés (ce qu'ils disent explicitement), et il est donc nécessaire de les présenter d'une façon atténuée :

16. On n'observe pas non plus d'amélioration des fonctions surrénale et testiculaire sous traitement par l'huile de Lorenzo. En revanche, **il semble que** l'administration de GTO-GTE avant l'apparition des premiers symptômes de la maladie ralentisse sa progression; en effet, les premiers résultats de l'étude menée par Moser sur 52 malades asymptomatiques suivis pendant 33 mois, montrent que 10 % d'entre eux développent une atteinte neurologique sévère, comparés aux 30 % prévus approximativement sur les données de l'histoire naturelle de l'adrénoleucodystrophie [5], [10], [36], [39], [40]. Une étude plus longue et portant sur un plus grand nombre de malades est indispensable avant de conclure, et est actuellement en cours.

Comme c'était le cas pour la section sur la description clinique, plusieurs des modalisateurs épistémiques dans la section sur la thérapeutique sont précédés ou suivis d'une remarque sur les connaissances limitées de la science médicale :

17. Bien que le renouvellement des cellules de la microglie soit **mal connu**, **il semble que** des monocytes provenant des cellules hématopoïétiques périphériques franchissent la barrière hémato-encéphalique, pénètrent dans le parenchyme cérébral et se différencient en cellules microgliales.

Il convient de noter que la présentation de résultats dans Frmed11 n'est pas systématiquement affectée d'une modalité épistémique d'incertitude, même si les exemples donnés jusqu'ici peuvent donner cette impression. Certains résultats sont introduits avec un verbe factif tel que *montrer* plutôt qu'avec des verbes de croyance comme *penser* et *suggérer*. Ainsi, les exemples cités ci-dessous, où l'emploi du verbe *montrer* indique que le contenu informatif transmis est considéré comme confirmé ou prouvé s'opposent à ce que l'on a vu par exemple dans 7 et 15. Cette variation dans la manière de présenter des résultats suggère que l'emploi des modalisateurs épistémiques est un choix conscient ; les modalisateurs ont un véritable contenu et ne sont pas là uniquement pour des raisons stylistiques.

18. Les résultats **montrent** que les AGTLC représentent environ 0,1 % des acides gras totaux dans le plasma pathologique contre 0,01 à 0,02 % dans le plasma normal.
19. Knazek et al. **ont montré** que la microviscosité de la membrane des érythrocytes des malades atteints est augmentée à cause de la présence des AGTLC qui entraînent une rigidité membranaire [31].

Cependant, plutôt que de refléter une opposition entre résultats confirmés et résultats non confirmés, l'opposition entre *montrer* et *suggérer / penser* etc. reflète une opposition entre la présentation « brute » de résultats d'un côté et une interprétation de ces résultats de l'autre (cf. la distinction que fait Skelton 1997 entre *evidential* et *interpreted truth*). Cette opposition est très claire dans les exemples suivants, où *montrer* est utilisé pour rapporter les résultats en eux-mêmes et *suggérer / probablement* pour rapporter ou proposer des interprétations de ces mêmes résultats :

20. Des études récentes **ont montré** que les macrophages dans cette zone de démyélinisation, et également les astrocytes, contenaient de fort taux de TNF alfa (tumor necrosis factor alfa) [27]. Les cellules périvasculaires sont constituées de lymphocytes T4, T8, monocytes et macrophages (avec prédominance des T4). Ces observations **suggèrent** que les mécanismes auto-immuns jouent un rôle dans la pathogénie des lésions dans la forme cérébrale de l'adrénoleucodystrophie.  
(cf. l'exemple 5)
21. On **a montré** que l'administration d'acides gras mono-insaturés diminuait la synthèse AGTLC, **probablement** par compétition au niveau du système enzymatique entre les précurseurs saturés et insaturés. Ces observations sont également retrouvées dans l'adrénoleucodystrophie.
22. Les essais thérapeutiques effectués **montrent** que l'administration d'acides gras mono-insaturés est plus efficace pour diminuer les taux d'AGTLC que la restriction diététique seule de ces acides, ce qui **suggère** que la synthèse endogène est une source plus importante d'AGTLC que l'apport diététique lui-même.
23. Une publication récente **a montré** que l'addition de 4-phénylbutyrate à des cultures de fibroblastes de malades atteints ou de souris malades normalise les taux d'AGTLC en 6 à 12 jours et augmente leur oxydation [47]. Le 4-phénylbutyrate augmente de 3,4 fois l'expression d'une protéine homologue à la protéine de transfert défectueuse dans l'adrénoleucodystrophie. Les auteurs **suggèrent** qu'il s'agit d'un exemple de "thérapie génique pharmacologique".

En ce qui concerne le choix des marqueurs, les auteurs de Frmed11 utilisent notamment le verbe *suggérer* (10 attestations) (dans la grande majorité des cas avec un sujet non animé), mais aussi le verbe *sembler* (7), l'adjectif *probablement* (6) et la forme verbale *pourrait* (3). Alors que *suggérer* s'utilise principalement pour présenter des hypothèses et des

résultats d'une manière atténuée, *probablement* sert avant tout à avancer des explications possibles. Il y a peu de marqueurs dans cet article qui n'aient pas été répertoriés dans la section 6.4, hormis quelques occurrences des verbes *supposer* et *penser* (cf. les exemples 7, 10, 11) et du conditionnel épistémique (cf. l'exemple 1).

Au niveau des fonctions pragmatiques, le nombre d'occurrences qui remplissent clairement une fonction contenu est frappant. La fréquence élevée de marqueurs assumant cette fonction s'explique sans doute par le sujet de l'article : il s'agit d'une maladie dont la médecine ne sait pas grand chose encore, la recherche semble en être à ses débuts et de nombreuses questions restent non résolues.

### 7.6.5 Bilan

Les études de cas menées ici montrent comment les modalisateurs épistémiques peuvent contribuer à la structure argumentative d'un article. Leur rôle dans l'argumentation est plus manifeste lorsqu'ils assument une fonction interpersonnelle, car ils visent alors, entre autres, à convaincre le lecteur et à promouvoir la recherche de l'auteur. Les modalisateurs remplissant cette fonction sont plus fréquents dans les deux articles de linguistique et on observe en fait une différence sensible entre l'emploi de modalisateurs épistémiques dans les articles de linguistique (Noling05 et Frling15) et les articles de médecine (Nomed08 et Frmed11). Dans ces derniers, c'est la fonction contenu qui domine, et il s'agit souvent d'une véritable incertitude – rendue plus explicite encore par des remarques évoquant des connaissances limitées de la science médicale en général. Dans les articles de linguistique, en revanche, on note un emploi mixte des deux principaux types de fonctions ; les occurrences semblent cependant majoritairement remplir une fonction interpersonnelle. La différence d'emploi des modalisateurs épistémiques entre les deux disciplines est particulièrement nette entre Frmed11 et la deuxième section de Noling05. Les modalisateurs épistémiques dans Frmed11 illustrent bien la fonction contenu : ils y sont utilisés pour exprimer une incertitude réelle et pour émettre les réserves d'usage en présentant les conclusions et les résultats. Les facteurs justifiant ces marqueurs combinent probablement le souci d'être précis et honnête et des conventions du genre. Dans la deuxième section de Noling05, la majorité des marqueurs épistémico-modaux présentent un caractère bien plus rhétorique et interpersonnel : l'auteur fait des réserves pour protéger l'autre (atténuation de critiques) ou pour convaincre le lecteur (par exemple en argumentant contre d'éventuelles remarques critiques qu'il anticipe). Il peut être tentant de caractériser le premier article comme typique à cet égard du discours médical et le second comme typique chez les linguistes, mais les choses ne sont évidemment pas si

simple. Comme nous l'avons vu, les linguistes utilisent aussi des modalisateurs épistémiques avec une fonction contenu et les médecins-chercheurs les emploient parfois avec une fonction plus rhétorique ou persuasive. En outre, il ressort de ces études de cas que le sujet d'un article et la méthode appliquée jouent un rôle important non seulement pour le nombre de modalisateurs épistémiques employés, mais aussi pour la fonction que sert leur utilisation. Il est clair qu'un article dont l'objectif est de contester une théorie contiendra plus de modalisateurs assumant une fonction interpersonnelle qu'un article dont l'objectif est de fournir une synthèse sur un phénomène encore peu investigué.

## 8 CONCLUSION

Dans cette thèse, nous avons examiné l'emploi de modalisateurs épistémiques d'incertitude dans des articles de recherche en adoptant une approche doublement comparative : la problématique principale consistait à analyser dans quelle mesure il existe à cet égard des différences entre 1) des articles rédigés en différentes langues et 2) des articles relevant de différentes disciplines. Trois langues (anglais, français, norvégien) et deux disciplines (linguistique, médecine) ont été comparées. Dans le présent chapitre, nous récapitulerons d'abord les observations principales résultant des analyses quantitatives et qualitatives (8.1), nous proposerons ensuite quelques pistes pour des recherches ultérieures (8.2) et nous finirons par quelques mots sur les applications possibles des études de ce type (8.3).

### 8.1 Synthèse

Dans l'introduction, nous avons posé, à l'instar du projet KIAP, la question de savoir si les identités culturelles dans le discours scientifique sont principalement liées à l'appartenance langagière ou à l'appartenance disciplinaire des chercheurs. A la suite des analyses quantitatives et qualitatives que nous avons effectuées, nous pouvons conclure que pour le phénomène de la modalité épistémique, il n'est pas possible de répondre à cette question de manière univoque, car les deux paramètres ne sont pas d'importance égale pour les deux niveaux d'analyse (quantitatif et qualitatif). Au niveau quantitatif, c'est le paramètre de la langue qui l'emporte clairement sur la discipline: la fréquence de modalisateurs épistémiques présente une différence statistiquement significative entre d'une part les articles anglais et norvégiens et de l'autre les articles français. Les résultats montrent que les auteurs d'articles français recourent moins à la modalisation épistémique d'incertitude que leurs collègues anglophones et norvégiens. Les auteurs norvégiens se situent entre les francophones et les anglophones, mais s'avèrent beaucoup plus proche des anglophones, ce qui corrobore l'observation faite par l'équipe KIAP, selon laquelle le discours scientifique norvégien est à bien des égards relativement similaire au discours anglo-américain, du moins pour les phénomènes étudiés dans ce projet. Par ailleurs, il est possible que le taux de modalisateurs épistémiques obtenu dans notre étude pour le corpus norvégien soit sous-estimé : les études de cas et l'analyse exploratoire attestent que d'autres marqueurs, outre ceux retenus, sont assez souvent utilisés dans le corpus norvégien. Les études de cas et l'analyse exploratoire permettent ainsi d'avancer que la sélection de marqueurs est vraisemblablement plus exhaustive et plus représentative pour le français et l'anglais que pour le norvégien. Si tel est

le cas, les articles norvégiens sont alors plus proches encore des articles en anglais que ne l'indique la présente étude. Quoi qu'il en soit, la différence la plus nette est celle qui distingue les auteurs francophones d'un côté et les auteurs norvégiens et anglophones de l'autre, en ce sens que les auteurs francophones utilisent significativement moins de modalisateurs épistémiques que les deux autres groupes. Ce constat suggère que le phénomène d'atténuation, si souvent présenté comme un trait caractéristique du discours scientifique, n'est pas nécessairement un trait saillant de tout discours scientifique, bien qu'il le soit, indiscutablement, pour le discours scientifique anglo-américain.

Entre les deux disciplines en revanche, aucune différence forte ne se manifeste au niveau de la fréquence des marqueurs. Une comparaison entre deux disciplines ne permet évidemment pas de conclure de façon décisive en faveur de l'hypothèse avancée par Markkanen et Schröder (1997 : 10), selon laquelle les différences entre sciences naturelles et sciences humaines et sociales ne sont pas si grandes qu'on ne l'a traditionnellement cru. Comme nous l'avons mentionné, les disciplines investiguées ne sont pas les représentants les plus typiques de leurs domaines scientifiques respectifs, et un autre choix de disciplines aurait pu donner lieu à des résultats plus contrastés. Cependant notre étude montre qu'en ce qui concerne le phénomène de l'atténuation, les différences entre langues sont aussi importantes, peut-être même plus, que les différences entre disciplines. Nous avons toutefois mis en évidence une interdépendance importante des deux paramètres : la différence de fréquence des marqueurs est plus saillante pour le corpus linguistique que pour le corpus médical. Les différences entre les langues sont donc plus nettes en linguistique qu'en médecine, ce qui peut s'expliquer par le fait que la médecine est une discipline gouvernée par des normes hautement standardisées qui dépassent souvent les frontières linguistiques et nationales. Les articles des médecins-chercheurs francophones de notre corpus présentent un taux de modalisateurs épistémiques qui se situe entre celui des articles linguistiques francophones et celui des articles des chercheurs norvégiens et anglophones. Nous pouvons en conclure que les médecins-chercheurs francophones sont probablement plus influencés par le discours scientifique anglo-américain que ne le sont les linguistes francophones. Il est toutefois clair qu'ils appartiennent à une culture scientifique française, en ce qu'ils emploient considérablement moins de modalisateurs épistémiques que leurs collègues anglophones et norvégiens.

La fréquence des modalisateurs épistémiques est donc à peu près la même dans les deux disciplines, alors qu'il y a une différence nette entre les langues. En ce qui concerne le niveau qualitatif, on observe en revanche que les fonctions pragmatiques que ces

modalisateurs remplissent semblent varier davantage entre les disciplines qu'entre les langues. Nous avons vu que dans la discipline de médecine, la grande majorité des modalisateurs épistémiques assument une fonction « contenu », c'est-à-dire qu'ils portent sur le contenu textuel lui-même, plutôt que d'être orientés vers le lecteur. Ils servent à introduire des hypothèses et des interprétations de résultats et à atténuer des conclusions. Ces usages sont fréquents en linguistique aussi, mais on y trouve en outre un nombre considérable de modalisateurs qui remplissent une fonction clairement interpersonnelle : une partie importante des modalisateurs servent à atténuer les critiques formulées à l'encontre des recherches d'autres ou à faire preuve de prudence en interprétant ces recherches. De surcroît, certains modalisateurs entrent dans des séquences argumentatives où l'auteur répond à d'éventuelles remarques critiques qu'il envisage à l'avance – et c'est dans cet usage que l'aspect persuasif se manifeste la plus explicitement. Les modalisateurs qui remplissent une fonction clairement interpersonnelle ne sont pas absents du corpus médical, mais ils y sont beaucoup plus rares que dans les articles de linguistique. Il convient cependant de noter que cette observation se base principalement sur une étude systématique de six marqueurs uniquement (trois pour chacune des langues, à savoir *sembler*, *paraître*, *pouvoir*, *synes*, *se ut til/som* et *kunne*). Cela étant, ce sont les marqueurs les plus fréquents, et des observations similaires ont par ailleurs été faites pour le corpus anglais (Vold 2006b).

