

Fréquence des marqueurs discursifs en français parlé : quelques problèmes de méthodologie

Catherine CHANET
Université de Provence

Résumé

Cet article a pour double objectif de recenser les problèmes spécifiques posés par l'étude des marqueurs discursifs du français parlé dans les gros corpus, et de donner les fréquences des marqueurs discursifs les plus utilisés à l'oral. Après avoir tenté de définir ce qu'on entend généralement en pragmatique par "marqueurs discursifs", l'article recense les problèmes liés à la catégorisation de ces marqueurs en discours. Il donne alors les fréquences des 85 formes les plus utilisées comme marqueurs discursifs en français parlé, à partir d'un corpus de 1 500 000 mots. Une étude des différents statuts (marqueur ou non) que peuvent revêtir ces formes est ensuite conduite, pour détailler les statuts des quatre formes ambiguës les plus utilisées : *bon*, *bien*, *quoi*, et *enfin*, pour lesquelles les fréquences d'occurrences en tant que particules sont calculées sur le « Corpus de Référence du Français Parlé » (450 000 mots), détaillé dans ce numéro.

1. Objectifs et position du problème

1.1. Objectifs

L'objectif de cet article est de recenser les problèmes méthodologiques que pose une description quantitative des marqueurs discursifs du français parlé sur gros corpus. Deux corpus d'interactions orales seront utilisés ici : le corpus CORPAIX recueilli entre 1977 et 1999 par le G.A.R.S., qui

comporte 1 050 000 mots, et le Corpus de Référence du Français Parlé recueilli par l'équipe DELIC et décrit dans ce numéro, qui en comporte 450 000. Sauf indication contraire, les exemples cités seront tous extraits du Corpus de Référence du Français Parlé.

L'étude des marqueurs discursifs sur ce type de données soulève des questions intéressantes qui n'apparaissent pas nécessairement lorsque les corpus d'étude sont de taille moindre et ne requièrent pas de procédures automatiques d'analyse. Ces questions sont de trois types :

- on rencontre d'abord des problèmes de reconnaissance, et donc de catégorisation, de ces marqueurs en discours : la taille des corpus exclut un relevé manuel des occurrences recherchées, et l'on ne dispose pas encore d'outils permettant d'extraire celles-ci automatiquement de façon satisfaisante ;
- en second lieu, la taille des corpus permet quelques constats d'ordre quantitatif sur la fréquence de ces marqueurs, qui vont parfois à l'encontre des idées reçues en linguistique, et qui, en tout cas, permettent de dégager des pistes de recherche concernant le fonctionnement de ces marqueurs à l'oral ;
- enfin, les emplois relevés apparaissent suffisamment variés pour envisager une description qualitative des fonctions de ces « marqueurs », en se posant la question des méthodologies adaptées à un tel objectif.

Je traiterai ici des deux premières questions, laissant le problème d'une description qualitative de ces marqueurs à des études ultérieures.

1.2. Qu'est-ce qu'un « marqueur discursif » ?

Il est difficile de donner d'emblée une définition de ce qui est entendu ici par « marqueurs discursifs » : c'est précisément l'établissement d'une telle définition qui soulève des questions de méthodologie. Pour néanmoins se faire une idée plus précise de cet objet d'étude, on conviendra que le terme « marqueurs discursifs » (désormais : « MD ») utilisé dans ces pages possède les deux caractéristiques suivantes :

- le terme « marqueurs discursifs », tel qu'il est employé ici, **ne recouvre pas** ce que l'analyse conversationnelle appelle traditionnellement « phatiques » et « régulateurs », lesquels constituent des signaux de pilotage de l'interaction destinés à maintenir / orienter l'échange. On distingue généralement (Cosnier,

1988 ; De Gaulmyn, 1987) les éléments « phatiques » (*tu sais, tu vois, hein, ...*), émis par le locuteur, des éléments « régulateurs » (*hm hm, oui, d'accord, je vois, ...*) émis par l'allocutaire et assurant un rôle de feed-back interactionnel. Ces signaux de régulation ne sont pas seulement verbaux : les gestes et les mimiques (regards, postures, hochements de tête, sourires) jouent un rôle très important dans la synchronisation de l'interaction.

- le terme « marqueurs discursifs » **recouvre** à la fois ce que la littérature pragmatique appelle habituellement « connecteurs », et ce qu'elle appelle « particules » (au sens de Fernandez, 1994). Par exemple, *mais, alors, donc* (habituellement considérés comme connecteurs) font partie des MD, et *ben, voilà, ou en fait* (habituellement considérés comme particules), également.

Après avoir longtemps été définis comme reliant des propositions, des énoncés, ou des actes de langage, les connecteurs semblent désormais décrits par un certain nombre de linguistes (dont Berrendonner, 1983 : 237 ; Mosegaard-Hansen, 1998 : 73) comme des items reliant une information produite par le discours à l'ensemble des représentations mentales antérieurement construites par ce même discours. Quant aux particules, elles sont appréhendées (Fernandez, 1994 : 3 *sqq*) comme des items n'assurant pas nécessairement ce rôle de « liaison », mais donnant des informations sur les opérations conduites par le locuteur dans la construction de son discours.