Les différences de fonctions des marqueurs relèvent sans doute de pratiques argumentatives différentes. Bien que tout article scientifique ait entre autres visées celle de convaincre, il est clair que cet aspect se manifeste différemment et de manière plus ou moins explicite selon les cultures scientifiques (cf. Breivega 2003, Fløttum et al. 2006). Les normes de l'argumentation ne sont pas les mêmes dans les deux disciplines investiguées, et cela se reflète dans l'emploi des modalisateurs épistémiques. Dans les articles de linguistique, où l'argumentation ouverte et la critique explicite envers les travaux d'autres collègues sont des stratégies courantes, le nombre d'atténuateurs assumant une fonction interpersonnelle et associables à la notion de politesse est plus élevé.

Comme nous l'avons mentionné plus haut, nous n'avons pas observé de différences nettes entre les langues examinées (le français et le norvégien) au niveau des fonctions des marqueurs. Les différences interlinguistiques dans l'emploi de modalisateurs épistémiques jouent donc sur les fréquences mais non sur les cotextes d'emploi et leurs effets pragmatiques.

Pour répondre de façon succincte à la question centrale de notre étude, on peut résumer les résultats comme suit :

- En ce qui concerne l'aspect quantitatif, c'est le paramètre de la langue qui est le plus important.
- Pour ce qui est de l'aspect pragmatico-fonctionnel, c'est le paramètre de la discipline qui influence le plus l'usage des marqueurs épistémico-modaux d'incertitude.

En ce qui concerne enfin le variable du sexe de l'auteur, nous n'avons pas trouvé de différences significatives dans la fréquence des modalisateurs épistémiques. La conception traditionnelle selon laquelle l'atténuation serait une caractéristique plus marquée du discours féminin ne peut donc pas être corroborée.

Si nous revenons aux trois questions principales posées par le projet KIAP, les analyses du phénomène investigué nous permettent d'apporter quelques éléments de réponse<sup>1</sup>. La première question concerne la voix de l'auteur : dans quelle mesure et de quelle manière l'auteur se manifeste-t-il dans son texte ? Comme nous l'avons mentionné dans l'introduction, un modalisateur épistémique est une trace de l'auteur, en ce qu'il exprime un jugement émis par ce dernier. Cela étant, par rapport à des traces énonciatives explicites telles que les pronoms personnels de la première personne, les modalisateurs épistémiques sont des traces moins immédiatement reconnues comme telles. Ils acquièrent un caractère plus personnel dès qu'ils sont utilisés en association avec un pronom personnel de la première personne. Cependant, l'analyse quantitative et les analyses du cotexte immédiat permettent de conclure que les auteurs préfèrent le caractère moins directement personnel de ces traces, et que cela vaut notamment pour les médecins-chercheurs. D'abord, les analyses exploratoire et quantitative du corpus ont révélé qu'il y a en général peu de marqueurs privilégiant l'emploi de la première personne – des exceptions notables avec l'usage fréquent de *anta* et *assume* ('supposer') se manifestent cependant dans les corpus Noling et Engling. Dans le corpus anglais, nous avons repéré une différence notable entre les disciplines dans le choix de marqueurs : les médecins-chercheurs préfèrent des marqueurs « objectivants » (par exemple *may*, *could*) alors que les linguistes utilisent aussi des marqueurs plus subjectifs, c'est-à-dire des marqueurs qui présupposent, de par leur sémantisme même, une évaluation personnelle (par exemple *seem*, *assume*). Par ailleurs, un examen du cotexte des marqueurs les plus fréquents révèle que la source du jugement est rarement explicitée ; on peut avancer des

---

<sup>1</sup> Soulignons qu'il ne s'agit que d'éléments de réponse, car la modalité épistémique n'est qu'un aspect parmi d'autres qui peuvent contribuer à des réponses à ces questions. Pour de plus amples développements, voir Fløttum et al. 2006.



raisons « traditionnelles » d'impersonnalisation et d'objectivation – ce qui correspond à l'explication admise pour l'emploi de *suggest* avec des passifs et des sujets non animés chez les médecins-chercheurs (cf. *supra*, 6.4.2.2). Cependant, pour ce qui est des marqueurs retenus pour les analyses qualitatives, la raison peut simplement être qu'une telle explicitation serait superflue (cf. *synes, se ut (supra 7.3.3.3)*) ou amènerait à un changement non souhaité dans le marquage du degré de certitude (cf. *sembler (supra 7.2.5.3)*). En bref, nous pouvons conclure que dans la plupart des cas, les modalisateurs épistémiques restent des traces relativement indirectes de la voix de l'auteur, et que l'on tend à éviter – du moins en médecine – la mise en relief de la subjectivité du jugement. En linguistique, la mise en relief de la subjectivité du jugement semble plus acceptable : la source du jugement est parfois rendue explicite par la combinaison de marqueurs épistémico-modaux avec des pronoms personnels de la première personne (par exemple *vi antar* ('nous supposons'), *I assume, we suggest* et *il me semble*).

Pour ce qui est des différences interlinguistiques, nous avons vu qu'il y a un haut degré de correspondance entre les trois langues en ce qui concerne les types de marqueurs utilisés – on observe cependant que les auxiliaires modaux ont plus de poids comme marqueurs épistémiques dans les corpus anglais et norvégien que dans le corpus français. Par ailleurs, la présence des verbes de cognition *anta* et *assume* dans les corpus norvégien et anglais ainsi que l'absence d'un item correspondant dans la liste des marqueurs français pourraient indiquer que la source des jugements épistémiques est rendue explicite plus souvent dans les articles anglais et norvégiens que dans les articles français. Le cas échéant, cela corroborerait les observations du projet KIAP, selon lesquelles les auteurs anglophones et norvégiens se manifestent plus explicitement dans leurs textes que ne le font les auteurs francophones. Il faut cependant souligner que les listes de marqueurs fréquents élaborées dans cette étude ne suffisent pas à conclure de manière définitive que la subjectivité dans la modalisation est moins saillante dans les articles français que dans les autres. On pourrait en effet arguer qu'elle est également manifeste dans les articles français, où *sembler*, qui sémantiquement présuppose une instance modalisante, est le premier marqueur épistémico-modal, alors que dans les articles anglais et norvégiens, ce sont les marqueurs « objectifs » *may* et *kan* qui sont les plus utilisés (sauf dans le corpus Engling, où *seem* est le marqueur le plus fréquent). Dans le même sens, l'examen du cotexte des six marqueurs soumis à des analyses qualitatives ne permet pas de conclure à des différences de mise en relief de la subjectivité du jugement entre le norvégien et le français. (cf. *supra 7.5*).

La deuxième question posée par le projet KIAP concerne les voix d'autres chercheurs : dans quelle mesure les points de vue (ou voix) d'autres chercheurs se manifestent-ils ? Nous avons commenté cette question au chapitre 7, où nous avons vu que l'aspect interactif de la modalisation épistémique est beaucoup plus saillant dans les articles linguistiques que dans les articles médicaux. Cela est dû au fait que les linguistes commentent, argumentent contre et critiquent de façon explicite les idées d'autres chercheurs plus souvent que ne le font les médecins-chercheurs. Cette dimension polémique les oblige à une certaine réserve afin de préserver leur propre face et celle des autres. La conception traditionnelle de l'atténuation comme stratégie de politesse semble donc davantage valoir pour la linguistique que pour la médecine. Cependant, l'aspect interactif se manifeste également dans les cas où la modalisation épistémique est liée à la présentation de résultats issus d'études antérieures. Evidemment, le fait de faire référence à des études antérieures est une manière d'inclure, voire d'évaluer, les points de vue d'autres chercheurs. Il s'agit là d'une stratégie commune aux médecins-chercheurs et aux linguistes, et nous pouvons donc conclure que la modalisation liée aux points de vue d'autres se présente de manière plus diversifiée dans les articles de linguistique.

La troisième question du projet KIAP concerne, comme la première, la voix de l'auteur. Comment l'auteur promeut-il sa recherche ?<sup>2</sup> Nous avons vu qu'il est plus légitime en linguistique qu'en médecine de promouvoir sa propre recherche en argumentant explicitement en faveur de son point de vue et en avançant des remarques critiques contre les points de vue d'autres chercheurs. Cela explique le taux plus élevé en linguistique de modalisateurs « persuasifs » que le linguiste utilise souvent pour défendre son point de vue en avançant des remarques critiques potentielles, et de modalisateurs de politesse qu'il utilise pour atténuer ses propres remarques critiques contre d'autres. En médecine, la promotion de la recherche ne se fait pas de manière aussi explicite. Les multiples modalisateurs épistémiques liés à l'interprétation des résultats et à la présentation des limites méthodologiques, à priori davantage tournés vers le contenu que vers le lecteur, contribuent cependant à l'image d'un chercheur sérieux et honnête qui a pris en compte tout facteur susceptible d'intervenir et toute source d'erreur possible.

Une fois formulées ces remarques générales sur les langues et les disciplines, il convient de rappeler l'importante variation individuelle observée au cours de notre étude. Au

---

<sup>2</sup> Rappelons encore une fois que notre réponse vaut uniquement pour ce qui se rapporte au phénomène de l'atténuation – la promotion de la recherche par d'autres moyens linguistiques, par exemple à l'aide des adjectifs évaluatifs comme *important*, *intéressant* etc., n'a pas été examinée dans la présente étude.

niveau de la fréquence des marqueurs, nous avons vu que la variation individuelle est d'importance égale dans les deux disciplines, mais au niveau fonctionnel, les articles de linguistique semblent davantage varier entre eux que les articles de médecine. Les modalisateurs à fonction « contenu » semblent dominer dans tous les articles de médecine, alors que dans le corpus de linguistique, certains articles, écrits sur un ton argumentatif et persuasif, contiennent de nombreux modalisateurs à fonction clairement interpersonnelle, tandis que d'autres adoptent un style plus neutre. L'hypothèse posée au début, selon laquelle la variation intragroupe serait plus importante dans le corpus linguistique, est donc partiellement validée par ces observations.

## **8.2 Pistes pour des recherches ultérieures**

Si la présente étude permet de tirer un certain nombre de conclusions, elle montre aussi que plusieurs questions restent à explorer de plus près. D'abord, la question de savoir quel paramètre détermine le plus l'écriture de l'article, celui de discipline ou celui de langue, ne saurait être entièrement résolue par l'étude de deux disciplines et de trois langues. On pourrait par exemple supposer que les différences disciplinaires sont plus importantes que ne le suggère notre étude, car les deux disciplines choisies – bien qu'elles relèvent de deux domaines scientifique différents – sont plus proches que d'autres ne le sont. Cette question mériterait d'être traitée d'une manière plus complète, en tenant compte d'autres langues et d'autres disciplines – et pourrait être élargie même à la question des sous-disciplines (cf. *supra*, 6.4.2.1). L'intérêt d'une comparaison de différentes sous-disciplines reste cependant à discuter. De notre point de vue, ce sont les différences linguistiques et culturelles qui demandent le plus à faire l'objet de nouvelles études, car c'est ici que le manque d'analyses est le plus flagrant et que le risque de malentendus et de problèmes de communication semble être le plus grand. En raison de la mondialisation des recherches scientifiques, la communication se fait pour beaucoup à travers les frontières nationales et culturelles, ce qui nécessite une compréhension entre les groupes (cf. *infra* 8.3). Les études sur les différences disciplinaires se justifient bien entendu aussi, par la pluridisciplinarité, qui connaît ces dernières années un essor considérable et appelle donc à mieux comprendre les proximités entre disciplines et leurs spécificités.

En ce qui concerne le paramètre du sexe de l'auteur, nous n'avons étudié que l'aspect quantitatif, par rapport auquel nous n'avons pas trouvé de différences significatives. Selon Holmes (1986) cependant, les différences entre sexes tiennent principalement au fait que les

hommes et les femmes tendent à utiliser des atténuateurs dans des buts différents ; c'est-à-dire que c'est l'aspect pragmatico-fonctionnel qui varie le plus. Afin de répondre de façon plus claire à la question des différences de sexe, on devrait donc étudier de manière plus approfondie les fonctions pragmatiques des atténuateurs employés par les auteurs féminins et masculins.

Par ailleurs, nous avons étudié l'usage de modalisateurs épistémiques d'incertitude, mais nous n'avons qu'effleuré la question de l'interaction de ce type de marqueurs avec des marqueurs de certitude ou avec ce que Fløttum et al. (à paraître 2007) appellent les verbes de positionnement (*argue, claim*), qui sont des verbes signalant la présence d'un auteur affirmatif et polémique. La méthode la plus pertinente pour aborder une telle problématique est évidemment une analyse détaillée de textes individuels, plutôt que l'étude d'un large corpus. Nous avons effleuré ces questions dans la section 7.6 des études de cas, mais elles méritent nettement des études plus approfondies, susceptibles d'améliorer notre compréhension des rapports qu'entretiennent les différents types d'éléments linguistiques et leur influence sur l'interprétation d'un texte dans son ensemble.

Enfin, d'autres paramètres pourraient être étudiés, par exemple celui du genre d'écrit auquel on a affaire et celui de l'âge de l'auteur ou de son statut dans sa carrière. Les tendances observées pour les articles de recherche ne valent pas nécessairement pour d'autres genres scientifiques. La présente thèse (qui est rédigée par une norvégienne, inévitablement influencée en cela par le discours anglo-américain) contient probablement plus d'atténuation qu'un article français, mais la question de savoir si elle contient beaucoup d'atténuateurs par rapport aux thèses françaises rédigées par des locuteurs natifs reste à vérifier. Il est raisonnable de supposer que les thèses, qui sont un genre généralement rédigé par de jeunes chercheurs au début de leur carrière scientifique, contiennent plus d'atténuation que les articles de revue rédigés par des chercheurs confirmés (pour en savoir plus sur la spécificité des pratiques d'écriture des néophytes, voir Rinck 2006).

### **8.3 Applications pratiques**

Cette thèse s'inscrit dans le champ des recherches sur le discours scientifique, qui connaît un essor important depuis quelques décennies. Nous aimerions conclure en disant quelques mots des applications pratiques que pourraient avoir ce type de recherches en général et la recherche présentée dans cette thèse en particulier. La recherche sur le discours scientifique nous aide à comprendre ce qu'est la science, et comment se fait la construction, la

transmission et la négociation du savoir. Il est désormais bien établi que la conception traditionnelle de la science comme « objective » et « neutre » est trop simpliste. Les « vérités scientifiques » sont rarement découvertes ou démontrées de manière immuable ; elles ont le statut d'hypothèses jusqu'à ce que la communauté scientifique concernée les accepte (à des degrés divers) comme vraies. La construction des savoirs dépend probablement plus de négociations entre chercheurs et d'échanges entre pairs qu'on n'a pu traditionnellement le penser (comme le montre Breivega 2003). Et pour faire reconnaître sa recherche dans un monde scientifique qui devient de plus en plus compétitif, il ne faut pas seulement être en mesure de présenter des résultats solides et originaux, mais posséder aussi le savoir-faire nécessaire pour les transmettre d'une manière convaincante.

Afin d'enseigner aux étudiants ou aux chercheurs novices ce type d'habiletés, des cours de rédaction scientifique ont été mis en place dans beaucoup de pays. Cette tradition est relativement récente, du moins en Norvège et en France et pour ce qui concerne les sciences humaines. Ces cours visent à expliciter ce qui jusque là est resté une sorte de curriculum caché – les normes tacites que les étudiants étaient supposés acquérir au cours de leur carrière, mais qui n'ont jamais fait partie de l'ordre du jour officiel en classe. Pour le courant rhétorico-linguistique (surtout anglo-américain), c'est la visée didactique qui constitue la première justification des recherches sur le discours scientifique. Ces recherches permettent d'explicitier les normes de rédaction scientifique qui ont restées implicites non seulement pour les étudiants, mais sans doute aussi pour les enseignants-chercheurs.

Dans cette perspective, les études comparatives sont particulièrement utiles, car la rédaction scientifique dans une langue étrangère est toujours plus difficile que celle en langue maternelle, a fortiori quand on sait que des différences culturelles existent dans la rédaction d'un genre spécifique et en l'occurrence d'un article de recherche. Plusieurs études ont souligné le besoin de faire de l'usage des atténuateurs un chapitre des cours de rédaction scientifique pour les étudiants étrangers (Hyland 2000, Markkanen et Schröder 1997). Mauranen (1993 : 252) a observé que le simple fait d'exposer les auteurs à des textes de la langue cible n'est pas suffisant pour que ceux-ci adoptent les normes propres à la langue cible, et une explicitation des normes semble donc requise. Cependant, si l'on se propose d'enseigner les normes de l'atténuation, le risque est de devenir trop prescriptif, alors que dans les usages on observe une grande diversité. La meilleure solution est probablement de sensibiliser les étudiants au phénomène de l'atténuation, à la fois au niveau de ses manifestations linguistiques et de ses effets pragmatiques, par exemple à travers des exercices d'analyse de texte (voir aussi Skelton 1997 : 135 et Markkanen et Schröder 1997 : 12).

Le fait de posséder des connaissances sur les différences linguistiques et culturelles est important non seulement pour les étudiants et les chercheurs qui présentent et publient dans une langue étrangère, mais aussi pour ceux qui évaluent : c'est-à-dire pour la communauté scientifique concernée en général et pour certains groupes en particulier, par exemple les éditeurs de revues internationales, qui reçoivent beaucoup de textes écrits par des chercheurs qui ne possèdent pas toujours des connaissances approfondies sur la pratique d'écriture et le style rhétorique de la langue en question. Une meilleure connaissance des différences linguistiques et culturelles chez l'éditeur est donc nécessaire pour éviter les malentendus et, plus grave encore, le refus de recherches importantes, basé sur un fondement inapproprié.

## BIBLIOGRAPHIE

- Andersen, A.L. 2003. *Oversettelse av modalpartiklene JO, NOK og VEL til italiensk. En korpusbasert kontrastiv studie med et psykolingvistisk tilsnitt*. Hovedoppgave, Universitetet i Oslo.
- Bally, C. 1965. *Linguistique générale et linguistique française*. Berne : Editions Francke.
- Banks, D. 1994. Hedges and how to trim them. In Brekke, M., Ø. Andersen, T. Dahl & J. Myking (eds.): *Applications and implications of current LSP research*. Proceedings of the 9<sup>th</sup> European LSP symposium. Bergen : Universitetet i Bergen, 587-593.
- Barnes, J. 1984. Evidentials in the Tuyuca verb. *International Journal of American Linguistics* 50, 255-271.
- Bazerman, C. 1988. *Shaping written knowledge: The genre and activity of the experimental article in science*. Madison, Wis. : The University of Wisconsin Press.
- Beaufrère-Bertheux, C. 1997. L'anglais de la recherche médicale : une grande diversité. *ASp (Anglais de Spécialité)* 15–18, 223-238.
- Bentley, D. 2003. Sur la force d'une approche non-dérivationnelle de l'analyse linguistique : quelques données de l'italo-roman. *Cahiers du CRISCO* 13. Université de Caen. Disponible sur <http://elsap1.unicaen.fr/cahiers/> (Consulté le 25/01/07).
- Benveniste, E. 1966. *Problèmes de linguistique générale*. 5<sup>e</sup> partie : L'homme dans la langue. Paris : Gallimard.
- Birkelund, M., G. Boysen & P. Kjærsgaard (éds.) 2003. *Aspects de la modalité*. Tübingen : Niemeyer.
- Björk, L. & C. Räisänen. 2003. *Academic writing. A university writing course*. Lund : Studentlitteratur.
- Blanché, R. 1970. *La logique et son histoire d'Aristote à Russell*. Paris : Colin.
- Borillo, A. 2004. Les 'Adverbes d'opinion forte' SELON MOI, A MES YEUX, A MON AVIS,... : point de vue subjectif et effet d'atténuation. *Langue française* 142, 31-41.
- Bourdin, P. 1989. SEMBLER et PARAÎTRE, ou les deux visages de l'apparence. *Semantikos* 10/1-2, 45-68.
- Bouscaren, J. & J. Chuquet. 1987. *Grammaire et textes anglais. Guide pour l'analyse linguistique*. Paris : Ophrys.
- Boye, K. 2002. *Modalitet, evidentialitet og epistemicitet. En teori om de tre sproglige kategorier og en analyse af deres verbale manifestationer i dansk*. Specialeafhandling, Københavns Universitet.
- Boysen, G. 2003. Deux facteurs fondamentaux pour la syntaxe modale des complétives : personne grammaticale et fonction syntaxique. In Birkelund, M., G. Boysen & P. Kjærsgaard (éds.), 45-48.
- Breivega, K.R. 2003. *Vitskaplege argumentasjonsstrategiar. Ein komparativ analyse av superstrukturelle konfigurasjonar i medisinske, historiske og språkvitskaplege artiklar*. Nr. 8 i skriftserien *Sakprosa*, Universitetet i Oslo.
- Brown, P. & S. Levinson 1987. *Politeness: Some universals in language use*. Cambridge : CUP.
- Butler, C.S. 1990. Qualifications in science : Modal meanings in scientific texts. In Nash, W. (ed.) *The Writing Scholar: Studies in Academic Discourse*. Newbury Park : Sage, 137-170.
- Chafe, W. & Nichols, J. (eds.) 1986. *Evidentiality : The linguistic coding of epistemology*. Norwood : Ablex Publishing Corporation.
- Clyne, M. 1991. The socio-cultural dimension: the dilemma of the German-speaking scholar. In Schröder, H (ed.) *Subject-oriented texts*. Berlin : de Gruyter, 49-67.



- Clyne, M. 1998. Cultural differences in the organization of academic texts: English and German. In Cheshire, J. & P. Trudgill (eds.) : *The sociolinguistics reader*, vol. 2. London : Arnold, 315-347.
- Coates, J. 1983. *The semantics of the modal auxiliaries*. London : Croom Helm.
- Connor, U. 2004. Introduction. *Journal of English for Academic Purposes* 3, 271-276.
- Conte, M-E. 1998. Enoncés modaux et reprise anaphorique. In Jonasson, K., M. Forsgren & H. Kronning (éds.) : *Prédication, assertion, information*. Studia Romanica Upsaliensia 56. Uppsala : Acta Universitatis Upsaliensis.
- Cotte, P. 2002. *Le possible, le nécessaire et la volonté*. Disponible sur : <http://www.univ-pau.fr/ANGLAIS/alaes/modaux.html> (Consulté le 21.09.05).
- Crismore, A., R. Markkanen & M. S. Steffensen 1993. Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students. *Written Communication* 10/1, 39-71.
- Crompton, P. 1997. Hedging in academic writing: Some theoretical problems. In *English for specific purposes* 16, 271-287.
- Dahl, T. 2003. Metadiscourse in research articles. In Fløttum, K. & F. Rastier (éds.) : *Academic discourse. Multidisciplinary approaches*. Oslo : Novus, 120-138.
- Dahl, T. 2004. Textual metadiscourse in research articles: A marker of national culture or of academic discipline? *Journal of Pragmatics* 36/10, 1807-1825.
- Darian, S. 1995. Hypotheses in introductory science texts. *IRAL (International review of applied linguistics in language teaching)* 33/2, 83-108.
- David, J. & G. Kleiber (éds.) 1983. *La notion sémantico-logique de modalité*. Paris : Klincksieck.
- Dendale, P. 1993. Le conditionnel de l'information incertaine : marqueur modal ou marqueur évidentiel ? In Hilty, G. et al. *Actes du XX<sup>e</sup> Congrès International de Linguistique et Philologie Romanes*, Tome 1. Tübingen : Francke Verlag, 165-176.
- Dendale, P. 1994. DEVOIR épistémique : marqueur modal ou évidentiel ? *Langue française* 102, p. 24-40.
- Dendale, P. 1999. DEVOIR au conditionnel : valeur évidentio-modale et origine du conditionnel. In Vogeleer, S et al. (éds.), 7-28.
- Dendale, P. & J. Auwera (éds.) 2001. *Les verbes modaux*. Amsterdam : Rodopi.
- Dendale, P. & D. Coltier 2003. Point de vue et évidentialité. *Cahiers de praxématique* 41, 105-130.
- Dendale, P. & L. Tasmowski (éds.) 1994. *Langue française 102 : Les sources du savoir et leurs marques linguistiques*.
- Dendale, P. & L. Tasmowski (eds.) 2001a. *Journal of Pragmatics* 33. Amsterdam : Elsevier Science.
- Dendale, P. & L. Tasmowski 2001b. Introduction : Evidentiality and related notions. In Dendale, P. & L. Tasmowski (éds.) 2001a, 339-348.
- Dendale, P. & L. Tasmowski (éds.) 2001c : *Le conditionnel en français*. Paris : Klincksieck.
- Didriksen, A. 2004. *DONC et autres connecteurs argumentatifs dans les articles de recherche français*. Hovedoppgave, Universitetet i Bergen.
- Douay, C. 2003. Des modalités de l'interlocution au système des modaux. *CORELA* 1/1. Disponible en ligne sur : <http://edel.univ-poitiers.fr/corela/document.php?id=37> (Consulté le 01.03.2007).
- Drubig, H. 2001. *On the syntactic form of epistemic modality*. University of Tübingen. Disponible en ligne sur : <http://www.sfb441.uni-tuebingen.de/b2/papers/DrubigModality.pdf> (Consulté le 08.03.07.)
- Ducrot, O. 1972. *Dire et ne pas dire*. Paris : Hermann.



- Ducrot, O. 1993. A quoi sert le concept de modalité ? In Dittmar, N. & A. Reich (éds.) : *Modalités et acquisition des langues*. Berlin : Walter de Gruyter, 111-129.
- Dyvik, H. 1980. *Grammatikk og empiri. En syntaktisk modell og dens forutsetninger*. Doktoravhandling, Universitetet i Bergen.
- Faarlund, J.T., S. Lie, & K.I. Vannebo. 2002 [1997] : *Norsk referansegrammatikk*. Oslo : Universitetsforlaget.
- Fløttum, K. 2003a. JE et le verbe. *Tribune 14*. Skriftserie for Romansk institutt, Universitetet i Bergen, 7-14.
- Fløttum, K. 2003b. Personal English, indefinite French and plural Norwegian scientific authors? *Norsk Lingvistisk Tidsskrift* 21/1, 21-55.
- Fløttum, K. 2003c. Bibliographical references and polyphony in research articles. In K. Fløttum & F. Rastier (eds.), 97-119.
- Fløttum, K. 2004. La présence de l'auteur dans les articles scientifiques : étude des pronoms JE, NOUS et ON. In Auchlin, A. et al. (éds.) : *Structures et discours*. Québec : Ed. Nota Bene, 401-416.
- Fløttum, K. 2005a. The self and the others – polyphonic visibility in research articles. *International Journal of Applied Linguistics* 15/1, 29-44.
- Fløttum, K. 2005b. MOI et AUTRUI dans le discours scientifique : l'exemple de la négation NE...PAS. In J. Bres, P.P. Haillet, S. Mellet, H. Nølke & L. Rosier (éds.) : *Dialogisme, polyphonie : approches linguistiques*. Bruxelles : De Boeck-Duculot, 323-337.
- Fløttum, K. 2006. En språklig analyse av norsk medisinsk prosa. *Tidsskrift for Den Norske Lægeforening* 126/1, 65-68.
- Fløttum, K. & F. Rastier, 2003. (eds.): *Academic discourse. Multidisciplinary approaches*. Oslo : Novus.
- Fløttum, K, T. Dahl & T. Kinn 2006. *Academic voices - Across languages and disciplines*. Amsterdam : John Benjamins.
- Fløttum, K., T. Dahl, T. Kinn, A.M. Gjesdal & E.T. Vold. 2007. Cultural identities and academic voices. In Fløttum, K. (éd.): *Language and discipline perspectives on academic discourse*. Cambridge : Cambridge Scholars Press, 14-40.
- Fløttum, K., K. Jonasson & C. Norén 2007. *ON – pronom à facettes*. Bruxelles : Duculot-DeBoeck.
- Foullioux, C & D. Tejedor de Felipe. 2004. A propos du mode et de l'atténuation. *Langue française* 142, 112-126.
- Fretheim, T. 1981. "Ego"-dempere og "Alter"-dempere. *Maal og Minne*, 86-100.
- Fretheim, T. 1988. Noen kommentarer til Askedals 'Beschreibung norwegischer Partikelstrukturen'. *Norskraft* 59, 54-81.
- Fuchs, C. 1996. *Les ambiguïtés du français*. Paris : Ophrys.
- Gardies, J-L. 1983. Tentative d'une définition de la modalité. In David, J. & G. Kleiber (éds), 13-24.
- Gilhus, N.E. 2003. Medisinernes publisering; eksakt overføring av ny og nyttig informasjon eller selvhevdende drøvtygging av gamle resultater? *Akademisk Prosa* 1, 15-26.
- Gosselin, L. 2000. Le statut du temps et de l'aspect dans la structure modale de l'énoncé. Esquisse d'un modèle global. *Syntaxe et Sémantique* 2, Presses Universitaires de Caen, 57-80.
- Granger, G. G. 1976. *La théorie aristotélicienne de la science*. Paris : Aubier Montaigne.
- Gresset, S. 2002. MIGHT, COULD et le possible épistémique ? Ou l'interchangeabilité à l'épreuve du texte. Version remaniée d'une communication faite au colloque *Modality in Contemporary English* (Vérone 6-8 septembre 2001). Disponible en ligne sur : [www.univ-pau.fr/ANGLAIS/alaes/modaux02/gresset.pdf](http://www.univ-pau.fr/ANGLAIS/alaes/modaux02/gresset.pdf) (Consulté le 02.03.07).

- Grossmann, F. & F. Rinck. 2004. La surénonciation comme norme du genre : l'exemple de l'article de recherche et du dictionnaire en linguistique. *Langages* 156, 34-50.
- Guentchéva, Z. 1994. Manifestations de la catégorie du médiatif dans les temps du français. *Langue française* 102, 8-23.
- de Haan, F. 2001. The Relation between Modality and Evidentiality. In R. Müller et M. Reis (eds.) : *Modalität und Modalverben im Deutschen*. Linguistische Berichte, Sonderheft 9. Hamburg : Buske, 201-216.  
Disponible en ligne sur [www.u.arizona.edu/~fdehaan/papers/SWJL99.pdf](http://www.u.arizona.edu/~fdehaan/papers/SWJL99.pdf) (Consulté le 08.03.2007).
- Halliday, M.A.K. 1994. *An Introduction to Functional Grammar*. London : Arnold.
- Hardman, M.J. 1986. Data-Source Marking in the Jaqi Languages. In Chafe, W. et J. Nichols (eds.), 113-136.
- Herslund, M. 1989. Modality: A presentation. In Herslund, M. (ed.): *On modality. Papers from meetings and discussions in the linguistic circle of Copenhagen*. København : Akademisk forlag, 7-16.
- Hinds, J. 1987. Reader versus writer responsibility. A new typology. In Connor, U. et R. Kaplan (eds.) : *Writing across languages: analysis of L2 text*. Reading, MA : Addison-Wesley, 141-152.
- Holmes, J. 1986. Functions of YOU KNOW in women's and men's speech. *Language in Society* 15/1, 1-21.
- Holmes, J. 1990. Hedges and boosters in women's and men's speech. *Language and Communication* 10/3, 185-205.
- Hopper, P. & E. Traugott. 1993. *Grammaticalization*. Cambridge : Cambridge University Press.
- Hughes, G.E. & M.J. Cresswell. 1968. *An introduction to modal logic*. London : Methuen & CO.
- Hunston, S. 1994. Modal meanings in the structure of argumentation. In Brekke, M., Ø. Andersen, T. Dahl & J. Myking (eds.): *Applications and implications of current LSP research*. Proceedings of the 9<sup>th</sup> European LSP Symposium. Bergen : Universitetet i Bergen, 619-629.
- Huot, H. 1974. *Le verbe DEVOIR. Etude synchronique et diachronique*, Paris : Klincksieck.
- Huot, H. 1986. Le subjonctif dans les complétives : subjectivité et modalisation. In Ronat, M. & D. Couquaux (éds.) : *La grammaire modulaire*. Paris : Minuit, 81-111.
- Hyland, K. 1996a. Talking to the academy. Forms of hedging in science research articles. *Written Communication* 13/2, 251-281.
- Hyland, K. 1996b. Writing without conviction? Hedging in science research articles. *Applied Linguistics* 17/4, 433-454.
- Hyland, K. 1998. *Hedging in scientific research articles*. Amsterdam : John Benjamins Publishing Company.
- Hyland, Ken 1999a. Talking to students: Metadiscourse in introductory textbooks. *English for Specific Purposes* 18/1, 3-26.
- Hyland, Ken 1999b. Disciplinary discourses: writer stance in research articles. In Candlin, C & K. Hyland (eds.): *Writing: texts, processes and practices*. London : Longman, 99-122.
- Hyland, K. 2000. Hedges, boosters and lexical invisibility: Noticing modifiers in academic texts. *Language Awareness* 9/4, 179-197.
- Hyland, K. 2001. Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes* 20/3, 207-226.
- Hyland, K. 2004 [2000]. *Disciplinary discourses : social interactions in academic writing*. Ann Arbor : University of Michigan Press.
- Hyland, K. & M. Bondi (eds.) 2006: *Academic discourse across disciplines*. Peter Lang.

- Johansson, S. 2001. The English verb SEEM and its correspondences in Norwegian: What seems to be the problem? In Aijmer, K. (ed): *A wealth of English. Studies in honour of Göran Kjellmer*. Gothenburg Studies in English 81. Göteborg : Acta Universitatis Gothoburgensis, 221-245.
- Johansson, S. 2003. Contrastive linguistics and corpora. In: Granger, S., J. Lerot & S. Petch-Tyson (eds.), *Corpus-based approaches to contrastive linguistics and translation studies*. Approaches to translation studies 20. Amsterdam : Rodopi.
- Johns, T. 2001. From evidence to conclusion: The case of INDICATE THAT. In Hewings, M. (ed.) *Academic writing in context. Implications and applications*. Birmingham: University of Birmingham Press, 55-62.
- Kerbrat-Orecchioni, C. 2002 [1980]: *L'énonciation : de la subjectivité dans le langage*. Paris : Armand Colin.
- Kibbee, D. 1995. Assertion/atténuation, subjectivité/objectivité en anglais et en français : SEEM/SEMBLER. In Ballard, M. (éd) : *Relations discursives et traduction*. Lille : Presses universitaires de Lille, 73-88.
- Kolflaath, E. 2003. *Språk og argumentasjonsteori – med eksempler fra juss*. Kompendium, Universitetet i Bergen.
- Kourilovà, M. 1994. Some problems posed by the system of epistemic modality in written scientific discourse. In Brekke, M., Ø. Andersen, T. Dahl & J. Myking (eds.): *Applications and implications of current LSP research*. Proceedings from the 9<sup>th</sup> European symposium on LSP. Bergen : Fagbokforlaget, 639-646.
- Kronning, H. 1996. *Modalité, cognition et polysémie : sémantique du verbe modal DEVOIR*. Uppsala : Acta Universitatis Upsaliensis.
- Kronning, H. 2001a. Pour une tripartition des emplois du modal DEVOIR. In Dendale, P. & J. Auwera (éds.), 67-84.
- Kronning, H. 2001b. Au-delà du déontique et de l'épistémique. In Prandi, M. & P. Ramat (éds) : *Semiotica e linguistica. Per ricordare Maria-Elisabeth Conte. Materiali Linguistici* 32, Università di Pavia. Milano : FrancoAngeli, 97-121.
- Kronning, H. 2002. Le conditionnel « journalistique » : médiation et modalisation épistémiques. *Romansk Forum* 16/2, 561-575.
- Kronning, H. 2003. Modalité et évidentialité. In Birkelund, M., G. Boysen & P. Kjærsgaard (éds.), 131-152.
- Kronning, H. 2005. *Epistemiska uttryck i språket – en fråga om ansvarstagande*. Cours donné à l'université de Bergen 08.11.05.
- Kronning, H., à paraître. Om epistemiska uttryck i de romanska språken. *Kungliga Vetenskaps-Societeten/Royal Society of Sciences*, 1-35.
- Lakoff, G. 1972. *Hedges : A study of meaning criteria and the logic of fuzzy concepts*. Osculd : Göteborg.
- Lakoff, R. 1975. *Language and Woman's Place*. New York : Harper and Row.
- Laurendeau, P. 2004. Modalité, opération de modalisation et mode médiatif. In Delamotte-Legrand, R. (éd.) : *Les médiations langagières*, Vol. 1. Dyalang, CNRS. Rouen : Publications de l'Université de Rouen, 83-95.  
Version électronique, disponible sur <http://www.yorku.ca/paull/articles/2004a.html>. (Consultée le 01.03.2007).
- Le Querler, N. 1996. *Typologie des modalités*. Caen : Presses universitaires de Caen.
- Le Querler, N. 2001. La place du verbe modal POUVOIR dans une typologie des modalités. In Dendale, P. & J. Auwera (éds.), 17-32.
- Lewin, B.A. 2005. Hedging : An exploratory study of authors' and readers' identification of 'toning down' in scientific texts. *Journal of English for Academic Purposes* 4, 163-178.

- Li, Ch.N. 1986. Direct speech and indirect speech: A functional study. In Coulmas, F. (ed.): *Direct and indirect speech*. Berlin : Mouton de Gruyter, 29-46.
- Loffler-Laurian, A.-M. 1980. L'expression du locuteur dans les discours scientifiques. *Revue de linguistique romane* 44, 135-157.
- Lundquist, L. 1990. *L'analyse textuelle : méthode, exercice*. 2<sup>nd</sup> éd. København : Handelshøjskolens forlag.
- Lundquist, L. 1998. Cadres et espaces mentaux dans la structuration cognitive des connaissances – et des discours – juridiques. In Gambier, Y. (éd.): *Discours professionnels en français*. Frankfurt am Main : Peter Lang, 123-150.
- Lyons, J. 1981. *Language, Meaning and Context*. London : Fontana.
- Lyons, J. 1994 [1977] : *Semantics*. Cambridge: Cambridge University Press.
- Løken, B. 1997. Expressing possibility in English and Norwegian. *ICAME Journal* 21, 43-61.
- Løken, B. 2004. English and Norwegian expressions of modal possibility: Outline of a project and preliminary findings. In Aijmer, K. & H. Hasselgård (eds.): *Translation and Corpora*. Selected papers from the Göteborg-Oslo symposium 18-19 October 2003. Göteborg : Acta Universitatis Gothoburgensis, 33-42.
- Maingueneau, D. 1993. *Éléments de linguistique pour le texte littéraire*. Paris : Dunod.
- Markkanen, R. & H. Schröder (eds.) 1997. *Hedging and discourse : approaches to the analysis of a pragmatic phenomenon in academic texts*. Research in text theory, vol. 24. Berlin : Walter de Gruyter.
- Martin, R. 1987. *Langage et croyance : les « univers de croyance » dans la théorie sémantique*. Bruxelles : P. Mardaga.
- Mauranen, A. 1993. *Cultural differences in academic rhetoric: a textlinguistic study*. Frankfurt am Main : Peter Lang.
- Mauranen, A. 1997. Hedging in language revisers' hands. In Markkanen, R. & H. Schröder (eds.), 115-133.
- McCawley, J.D. 1993. *Everything that linguists have always wanted to know about logic (but were ashamed to ask)*. Chicago : The University of Chicago Press.
- McEnery, T. & N.A Kifle 2002. Epistemic modality in argumentative essays of second-language writers. In Flowerdew, J. (ed.) *Academic Discourse*. London : Longman, 182-195.
- Melander, B. 1998. Culture or genre? Issues in the interpretation of cross-cultural differences in scientific papers. In Fortanet, I., S. Posteguillo, J.C Palmer & J.F Coll (eds.): *Genre studies in English for academic purposes*. Col·lecció "Summa", Sèrie Filologia 9. Castello de la Plana: Universitat Jaume I, 211-226.
- Melander, B., J. M. Swales & K. M. Fredrickson 1997. Journal abstracts from three academic fields in the United States and Sweden : national or disciplinary proclivities ? In Duszak, A. (ed.) : *Culture and Styles of Academic Discourse*. Trends in Linguistics 104. Berlin : Mouton de Gruyter, 251-273.
- Meunier, A. 1981. Grammaires du français et modalités. Matériaux pour une nébuleuse. *DRLAV* 25, Université de Paris 8.
- Motchane, J.L. 1990. Chercher, inventer, innover dans la langue. In Cassen, B., (éd.) : *Quelles langues pour la science?* Paris : Éditions de la Découverte, 4-62.
- Myers, G. 1989. The pragmatics of politeness in scientific articles. *Applied Linguistics* 10, 1-35.
- Nuyts, J. 2001. Subjectivity as an evidential dimension in epistemic modal expressions. In Dendale, P. & L.Tasmowski (eds.) 2001a, 383-400.
- Nølke, H. 1989. Modality and polyphony. A study of some French adverbials. In Herslund, M. (ed.): *On modality. Papers from meetings and discussions in the linguistic circle of Copenhagen*. København : Akademisk forlag, 45-64.

- Nølke, H. 1991. Contrastive pragmatic linguistics. In Lauridsen, K. M & O. Lauridsen (eds.) : *Contrastive linguistics. Papers from the CL symposium, 28-30 August 1989*. Århus : Handelshøjskolen i Århus, 199-231.
- Nølke, H. 1993. *Le regard du locuteur*. Paris : Kimé.
- Nølke, H. 2001. La dilution linguistique des responsabilités. Essai de description polyphonique des marqueurs évidentiels IL SEMBLE QUE et IL PARAÎT QUE. In Nølke, H. *Le regard du locuteur 2*. Paris : Kimé.
- Nølke, H., K. Fløttum & C. Norén 2004: *ScaPoLine. La théorie scandinave de la polyphonie linguistique*. Paris : Kimé.
- Palmer, F.R. 1986. *Mood and modality*. Cambridge : Cambridge University Press.
- Palmer, F.R. 2001. *Mood and modality*. 2<sup>nd</sup> ed. Cambridge : Cambridge University Press.
- Papafragou, A. 2000. *Modality: Issues in the semantics-pragmatics interface*. Amsterdam : Elsevier.
- Pedersen, J., E. Spang-Hanssen & C. Vikner 1992. *Fransk Grammatik*. København : Akademisk forlag.
- Perkins, M. 1983. *Modal expressions in English*. London : Frances Pinter.
- Picavez, M. H. 2003. *Modalisation et verbes de connaissance: Une approche linguistique des « Pensées » de Pascal*. Thèse de doctorat, Université de Nantes. Disponible en ligne sur [http://marg.lng3.free.fr/documents/the0022\\_picavez\\_h/the0022.pdf](http://marg.lng3.free.fr/documents/the0022_picavez_h/the0022.pdf) (Consulté le 08.03.06.)
- Pottier, B. 1983. Chronologie des modalités. Dans David, J. & G. Kleiber (éds.), 55-64.
- Pottier, B. 1985 [1974]. *Linguistique générale, théorie et description*. Paris : Klincksieck.
- Poudat, C. 2006. *Étude contrastive de l'article scientifique de revue linguistique dans une perspective d'analyse des genres*. Thèse de doctorat, Université d'Orléans. Publié dans *Texto! Textes et cultures* [revue électronique], vol. XI, n°3-4. Disponible en ligne sur : <http://www.revue-texto.net/Corpus/Publications/Poudat/Etude.html> (Consulté le 20.02.07.)
- Preisler, B. 1986. *Linguistic Sex Roles in Conversation: Social Variation in the Expression of Tentativeness in English*. Berlin : Mouton de Gruyter.
- Prince, E.F., J. Frader & C. Bosk. 1982. On hedging in physician-physician discourse. In Di Pietro, R. J. (ed.): *Linguistics and the Professions*. New Jersey : Ablex, 83-97.
- Quirk, R., S. Greenbaum, G. Leech & J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London : Longman.
- Régent, O. 1994. L'article scientifique : un produit culturel. *ASp (Anglais de Spécialité)* 5-6, 55-61.
- Rescher, N. 1968. *Topics in philosophical language*. Dordrecht : Reidel.
- Riegel, M., J-C. Pellat & R. Rioul 1998 [1994]. *Grammaire méthodique du français*. Paris : PUF.
- Rinck, F. 2006. *L'article de recherche en Sciences du Langage et en Lettres : Figure de l'auteur et approche disciplinaire du genre*. Thèse de doctorat, Université Stendhal, Grenoble 3.
- Salager-Meyer, F. 1994. Hedges and textual communicative function in medical English written discourse. *English for Specific Purposes* 13, 149-170.
- Salager-Meyer, F. 1997. I think that perhaps you should: A study of hedges in written scientific discourse. In Miller, T. (ed.): *Functional Approaches to Written Text: Classroom Applications*. Washington DC : USIA, 105-118.
- Salager-Meyer, F. 1998. Language is not a physical object. *English for Specific Purposes* 17, 295-302.
- Salager-Meyer, F. 2000. Debate-creating vs. accounting references in French medical journals. *Technical Communication Quarterly* 9/3, 291-310.



- Salager-Meyer, F. & G. Defives 1998. From the gentleman's courtesy to the expert's caution: a diachronic analysis of hedges in academic writing (1810-1995). In Fortanet, I., S. Posteguillo, J.C Palmer & J.F Coll (eds.): *Genre Studies in English for Academic Purposes*. Col·lecció "Summa", Sèrie Filologia 9. Castello de la Plana: Universitat Jaume I, 133-171.
- Salager-Meyer, F., M. Angeles Alcaraz Ariza & N. Zambrano 2003. The scimitar, the dagger and the glove: intercultural differences in the rhetoric of criticism in Spanish, French and English medical discourse (1930-1995). *English for Specific Purposes* 22, 223-247.
- Salager-Meyer, F., M. Angeles Alcaraz Ariza & M. Pabón 2007. The prosecutor and the defendant: contrasting critical voices in French- and English-written academic book reviews. In Fløttum, K. (ed.): *Language and discipline perspectives on academic discourse*. Cambridge : Cambridge Scholars Press, 115-137.
- Sionis, C. 1997. Writing scientific discourse beyond words. *Asp (Anglais de Spécialité)*, 15-18, 339-357.
- Skelton, J. 1988. The care and maintenance of hedges. *English Language Teaching Journal* 42/1, 37-43.
- Skelton, J. 1997. The representation of truth in academic medical writing. *Applied Linguistics* 18/2, 121-140.
- Squartini, M. 2004. Disentangling evidentiality and epistemic modality in Romance. *Lingua* 114, 873-895.
- Stubbs, M. 1986. A matter of prolonged fieldwork. Notes towards a modal grammar of English. *Applied Linguistics* 7/1, 1-25.
- Sueur, J.P. 1975. *Etude sémantique et syntaxique des verbes DEVOIR et POUVOIR*. Thèse de doctorat, Université de Paris X, Nanterre.
- Sueur, J.P. 1979. Une analyse sémantique des verbes *DEVOIR et POUVOIR*. *Le Français moderne* 2, 97-120.
- Sueur, J.P. 1983. Les verbes modaux sont-ils ambigus ? In David, J. & G. Kleiber (éds).
- Swales, J.M. 2002 [1990]. *Genre Analysis: English in Academic and Research Settings*. Cambridge : CUP.
- Swales, J.M. 2004. *Research Genres. Exploration and Applications*. Cambridge : CUP.
- Tasmowski, L. & P. Dendale. 1994. POUVOIR<sub>E</sub>, un marqueur d'évidentialité. *Langue française* 102, 41-56.
- Thompson, G. 1996. Voices in the text: Discourse perspectives on language reports. *Applied Linguistics* 17/4, 501-530.
- Toulmin, S. 1958. *The Uses of Argument*. London : Cambridge University Press.
- Tutescu, M. 1997. Du modalisateur épistémique au connecteur discursif. In Kleiber, G. & M. Riegel (éds.) : *Les formes du sens*. Louvain-la-Neuve : Duculot, 387-395.
- Uri, H. 1999. Pragmatiske partikler i et andrespråksperspektiv. *Norsk lingvistisk tidsskrift* 17/1, 139-154.
- Van der Auwera, J. & W. Plungian. 1998. On modality's semantic map. *Linguistic Typology* 2/1, 79-124.
- Varttala, T. 1999. Remarks on the communicative functions of hedging in popular scientific and specialist research articles on medicine. *English for Specific Purposes* 18/2, 177-200.
- Varttala, T. 2001. *Hedging in scientifically oriented discourse. Exploring variation according to discipline and intended audience*. Doctoral dissertation, University of Tampere. Published in Acta Electronica Universitatis Tamperensis 138. <http://acta.uta.fi/pdf/951-44-5195-3.pdf>

- Vassileva, I. 1997. Hedging in English and Bulgarian academic writing. In Duszak, A. *Culture and Styles of Academic Discourse*. Trends in Linguistics 104. Berlin : Mouton de Gruyter, 203-222.
- Vassileva, I. 1998. Who am I/who are we in academic writing? A contrastive analysis of authorial presence in English, German, French, Russian and Bulgarian. *International Journal of Applied Linguistics* 8/2, 163-190.
- Ventola, E. & A. Mauranen (eds.) 1996. *Academic writing: intercultural and textual issues*. Amsterdam : Benjamins.
- Ventola, E. 1997. Modalization : Probability – an exploration into its role in academic writing. In Duszak, A. *Culture and Styles of Academic Discourse*. Trends in Linguistics 104. Berlin : Mouton de Gruyter, 157-180.
- Vet, C. 1997. Modalités grammaticalisées et non-grammaticalisées. In Kleiber, G. & M. Riegel (éds.) : *Les formes du sens*. Louvain-la-Neuve : Duculot, 405-412.
- Vetters, C. 2001. Le conditionnel: ultérieur du non-actuel. In Dendale, P. & L. Tasmowski (éds.) 2001c, 169-207.
- Vihla, M. 2000. Epistemic Possibility: A study based on a medical corpus. In Kirk, J.M (ed.): *Corpora galore : analyses and techniques in describing English : papers from the Nineteenth International Conference on English Language Research on Computerised Corpora (ICAME 1998)*. Language and computers, no 30. Amsterdam : Rodopi, 209-224.
- Vion, R. 2001. Modalités, modalisations et activités langagières. *Marges linguistiques* 2. <http://www.marges-linguistiques.com>
- Vion, R. 2004. Modalités, modalisations et discours représentés. *Langages* 156, 96-111.
- Vogeleer, S., A. Borillo, M. Vuillaume & C. Vetters 1999. *La modalité sous tous ses aspects*. Amsterdam : Rodopi.
- Vold, E. T. (2006a). Epistemic modality markers in research articles: a cross-linguistic and cross-disciplinary study. *International Journal of Applied Linguistics* 16/1, 61-87.
- Vold, E. T. (2006b). The choice and use of epistemic modality markers in linguistics and medical research articles. In Hyland, K. & M. Bondi (eds.) : *Academic discourse across disciplines*. Bern : Peter Lang, 225-250.
- Wright, G.H. von 1951. *An essay in modal logic*. Amsterdam : North-Holland.
- Åfarli, T. A & K. M. Eide 2003. *Norsk generativ syntaks*. Oslo : Novus.





## Appendice 1 : le corpus

### English linguistics (Engling)

#### Revues:

*Journal of Linguistics* 1998-2001

*English for Specific Purposes* 2001-2002

*Linguistic Inquiry* 2000

*Language. Journal of the Linguistic Society of America* 2001

ENGLING01 (Nombre de mots: 6599)

Ellen Thompson (2001): Temporal dependency and the syntax of subjects. *Journal of Linguistics* 37/2, 287-312.

ENGLING02 (Nombre de mots: 7503)

Matthew Whelpton (2001): Elucidation of a telic infinitive. *Journal of Linguistics* 37/2, 313-337.

ENGLING03 (Nombre de mots: 9935)

John A. Hawkins (2001): Why are categories adjacent? *Journal of Linguistics* 37/1, 1-34.

ENGLING04 (Nombre de mots: 6863)

Frederick J. Newmeyer (2001): The Prague School and North American functionalist approaches to syntax. *Journal of Linguistics* 37/1, 101-126.

ENGLING05 (Nombre de mots: 8252)

Diane Blakemore (2000): Indicators and procedures: NEVERTHELESS and BUT. *Journal of Linguistics* 36/3, 463-486.

ENGLING06 (Nombre de mots: 7960)

Peter Trudgill, Elizabeth Gordon, Gillian Lewis & Margaret Maclagan (2000): Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics* 36/2, 299-318.

ENGLING07 (Nombre de mots: 11692)

John Harris, Sally Bates & Jocelyne Watson (1999): Prosody and melody in vowel disorder. *Journal of Linguistics* 35/3, 489-525.

ENGLING08 (Nombre de mots: 15023)

Robyn Carston (1998): Negation, 'presupposition' and the semantics/pragmatics distinction. *Journal of Linguistics* 34/2, 309-350.

ENGLING09 (Nombre de mots: 9131)

Alex Lascarides & Ann Copestake (1998): Pragmatics and word meaning. *Journal of Linguistics* 34/2, 387-414.

ENGLING10 (Nombre de mots: 4725)

Kate Kearns (1998): Extraction from make the claim constructions. *Journal of Linguistics* 34/1, 53-72.

ENGLING11 (Nombre de mots: 5114)

Brian Paltridge (2002): Thesis and dissertation writing: an examination of published advice and actual practice. *English for Specific Purposes* 21/2, 125-143.

ENGLING12 (Nombre de mots: 6395)

Elisabeth Rowley-Jolivet (2002): Visual discourse in scientific conference papers. A genre-based study. *English for Specific Purposes* 21/1, 19-40.

ENGLING13 (Nombre de mots: 5673)

Glenda Crosling & Ian Ward (2002): Oral communication: the workplace needs and uses of business graduate employees. *English for Specific Purposes* 21/1, 41-57.

ENGLING14 (Nombre de mots: 4179)

Thomas A. Upton & Ulla Connor (2001): Using computerized corpus analysis to investigate the textlinguistic discourse moves of a genre. *English for Specific Purposes* 20/4, 313-329.

ENGLING15 (Nombre de mots: 5187)

Ken Hyland (2001): Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes* 20/3, 207-226.

ENGLING16 (Nombre de mots: 11893)

Mark Steedman (2000): Information Structure and the Syntax-Phonology Interface. *Linguistic Inquiry* 31/4, 649-685.

ENGLING17 (Nombre de mots: 16997)

Andrea Tyler & Vyvyan Evans (2001): Reconsidering Prepositional Polysemy Networks: The Case of OVER. *Language. Journal of the Linguistic Society of America* 77/4, 724-765.

ENGLING18 (Nombre de mots: 11414)

Malka Rappaport Hovav & Beth Levin (2001): An Event Structure Account of English Resultatives. *Language. Journal of the Linguistic Society of America* 77/4, 766-797.

ENGLING19 (Nombre de mots: 5028)

Helen Goodluck (2001): The Nominal Analysis of Childrens's Interpretations of Adjunct Pro Clauses. *Language. Journal of the Linguistic Society of America* 77/3, 494-509.

ENGLING20 (Nombre de mots: 11418)

Rebecca Clift (2001): Meaning in Interaction. The Case of ACTUALLY. *Language. Journal of the Linguistic Society of America* 77/2, 245-291.

## **English medicine (Engmed)**

### **Revues:**

*British Medical Journal* 2000-2001.

*International Journal of Cancer* 1999, 2001.

*Journal of the American Medical Association (Jama)* 2002.

ENGMED01 (Nombre de mots: 2241)

Gordon C.S. Smith & Jill P Pell (2001): Teenage pregnancy and risk of adverse perinatal outcomes associated with first and second births: population based retrospective cohort study. *British Medical Journal* 323/7311, 476-479.

ENGMED02 (Nombre de mots: 2568)

Helen Bedford, John de Louvois, Susan Halket, Catherine Peckham, Rosaline Hurley & David Harvey (2001): Meningitis in infancy in England and Wales: follow up at age 5 years. *British Medical Journal* 323/7312, 533-535.

ENGMED03 (Nombre de mots: 1989)

Georgina Stamp, Gillian Kruzins & Caroline Crowther (2001): Perineal massage in labour and prevention of perineal trauma: randomised controlled trial. *British Medical Journal* 322/7297, 1277-1280.

ENGMED04 (Nombre de mots: 2266)

Susan F. Murray (2000): Relation between private health insurance and high rates of caesarean section in Chile: qualitative and quantitative study. *British Medical Journal* 321/7275, 1501-1505.

ENGMED05 (Nombre de mots: 2135)

S. Rachel Thomas, D. R. S. Jamieson & Keith W. Muir (2000): Randomised controlled trial of atraumatic versus standard needles for diagnostic lumbar puncture. *British Medical Journal* 321/7267, 986-990.

ENGMED06 (Nombre de mots: 1845)

Gregory S. Merrick, Kent Wallner, Wayne M. Butler, Jonathan H. Lief & Steven Sutlief (2001): Short-term sexual function after prostate brachytherapy. *International Journal of Cancer* 96/5, 313-319.

ENGMED07 (Nombre de mots: 2706)

Tracey A. Holland, Jackie Elder, Jonathan M. McCloud, Christine Hall, Mark Deakin, Anthony A. Fryer, James B. Elder & Paul R. Hoban (2001): Subcellular localisation of cyclin D1 protein in colorectal tumours is associated with p21WAF1/CIP1 expression and correlates with patient survival. *International Journal of Cancer* 95/5, 302-306.

ENGMED08 (Nombre de mots: 3203)

Miriam V. Dwek, Heidi A. Ross, Andrew J. Streets, Susan A. Brooks, Elizabeth Adam, Ann Titcomb, Jayne V. Woodside, Udo Schumacher & Anthony J. Leathem (2001): Helix pomatia agglutinin lectin-binding oligosaccharides of aggressive breast cancer. *International Journal of Cancer* 95/2, 79-85.

ENGMED09 (Nombre de mots: 2596)

William J. Merchant, Rosemary R. Millis, Paul Smith, Murid A. Chaudary & Diana M. Barnes (1999): Expression of c-erbB2 and p53 protein is similar in breast cancer from British and Japanese women. *International Journal of Cancer* 84/3, 278-283.

ENGMED10 (Nombre de mots: 2846)

Daniel W. Cramer, Hannah Kuper, Bernard L. Harlow & Linda Titus-Ernstoff (2001): Carotenoids, antioxidants and ovarian cancer risk in pre- and postmenopausal women. *International Journal of Cancer* 94/1, 128-134.

ENGMED11 (Nombre de mots: 3251)

Mark A. Hlatky, Derek Boothroyd, Eric Vittinghoff, Penny Sharp & Mary A. Whooley (2002): Quality-of-Life and Depressive Symptoms in Postmenopausal Women After Receiving Hormone Therapy. Results From the Heart and Estrogen/Progestin Replacement Study (HERS) Trial. *Journal of the American Medical Association (Jama)* 287/5, 591-597.

ENGMED12 (Nombre de mots: 3104)

Jane F. Seward, Barbara M. Watson, Carol L. Peterson, Laurene Mascola, Jan W. Pelosi, John X. Zhang, Teresa J. Maupin, Gary S. Goldman, Laura J. Tabony, Kimberly G. Brodovicz, Aisha O. Jumaan & Melinda Wharton (2002): Varicella Disease After Introduction of Varicella Vaccine in the United States, 1995-2000. *Journal of the American Medical Association (Jama)* 287/5, 606-611.

ENGMED13 (Nombre de mots: 3839)

Diane Feskanich, Vishwa Singh, Walter C. Willett & Graham A. Colditz (2002): Vitamin A Intake and Hip Fractures Among Postmenopausal Women. *Journal of the American Medical Association (Jama)* 287/1, 47-54.

ENGMED14 (Nombre de mots: 4422)

Arthur Margolin, Herbert D. Kleber, S. Kelly Avants, Janet Konefal, Frank Gawin, Elena Stark, James Sorensen, Eleanor Midkiff, Elizabeth Wells, T. Ron Jackson, Milton Bullock, Patricia D. Culliton, Sharon Boles & Roger Vaughan (2002): Acupuncture for the Treatment of Cocaine Addiction. A Randomized Controlled Trial. *Journal of the American Medical Association (Jama)* 287/1, 55-63.

ENGMED15 (Nombre de mots: 4238)

Gregory P. Geba, Arthur L. Weaver, Adam B. Polis, Mary E. Dixon & Thomas J. Schnitzer (2002): Efficacy of Rofecoxib, Celecoxib, and Acetaminophen in Osteoarthritis of the Knee. A Randomized Trial. *Journal of the American Medical Association (Jama)* 287/1, 64-71.

ENGMED16 (Nombre de mots: 2441)

Steven R. Cummings, Tu Duong, MA; Emily Kenyon, Jane A. Cauley, Malcolm Whitehead & Kathryn A. Krueger (2002): Serum Estradiol Level and Risk of Breast Cancer During Treatment With Raloxifene. *Journal of the American Medical Association (Jama)* 287/2, 216-220.

ENGMED17 (Nombre de mots: 2408)

Dennis H. Osmond, Susan Buchbinder, Amber Cheng, Alison Graves, Eric Vittinghoff, Cynthia K. Cossen, Bagher Forghani & Jeffrey N. Martin (2002): Prevalence of Kaposi Sarcoma-Associated Herpesvirus Infection in Homosexual Men at Beginning of and During the HIV Epidemic. *Journal of the American Medical Association (Jama)* 287/2, 221-225.

ENGMED18 (Nombre de mots: 3863)

Jane C. Weeks, Heidi Nelson, Shari Gelber, Daniel Sargent & Georgene Schroeder (2002): Short-term Quality-of-Life Outcomes Following Laparoscopic-Assisted Colectomy vs Open Colectomy for Colon Cancer. A Randomized Trial. *Journal of the American Medical Association (Jama)* 287/3, 321-328.

ENGMED19 (Nombre de mots: 3505)

Robert C. Green, L. Adrienne Cupples, Rodney Go, Kelly S. Benke, Timi Edeki, Patrick A. Griffith, Mary Williams, Yvonne Hipps, Neill Graff-Radford, David Bachman & Lindsay A. Farrer (2002): Risk of Dementia Among White and African American Relatives of Patients With Alzheimer Disease. *Journal of the American Medical Association (Jama)* 287/3, 329-336.

ENGMED 20 (Nombre de mots: 3944)

David W. Kaufman, Judith P. Kelly, Lynn Rosenberg, Theresa E. Anderson & Allen A. Mitchell (2002): Recent Patterns of Medication Use in the Ambulatory Adult Population of the United States. The Slone Survey. *Journal of the American Medical Association (Jama)* 287/3, 337-344.

## French linguistics (Frling)

### Reuves:

*Travaux de Linguistique*, Bruxelles, Belgique: Duculot. 2001-2002.

*Marges Linguistiques* (Revue en ligne: <http://www.marges-linguistiques.com/>). 2001.

FRLING01 (Nombre de mots : 2267)

Claire Blanche-Benveniste (2001): Préposition à éclipses. *Travaux de Linguistique* 42-43, 83-95.

FRLING02 (Nombre de mots : 3327)

Claude Cortier (2001): Les syntagmes prépositionnels prédicatifs dans les grammaires universitaires : un observatoire de la place accordée aux prépositions. *Travaux de Linguistique* 42-43, 121-140.

FRLING03 (Nombre de mots : 1675)

André Dugas (2001): Une analyse des constructions transitives indirectes en français. *Travaux de Linguistique* 42-43, 111-120.

FRLING04 (Nombre de mots : 1957)

David Gaatone (2001): Les prépositions: une classe aux contours flous. *Travaux de Linguistique* 42-43, 23-32.

FRLING05 (Nombre de mots : 2130)

Lucien Kupferman (2001) : Les déplacement des syntagmes en DE: un regard de troisième type. *Travaux de Linguistique* 42-43, 33-42.

FRLING06 (Nombre de mots : 2963)

Véronique Lagae (2001): Le pronom EN : les compléments adnominaux aux syntagmes quantificateurs. *Travaux de Linguistique* 42-43, 43-58.

- FRLING07 (Nombre de mots : 2514)  
Florence Lefeuve (2001): POUR QUOI ? *Travaux de Linguistique* 42-43, 199-210.
- FRLING08 (Nombre de mots : 3074)  
Ludo Melis (2001): La préposition est-elle toujours la tête d'un groupe prépositionnel? *Travaux de Linguistique* 42-43, 11-22.
- FRLING09 (Nombre de mots : 1610)  
Mireille Piot (2001): Relations entre prépositions et conjonctions ? L'apport de la comparaison en langues romanes. *Travaux de Linguistique* 42-43, 71-82.
- FRLING10 (Nombre de mots : 3329)  
Sylvie Porhiel (2001): AU SUJET DE et A PROPOS DE : une analyse lexicographique, discursive et linguistique. *Travaux de Linguistique* 42-43, 171-182.
- FRLING11 (Nombre de mots : 6965)  
Jacques Moeschler (2001): Pragmatique. Etat de l'art et perspectives. *Marges Linguistiques* 1 (<http://www.marges-linguistiques.com/>).
- FRLING12 (Nombre de mots : 7272)  
Frédéric François (2001): Langage et hors-langage. Quelques remarques. *Marges Linguistiques* 1 (<http://www.marges-linguistiques.com/>).
- FRLING13 (Nombre de mots : 4951)  
Jean-Claude Anscombe (2001): L'analyse de la construction EN TOUT N par D. Leeman : Quelques remarques. *Travaux de Linguistique* 42-43, 183-198.
- FRLING14 (Nombre de mots : 3364)  
Andrée Borillo (2001): Il y a prépositions et prepositions. *Travaux de Linguistique* 42-43, 141-156.
- FRLING15 (Nombre de mots : 4150)  
Pierre Cadiot (in press, 2002): Schémas et Motifs en sémantique prépositionnelle: vers une description renouvelée des prépositions dites 'spatiales'. *Travaux de Linguistique*.
- FRLING16 (Nombre de mots : 2560)  
Patrick Dendale (2001): L'emploi spatial de CONTRE: propositions pour un traitement unifié. *Travaux de Linguistique* 42-43, 229-240.
- FRLING17 (Nombre de mots : 3720)  
Marc Dominicy & Fabienne Martin (2001): À TRAVERS, AU TRAVERS (DE) et le point de vue. *Travaux de Linguistique* 42-43, 211-228.
- FRLING18 (Nombre de mots : 2927)  
Danielle Leeman (2001): TOUT CONTRE vs. TRES CONTRE. *Travaux de Linguistique* 42-43, 241-252.
- FRLING19 (Nombre de mots : 5596)  
Denis Paillard (in press, 2002): Prépositions et rection verbale. *Travaux de Linguistique*.
- FRLING20 (Nombre de mots : 2376)  
Laurence Rosier (in press, 2002): Genre: le nuancier de sa grammaticalisation. *Travaux de Linguistique*.

## French medicine (Frmed)

### Revue:

*Maladies chroniques au Canada* (also in English version), Ottawa, Canada. 1995-1999.

*Annales de Médecine Interne*, Paris, France: Masson. 2000-2001.

FRMED01 (Nombre de mots : 2497)

Robert Choinière, Michel Pageau & Marc Ferland (1999): Prévalence et disparités géographiques de certaines anomalies congénitales au Québec: Comparaison des méthodes d'estimation. *Maladies chroniques au Canada* 20/2.

FRMED02 (Nombre de mots : 3468)

Claudine Laurier, Wendy Kennedy, Jean-Luc Malo, Michèle Paré, Daniel Labbé, André Archambault & André-Pierre Contandriopoulos (1999): Taux et coût des hospitalisations pour l'asthme au Québec: Analyse des données de 1988-1989, 1989-1990 et 1994-1995. *Maladies chroniques au Canada* 20/2.

FRMED03 (Nombre de mots : 3225)

Claudine Laurier, Wendy Kennedy, Line Gariépy, André Archambault & André-Pierre Contandriopoulos (1997): Utilisation de médicaments contre l'asthme dans deux populations québécoises d'utilisateurs d'antiasthmatiques analyse de la base de données des ordonnances. *Maladies chroniques au Canada* 18/1.

FRMED04 (Nombre de mots : 2682)

Thomas Stephens (1995): Rapport d'atelier: Tendances dans la prévalence du tabagisme de 1991 à 1994. *Maladies chroniques au Canada* 16/1.

FRMED05 (Nombre de mots : 2691)

Judy A. Snider & Janet Beauvais (1998): Utilisation du frottis de Papanicolaou: Estimation basée sur une population admissible rajustée en fonction des antécédent d'hystérectomie. *Maladies chroniques au Canada* 19/1.

FRMED06 (Nombre de mots : 1657)

Patrice Cacoub, Abdallah Sbaï, Shuy Vong Toan, Jérôme Bellanger, Catherine Hoang, Pierre Godeau & Jean-Charles Piette (2001): La colite collagène. A propos d'une série de 11 cas. *Annales de Médecine Interne* 152/5, 299-303.

FRMED07 (Nombre de mots : 4267)

Luc Mouthon, Pascal Cohen, Claire Larroche, Marie Hélène André, Isabelle Royer, Philippe Casassus & Loïc Guillemin (1999): Déficit immunitaire commun variable: une ou plusieurs maladies? Trois observations cliniques. *Annales de Médecine Interne* 150/4, 275-282.

FRMED08 (Nombre de mots : 1756)

Emmanuel Andrès, Stéphane Vinzio, Frédéric Maloisel, Sophie Carre, Anne-Elisabeth Perrin, Bernard Goichot & Jean-Louis Schlienger (2001): Neuropathies périphériques auto-immunes à anticorps anti-MAG et hémopathies. À propos de 5 observations. *Annales de Médecine Interne* 152/3, 147-151.

FRMED09 (Nombre de mots : 1160)

Olivier Danon, Ali Mofredj, François Morsli, Marc Duval-Arnould, Danièle Pariente, Jérôme Waguet, Monique Fabre, Christophe Chardot & François Nocton (2000): Tumeur papillaire solide du pancréas chez un enfant. *Annales de Médecine Interne* 151/7, 606-608.

FRMED10 (Nombre de mots : 1309)

Vincent Le Moing, Zeina Eid, Jean-Luc Ecobichon, Xavier Duval, Pascale Longuet, Carole Elbim, Françoise Brun-Vezinet, Catherine Leport & Jean-Louis Vilde (2001): Réponse immunologique à long terme chez des malades infectés par le VIH avec un taux de CD4 < 50/mm<sup>3</sup> au début du traitement par inhibiteurs de protéase. *Annales de Médecine Interne* 152/1, 11-14.

FRMED11 (Nombre de mots : 7249)

Stéphanie Girard, Eric Bruckert & Gérard Turpin (2001): Généralités et atteintes endocriniennes de l'adrénoleucodystrophie. *Annales de Médecine Interne* 152/1, 15-26.

FRMED12 (Nombre de mots : 2070)

Jean-Pierre Cordray, Pierre Nys, Robert-Édouard Merceron & Aurélie Augusti (2001): Fréquence de l'hypothyroïdie après thyroïdite de De Quervain et intérêt de la mesure échographique du volume thyroïdien. *Annales de Médecine Interne* 152/2, 84-88.

FRMED13 (Nombre de mots : 5341)

Maurice Corcos, Nicolas Girardon, Sylvie Nezelof, Paul Bizouard, Jean-Luc Venisse, Gwénéolé Loas, François Lang, Olivier Halfon, Martine Flament & Philippe Jeammet (2000): Pertinence du concept d'addiction dans les troubles des conduites alimentaires. *Annales de Médecine Interne* 151, suppl. B, 53-60.

FRMED14 (Nombre de mots : 3419)

Jean-Paul Battesti & Elie Azoulay (2001): Formes atypiques de sarcoïdose. *Annales de Médecine Interne*.152/1, 51-57.

FRMED15 (Nombre de mots : 2994)

Badreddine Kilani, Lamia Ammari, Hanène Tiouiri, Fakher Kanoun, Ahmed Goubontini, Fayçal Zouiten & Taoufik Ben Chaabane (2001): Les spondylodiscites bactériennes: à pyogènes ou tuberculeuses? *Annales de Médecine Interne* 152/4, 236-241.

FRMED16 (Nombre de mots : 1766)

Imed Riahi, Mourad Mokni, Slim Haouet, Faika Cherif, Dalenda El Euch, Mohamed Iadh Azaiz, Mohamed Moncef Zitouna & Amel Ben Osman-Dhahri (2001): Pyoderma gangrenosum. À propos de 15 observations. *Annales de Médecine Interne* 152/1, 3-9.

FRMED17 (Nombre de mots : 3662)

Aline Tohmé, Dolla Karam-Sarkis, Rabab El-Rassi, Dania Chélala & Elie Ghayad (2001): Agents et conséquences des infections nosocomiales dans un centre hospitalier universitaire libanais. Étude rétrospective sur 2 ans. *Annales de Médecine Interne*. 152/2, 77-83.

FRMED18 (Nombre de mots : 3918)

Catherine Chapelon-Abric (2001): Sarcoïdoses cardiaques. *Annales de Médecine Interne* 152/2, 125-133.

FRMED19 (Nombre de mots : 3592)

Bénédicte Lepère, Laurent Gourarier, Mario Sanchez, Christian Adda, Emmanuelle Peyret, Frédéric Nordmann, Paul Ben Soussen, Mathilde Gisselbrecht & William Lowenstein (2001): Diminution du nombre de surdoses mortelles à l'héroïne, en France, depuis 1994. À propos du rôle des traitements de substitution. *Annales de Médecine Interne*. 152, suppl. au no 3, 5-12.

FRMED20 (Nombre de mots : 2457)

Eric Acquaviva, Laure Beaujouan, Thierry Sainte-Marie & François Chieze (2001): Orientation des patients toxicomanes " en crise " vus aux Urgences (janvier 1999, Assistance Publique - Hôpitaux de Paris). *Annales de Médecine Interne*. 152, suppl. au no 3, 13-17.

## **Norwegian linguistics (Noling)**

### **Revue:**

*Norsk lingvistisk tidsskrift* 1996-1999

NOLING01 (Nombre de mots : 4441)

Janne Bondi Johannessen (1996): DENNE. *Norsk lingvistisk tidsskrift* 14/1, 3-27.

- NOLING02 (Nombre de mots : 3798)  
Dagfinn Worren (1996): Materialgrunnlaget for morsmålsordbøker. *Norsk lingvistisk tidsskrift* 14/1, 47-62.
- NOLING03 (Nombre de mots : 1250)  
Rolf Theil Endresen (1996): Kognitiv lingvistikk og kognitiv grammatikk. *Norsk lingvistisk tidsskrift* 14/2, 97-103.
- NOLING04 (Nombre de mots : 8105)  
Rolf Theil Endresen (1996): Kognitiv morfologi i eit faghistorisk perspektiv. *Norsk lingvistisk tidsskrift* 14/2, 105-141.
- NOLING05 (Nombre de mots : 8431)  
Hans Olav Enger (1996): Ledeformer og telja-klassen. *Norsk lingvistisk tidsskrift* 14/2, 143-169.
- NOLING06 (Nombre de mots : 4197)  
Nesset, Tore (1996): Kasus i russisk, trekk og kognitiv grammatikk. *Norsk lingvistisk tidsskrift* 14/2, 171-188.
- NOLING07 (Nombre de mots : 3654)  
Kristin Meyer Bjerkan & Hanne Gram Simonsen (1996): Prosessering av preteritumsformer i norsk: Eksperimentell evidens fra barn og voksne. *Norsk lingvistisk tidsskrift* 14/2, 189-207.
- NOLING08 (Nombre de mots : 8006)  
Ottar Grønvik (1997): Tanker omkring et etterlatt manuskript av Ingerid Dal: negasjonssystemet i eldste norrønt (eddadikt og skaldekvad). *Norsk lingvistisk tidsskrift* 15/1, 3-32.
- NOLING09 (Nombre de mots : 3445)  
Øystein Alexander Vangsnes (1997): Intensiverende nominalfraser i norsk. *Norsk lingvistisk tidsskrift* 15/1, 35-51.
- NOLING10 (Nombre de mots : 7862)  
Thorstein Fretheim (1997): 'Enhedstryk' (Unit Accentuation): refleksjoner med utgangspunkt i to nye bøker om dansk prosodi. *Norsk lingvistisk tidsskrift* 15/1, 53-82.
- NOLING11 (Nombre de mots : 4444)  
Helene Uri (1997): Setningsforståelse og vanskelighetsfaktorer. *Norsk lingvistisk tidsskrift* 15/2, 139-164.
- NOLING12 (Nombre de mots : 2455)  
Jan Svennevig (1999): Innledning: Samtaleforskning og språkvitenskap. *Norsk lingvistisk tidsskrift* 17/1, 3-13.
- NOLING13 (Nombre de mots : 7262)  
Ragnar Rommetveit (1999): Om dialogisme og vitskapleg disiplinert diskurs- og samtaleanalyse. Dialog som kommunikasjons- og samlivsform. *Norsk lingvistisk tidsskrift* 17/1, 15-40.
- NOLING14 (Nombre de mots : 4051)  
Jan Svennevig (1999): Direkte og indirekte selvrepresentasjon i samtaler mellom ukjente. *Norsk lingvistisk tidsskrift* 17/1, 41-58.
- NOLING15 (Nombre de mots : 3785)  
Karianne Skovholt (1999): Kooperative fullføringer i samtale. *Norsk lingvistisk tidsskrift* 17/1, 59-78.
- NOLING16 (Nombre de mots : 3975)  
Hildegunn Otnes (1999): Oppmerksomhet og respons. Relasjonell lytting i klasseromssamtaler. *Norsk lingvistisk tidsskrift* 17/1, 79-96.
- NOLING17 (Nombre de mots : 2227)  
Rolf Wynn (1999): Røykesnakk. *Norsk lingvistisk tidsskrift* 17/1, 97-108.



NOLING18 (Nombre de mots : 2891)

Marianne Lind (1999): Deltakelsesfremmende strategier i samtale med en afasirammet taler. *Norsk lingvistisk tidsskrift* 17/1, 109-124.

NOLING19 (Nombre de mots : 3067)

Marit Gausdal Skarbø (1999): "Ja" og "nei" som emneinnledere i samtale. *Norsk lingvistisk tidsskrift* 17/1, 125-138.

NOLING20 (Nombre de mots : 3233)

Helene Uri (1999): Pragmatiske partikler i et andrespråksperspektiv. *Norsk lingvistisk tidsskrift* 17/1, 139-154.

## Norwegian medicine (Nomed)

### Revue:

*Tidsskrift for Den Norske Lægeforening* 2000-2001.

NOMED01 (Nombre de mots : 1761)

Nils-Einar Kløw, Bjørn Bendz, Jan Eritsland, Pavel Hoffman, Sindre Stavnes, Knut Haakon Stensæth, Edmund Søvik, Magne Brekke & Arild Mangschau (2001): Angiografiske resultater av primær angioplastikk ved akutt hjerteinfarkt. *Tidsskrift for Den Norske Lægeforening* 121/7, 780-783.

NOMED02 (Nombre de mots : 2442)

Hans H. Wasmuth, Roald J. Guleng, Peter Krane Ysteng, Erling Saltrøe & Trond Veldre Bogsrud (2001): Preoperativ parathyreoideascintigrafi ved primær hyperparatyroidisme. *Tidsskrift for Den Norske Lægeforening* 121/19, 2274-2277.

NOMED03 (Nombre de mots : 2676)

Bjørn Bratland (2001): Polymyalgia rheumatica i en allmennpraksis. *Tidsskrift for Den Norske Lægeforening* 121/19, 2270-2273.

NOMED04 (Nombre de mots : 1850)

Odd Gunnar Heitun (2001): Risperidon som hjelpemiddel ved vanskelig kontrollerbar aggresjon. *Tidsskrift for Den Norske Lægeforening* 121/19, 2265-2267.

NOMED05 (Nombre de mots : 2085)

Odd Mjåland, Thomas Clausen & Theis Tønnessen (2000): Lyskebrokkkirurgi på gammelt vis. *Tidsskrift for Den Norske Lægeforening* 120/1, 18-21.

NOMED06 (Nombre de mots : 2204)

Ola Didrik Saugstad, Terje Rootwelt & Odd O. Aalen (2000): Gjenoppliving av nyfødte med romluft eller oksygen. *Tidsskrift for Den Norske Lægeforening* 120/1, 25-28.

NOMED07 (Nombre de mots : 1933)

Ketil Størdal, Beint Bentsen, Helge Skulstad & Bjørn Moum (2000): Reflukssykdom og 24-timers pH-måling i spiserøret hos barn. *Tidsskrift for Den Norske Lægeforening* 120/2, 183-186.

NOMED08 (Nombre de mots : 2366)

Anne Line Bretteville-Jensen (2000): Kjønnsforskjeller i bruksmønster og helse blant intravenøse heroinmisbrukere i Oslo. *Tidsskrift for Den Norske Lægeforening* 120/2, 192-194.

NOMED09 (Nombre de mots : 2083)

Knut Rygve Jenssen, Lene Dæhlen Selnes, Per Erik Hæreid, Bente Brannsether Ellingsen & Trond Flægstad (2000): Invasive pneumokokkinfeksjoner hos barn fra Troms, Nordland og Sør-Trøndelag 1980-95. *Tidsskrift for Den Norske Lægeforening* 120/3, 308-311.

NOMED10 (Nombre de mots : 3568)

Finn Egil Skjeldestad (2000): P-pillesal, fødsler og svangerskapsavbrot før og etter "Marvelon-saken". *Tidsskrift for Den Norske Lægeforening* 120/3, 339-344.

NOMED11 (Nombre de mots : 2086)

Ludger Dorlöchter, Espen Kolsrud, Edda Olafsdottir, Gjermund Fluge, Karen Rosendahl & Jan Hatlebakk (2000): Oesophagusstriktur som komplikasjon til gastroøsofageal refluks hos barn. *Tidsskrift for Den Norske Lægeforening* 120/2, 187-189.

NOMED12 (Nombre de mots : 2018)

Jan-Erik Berdal & Nils Smith-Erichsen (2000): Fulminant influensapneumoni. *Tidsskrift for Den Norske Lægeforening* 120/3, 312-314.

NOMED13 (Nombre de mots : 1255)

Lasse Gunnar Gøransson, Terje Apeland & Roald Omdal (2000): Hypokalemisk induserte pareser sekundært til renal tubulær acidose. *Tidsskrift for Den Norske Lægeforening* 120/3, 324-325.

NOMED14 (Nombre de mots : 1373)

Tom Mala, Ole Christian Lunde, Arild Nesbakken, Erling Aadland & Morten Stiris (2000): Endoskopisk retrograd kolangiopankreatikografi - et fireårsmateriale. *Tidsskrift for Den Norske Lægeforening* 120/5, 560-562.

NOMED15 (Nombre de mots : 1251)

Svein Tore Baksaas, Odd R. Geiran & Jan L. Svennevig (2000): Svulster i hjertet. *Tidsskrift for Den Norske Lægeforening* 120/20, 2391-2393.

NOMED16 (Nombre de mots : 2164)

Odd Mjåland, Arnhild Flikke & Eldar Normann (2000): Primær hyperparatyreoidisme - et 16-årsmateriale fra et sentralsykehus. *Tidsskrift for Den Norske Lægeforening* 120/20, 2386-2389.

NOMED17 (Nombre de mots : 1055)

Karl-Otto Larsen, Njaal Stray, Vibeke Engh & Dagny Sandnes (2000): Oesophagusskader relatert til bisfosfonater. *Tidsskrift for Den Norske Lægeforening* 120/20, 2397-2399.

NOMED18 (Nombre de mots : 4556)

Claes G. Tropé, Mark Baekelandt, Tone Bjørge, Vera Abeler & Janne Kærn (2000): Borderlinetumorer i ovariene. *Tidsskrift for Den Norske Lægeforening* 120/23, 2764-2770.

NOMED19 (Nombre de mots : 2808)

A. Kathrine Lie (2000): Humant papillomavirus som årsak til kreftutvikling. *Tidsskrift for Den Norske Lægeforening* 120/23, 2771-2776.

NOMED20 (Nombre de mots : 1700)

Theis Tønnessen & Erik Carlsen (2001): Ulcus perforatum. *Tidsskrift for Den Norske Lægeforening* 121/7, 790-792.

## Appendice 2: la distribution des marqueurs entre articles

**Tableau 1 : Le corpus Frling** (frling 01-10 et 13-20 relèvent de *Travaux de Linguistique*, frling 11-12 de *Marges Linguistiques*)

marqueur	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	total
indiquer	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	2
paraître	2	0	0	5	0	2	0	0	0	1	0	0	1	0	6	0	3	0	1	2	23
peut	0	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	4
Peut-être	0	0	0	1	0	1	1	0	0	0	2	3	0	2	1	0	0	1	0	0	12
possible	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	2
pourrait	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	1	1	0	0	0	4
probable	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
probablement	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Sans doute	1	1	0	2	0	0	2	0	0	1	0	2	0	3	5	0	0	0	0	1	18
sembler	7	1	1	3	0	3	3	5	3	2	4	13	3	0	3	3	3	0	2	1	60
suggérer	0	0	0	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	4
<b>total</b>	<b>10</b>	<b>2</b>	<b>1</b>	<b>13</b>	<b>0</b>	<b>7</b>	<b>6</b>	<b>6</b>	<b>3</b>	<b>5</b>	<b>8</b>	<b>21</b>	<b>7</b>	<b>5</b>	<b>17</b>	<b>4</b>	<b>7</b>	<b>1</b>	<b>3</b>	<b>4</b>	<b>130</b>
<b>Fr.rél.</b>	<b>4,4</b>	<b>0,6</b>	<b>0,6</b>	<b>6,6</b>	<b>0</b>	<b>2,4</b>	<b>2,4</b>	<b>2,0</b>	<b>1,9</b>	<b>1,5</b>	<b>1,2</b>	<b>2,9</b>	<b>1,4</b>	<b>1,5</b>	<b>4,1</b>	<b>1,6</b>	<b>1,9</b>	<b>0,3</b>	<b>0,5</b>	<b>1,7</b>	<b>1,9</b>

**Tableau 2 : Le corpus Frmed** (frmed 01-05 relèvent de *Maladies chroniques au Canada*, frmed 06-20 de *Annales de Médecine Interne*.)

<b>marqueur</b>	<b>01</b>	<b>02</b>	<b>03</b>	<b>04</b>	<b>05</b>	<b>06</b>	<b>07</b>	<b>08</b>	<b>09</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>	<b>total</b>
indiquer	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	<b>5</b>
paraître	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	6	0	<b>8</b>
peut	0	2	0	0	4	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	<b>8</b>
Peut-être	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	<b>3</b>
possible	1	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	<b>7</b>
pourrait	0	4	4	0	1	0	6	0	0	1	3	0	0	0	1	0	0	0	1	0	<b>21</b>
probable	0	0	0	0	0	0	3	0	2	0	1	0	0	0	1	0	0	1	0	0	<b>8</b>
probablement	0	2	0	1	0	0	2	1	0	0	6	0	3	0	0	1	0	1	0	0	<b>17</b>
Sans doute	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	<b>2</b>
sembler	0	5	1	2	0	2	0	1	2	1	7	1	8	0	5	4	2	5	0	0	<b>46</b>
suggérer	0	0	0	0	0	3	0	0	0	3	10	0	0	1	0	0	0	0	0	0	<b>17</b>
<b>total</b>	<b>1</b>	<b>18</b>	<b>6</b>	<b>4</b>	<b>7</b>	<b>5</b>	<b>12</b>	<b>2</b>	<b>4</b>	<b>5</b>	<b>27</b>	<b>1</b>	<b>16</b>	<b>1</b>	<b>7</b>	<b>5</b>	<b>2</b>	<b>8</b>	<b>11</b>	<b>0</b>	<b>142</b>
<b>Fr.rél.</b>	<b>0,4</b>	<b>5,2</b>	<b>1,1</b>	<b>1,5</b>	<b>2,6</b>	<b>3,0</b>	<b>2,8</b>	<b>1,1</b>	<b>3,5</b>	<b>3,8</b>	<b>3,7</b>	<b>0,5</b>	<b>3,0</b>	<b>0,3</b>	<b>2,3</b>	<b>2,8</b>	<b>0,6</b>	<b>2,0</b>	<b>3,1</b>	<b>0</b>	<b>2,3</b>

**Tableau 3 : Le corpus Noling** (Tous les articles sont tirés de *Norsk lingvistisk tidsskrift*)

<b>marqueur</b>	<b>01</b>	<b>02</b>	<b>03</b>	<b>04</b>	<b>05</b>	<b>06</b>	<b>07</b>	<b>08</b>	<b>09</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>	<b>total</b>
Anta	8	0	0	0	4	4	8	6	0	2	5	0	0	0	0	0	0	0	0	3	<b>40</b>
Antagelse	1	0	0	0	2	1	0	0	0	2	3	2	0	0	0	0	2	0	0	0	<b>13</b>
Kan	7	2	0	1	10	1	5	7	3	4	3	0	0	5	2	1	9	0	1	1	<b>62</b>
Kanskje	2	0	0	2	3	1	0	4	2	2	3	0	3	0	1	2	1	1	1	2	<b>30</b>
Mulig	3	0	0	0	0	1	1	4	0	3	4	0	0	1	0	0	2	0	0	1	<b>20</b>
Muligens	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	1	0	0	0	1	<b>5</b>
Sannsynligvis	0	0	0	0	0	0	0	0	0	1	2	1	0	2	0	1	1	0	0	0	<b>8</b>
Se ut	9	2	0	6	8	3	6	1	0	2	3	0	0	4	1	1	0	2	0	0	<b>48</b>
Synes	0	1	0	0	0	2	3	4	8	5	0	0	10	0	0	0	0	0	0	0	<b>33</b>
Trolig	0	0	0	1	4	3	0	0	0	0	1	0	4	0	1	1	0	0	0	0	<b>15</b>
Tyde på	2	0	0	0	3	1	0	3	2	1	0	0	0	1	1	0	1	0	0	0	<b>15</b>
<b>total</b>	<b>32</b>	<b>5</b>	<b>0</b>	<b>10</b>	<b>35</b>	<b>17</b>	<b>24</b>	<b>29</b>	<b>15</b>	<b>22</b>	<b>24</b>	<b>3</b>	<b>17</b>	<b>13</b>	<b>7</b>	<b>7</b>	<b>16</b>	<b>3</b>	<b>2</b>	<b>8</b>	<b>289</b>
<b>Fr.rél.</b>	<b>7,2</b>	<b>1,3</b>	<b>0</b>	<b>1,2</b>	<b>4,2</b>	<b>4,1</b>	<b>6,6</b>	<b>3,6</b>	<b>4,4</b>	<b>2,8</b>	<b>5,4</b>	<b>1,2</b>	<b>2,3</b>	<b>3,2</b>	<b>1,8</b>	<b>1,8</b>	<b>7,2</b>	<b>1,0</b>	<b>0,7</b>	<b>2,5</b>	<b>3,2</b>

**Tableau 4 : Le corpus Nomed** (Tous les articles sont tirés de *Tidsskrift for den norske Lægeforening*)

<b>marqueur</b>	<b>01</b>	<b>02</b>	<b>03</b>	<b>04</b>	<b>05</b>	<b>06</b>	<b>07</b>	<b>08</b>	<b>09</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>	<b>total</b>
Anta	0	0	0	0	0	0	0	3	1	0	1	0	0	1	0	1	1	0	0	1	<b>9</b>
Antagelse	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	<b>1</b>
Kan	0	3	6	3	1	4	5	6	0	0	3	1	1	0	0	2	0	3	5	2	<b>45</b>
Kanskje	0	0	1	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	<b>4</b>
Mulig	0	0	3	1	0	0	3	0	0	0	1	0	0	0	0	0	0	1	1	2	<b>12</b>
Muligens	0	1	0	0	2	0	1	0	0	0	0	0	0	0	0	1	0	3	0	0	<b>8</b>
Sannsynligvis	1	0	1	0	0	1	0	2	1	0	1	0	0	0	0	0	0	4	2	0	<b>13</b>
Se ut til	0	0	1	2	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	1	<b>7</b>
Synes	0	2	2	0	2	2	0	2	0	0	1	0	0	2	0	0	0	6	1	0	<b>20</b>
Trolig	0	0	0	0	0	0	0	2	2	1	1	0	0	1	0	0	0	0	0	1	<b>8</b>
Tyde på	0	0	1	0	0	1	1	4	0	0	2	0	1	0	1	0	1	2	2	0	<b>16</b>
<b>total</b>	<b>1</b>	<b>6</b>	<b>15</b>	<b>6</b>	<b>6</b>	<b>8</b>	<b>10</b>	<b>21</b>	<b>5</b>	<b>3</b>	<b>10</b>	<b>1</b>	<b>2</b>	<b>4</b>	<b>1</b>	<b>5</b>	<b>2</b>	<b>19</b>	<b>11</b>	<b>7</b>	<b>143</b>
<b>Fr.rél.</b>	<b>0,6</b>	<b>2,5</b>	<b>5,6</b>	<b>3,2</b>	<b>2,9</b>	<b>3,6</b>	<b>5,2</b>	<b>8,9</b>	<b>2,4</b>	<b>0,8</b>	<b>4,8</b>	<b>0,5</b>	<b>1,6</b>	<b>2,9</b>	<b>0,8</b>	<b>2,3</b>	<b>1,9</b>	<b>4,2</b>	<b>3,9</b>	<b>4,1</b>	<b>3,3</b>

**Tableau 5 : Le corpus Engling** (engling 01- 10 sont tirés de *Journal of Linguistics*, engling 11-15 de *English for Specific Purposes*, engling 16 de *Linguistic Inquiry*, et engling 17- 20 de *Language*.)

<b>Marqueur</b>	<b>01</b>	<b>02</b>	<b>03</b>	<b>04</b>	<b>05</b>	<b>06</b>	<b>07</b>	<b>08</b>	<b>09</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>	<b>total</b>
Appear	2	11	8	1	0	1	6	5	2	1	0	0	3	1	3	10	9	6	1	4	<b>74</b>
Assume	7	9	2	0	11	4	9	4	13	4	1	0	1	0	0	10	11	3	3	0	<b>92</b>
Could	0	1	2	0	1	1	0	0	0	0	0	0	1	0	0	1	0	3	0	0	<b>10</b>
Indicate	0	0	0	0	2	1	2	2	2	0	0	0	4	0	2	0	2	1	2	0	<b>20</b>
May	0	4	0	0	2	8	2	8	0	6	9	3	5	6	0	4	10	1	7	11	<b>86</b>
Might	0	6	0	1	7	0	6	5	0	0	11	0	0	1	2	5	5	6	0	3	<b>58</b>
Perhaps	0	2	1	2	1	3	2	6	0	1	1	0	3	0	9	1	0	2	2	1	<b>37</b>
Possible	0	0	0	0	4	3	1	0	0	0	0	0	0	0	0	0	0	0	0	3	<b>11</b>
Probably	0	0	0	0	0	3	2	1	2	0	1	0	0	1	0	1	0	1	0	1	<b>13</b>
Seem	5	4	0	2	28	2	1	27	4	4	6	1	7	2	4	24	8	2	1	5	<b>137</b>
Suggest	0	12	2	2	18	8	7	1	4	6	1	0	0	0	9	4	21	16	2	8	<b>121</b>
<b>total</b>	<b>14</b>	<b>49</b>	<b>15</b>	<b>8</b>	<b>74</b>	<b>34</b>	<b>38</b>	<b>59</b>	<b>27</b>	<b>22</b>	<b>30</b>	<b>4</b>	<b>24</b>	<b>11</b>	<b>29</b>	<b>60</b>	<b>66</b>	<b>41</b>	<b>18</b>	<b>36</b>	<b>659</b>
<b>Fr.rél.</b>	<b>2,1</b>	<b>6,5</b>	<b>1,5</b>	<b>1,2</b>	<b>9,0</b>	<b>4,3</b>	<b>3,3</b>	<b>3,9</b>	<b>3,0</b>	<b>4,7</b>	<b>5,9</b>	<b>0,6</b>	<b>4,2</b>	<b>2,6</b>	<b>5,6</b>	<b>5,0</b>	<b>3,9</b>	<b>3,6</b>	<b>3,6</b>	<b>3,2</b>	<b>3,9</b>

**Tableau 6: Le corpus Engmed** (engmed 01-05 relèvent de *British Medical Journal*, engmed 06-10 de *International Journal of Cancer*, et 11-20 de *JAMA*.)

marqueur	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	total
Appear	0	0	0	0	0	0	0	2	1	1	0	0	0	0	0	1	0	0	1	0	<b>6</b>
Assume	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	2	<b>4</b>
Could	0	0	0	0	0	0	0	0	0	5	0	0	2	1	0	0	6	0	3	0	<b>17</b>
Indicate	0	0	0	0	1	1	0	2	0	0	1	2	0	0	1	0	0	0	2	1	<b>11</b>
May	1	2	2	0	1	3	9	3	1	12	8	3	6	5	3	5	6	1	2	3	<b>76</b>
Might	2	0	0	2	0	0	0	0	1	3	0	0	1	1	0	5	2	2	3	0	<b>22</b>
Perhaps	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	2	0	0	<b>4</b>
Possible	0	0	1	0	0	0	1	0	1	2	2	0	0	1	1	0	0	1	2	0	<b>12</b>
Probably	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	<b>3</b>
Seem	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	<b>1</b>
Suggest	4	0	0	1	0	1	8	3	1	8	2	1	0	1	0	1	5	6	6	0	<b>48</b>
<b>total</b>	<b>7</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>2</b>	<b>5</b>	<b>18</b>	<b>11</b>	<b>5</b>	<b>31</b>	<b>13</b>	<b>8</b>	<b>9</b>	<b>9</b>	<b>6</b>	<b>13</b>	<b>21</b>	<b>12</b>	<b>19</b>	<b>6</b>	<b>204</b>
<b>Fr.rél.</b>	<b>3,1</b>	<b>0,8</b>	<b>1,5</b>	<b>1,8</b>	<b>0,9</b>	<b>2,7</b>	<b>6,7</b>	<b>3,4</b>	<b>1,9</b>	<b>10,9</b>	<b>4,0</b>	<b>2,6</b>	<b>2,3</b>	<b>2,0</b>	<b>1,4</b>	<b>5,3</b>	<b>8,7</b>	<b>3,1</b>	<b>5,4</b>	<b>1,5</b>	<b>3,4</b>