Connecteurs et particules ont donc en commun le fait de constituer des unités non référentielles (n'ayant pas un signifié dénotatif mais plutôt instructionnel), et le fait d'agir sur les représentations cognitives construites par le discours, et dans la construction de ces représentations. Ces « marqueurs » donneraient des instructions sur la manière dont les interactants peuvent co-construire des représentations, les modifier, et les ajuster les unes aux autres (Chanet, 2001a). Autrement dit, les marqueurs discursifs n'interviendraient pas nécessairement dans la construction discursive d'un univers de référence, mais pourraient donner des indications sur la manière de construire cet univers, et, de façon plus générale, sur les opérations cognitives à conduire par les interactants dans l'activité discursive pour optimiser la communication.

Dans la pratique, la distinction connecteurs *vs* particules n'est pas si claire : d'une part, certains auteurs emploient le terme « particules » là où d'autres emploient celui de « marqueur discursif », et inversement ; d'autre part, il apparaît que, dans certains de leurs emplois, *mais, donc, ou alors*

ont des fonctionnements pragmatiques similaires à ceux des particules, tandis que des formes comme *cependant*, *car*, ou *puisque* sont toujours considérées comme des connecteurs dans la littérature. On n'a donc pas affaire à des catégories disjointes entre connecteurs d'une part et particules d'autre part, ce qui justifie la propension de certains linguistes à les regrouper à l'intérieur d'une classe unique « marqueurs discursifs ».

1.3. Position du problème

Le problème que posent des unités de ce type est de déterminer en quoi, ou de quel point de vue, elles constituent une classe. La terminologie employée dans la littérature est à cet égard révélatrice : selon les observables choisis, les objectifs des auteurs et leurs théories sous-jacentes, on rencontre une ou plusieurs des dénominations suivantes : « mots du discours » (Ducrot *et al.*, 1980), « connecteurs » (pragmatiques et/ou argumentatifs) (Roulet *et al.*, 1985, et bien d'autres), « organisateurs textuels » (que Adam, 1990 oppose aux connecteurs), « opérateurs » (généralement opposés aux connecteurs : Anscombe & Ducrot, 1983, et aussi, dans un autre sens, Rossari, 1989), « marqueurs de structuration de la conversation » (Auchlin, 1981a, 1981b), « marqueurs de reformulation paraphrastique » (Gülich & Kotchi, 1983), « ponctuants » (Vincent, 1993), « particules » (énonciatives ou discursives) (Fernandez, 1994, Mosegaard-Hansen, 1998, et d'autres), « marqueurs discursifs » (Shiffrin, 1987), « marqueurs méta discursifs » (Mosegaard-Hansen, 1995), « marqueurs métalinguistiques » (Cadiot *et al.* 1985), voire « petits mots de l'oral » (Bruxelles & Traverso, 2001) ...

Dans le domaine de la littérature américaine, Fraser (1999) recense encore les appellations suivantes¹ : « cue phrases », « discourse connectives », « discourse particles », « discourse signalling devices », « phatic connectives », « pragmatic connectives », « pragmatic expressions », « pragmatic formatives », « pragmatic markers », « pragmatic operators », « pragmatic particles », « semantic conjuncts », « sentence connectives ».

On le voit, se repérer dans cette multitude d'appellations demande de définir précisément l'extension de chacun de ces termes, et la façon dont ils se distinguent par ailleurs des termes de « modalisateurs », « régulateurs »,

¹ Je renvoie à l'article de Fraser (1999) pour les nombreuses références correspondantes.

et « phatiques » qui recouvrent des objets plus étudiés par les conversationnalistes.

Cette fastidieuse liste non exhaustive permet de dégager deux propriétés attribuées par la littérature aux « marqueurs discursifs » :

- d'une part, les unités en question sont considérées comme relevant de l'étude du discours, et donc (pour les études) de la pragmatique principalement. Très peu de travaux traitent du lien entre rôle pragmatique des marqueurs discursifs et syntaxe, ou entre marqueurs discursifs et prosodie, par exemple, du moins en ce qui concerne le français². Aucune étude, à ma connaissance, n'envisage ces marqueurs en liaison avec l'ensemble des différents « niveaux » traditionnellement distingués en linguistique. Il semble donc généralement admis que le fonctionnement pragmatique d'unités de ce type puisse se décrire à l'intérieur même d'un seul « niveau » pragmatique, ce qui constitue une prise de position épistémologique qui ne va pas de soi. Si l'on considère, à l'inverse, que la pragmatique doit rendre compte d'une construction de l'interprétation qui fait intervenir tous les niveaux traditionnellement distingués dans les sciences du langage, on ne voit guère pourquoi les phénomènes d'ordre discursif n'intéresseraient pas également la syntaxe, la sémantique, et la prosodie.
- d'autre part, dans ce cadre pragmatique, les études adoptent des démarches hétérogènes, et peuvent s'attacher à caractériser des fonctions, des formes, ou les deux à la fois, ce qui donne lieu à trois types d'approches privilégiées :
 - une approche « descendante », qui étudie certaines opérations discursives (opérations de reformulation, ou opérations de structuration de séquences d'énoncés) à travers leurs traces en discours, parmi lesquelles figurent des connecteurs et/ou des particules (Gülich & Kotchi, 1983 ; Rossari, 1997=1993 ; Auchlin, 1981a et 1981b ; Roulet *et al.*, 1985) ;
 - une approche « ascendante », qui cherche à caractériser le fonctionnement pragmatique d'une unité donnée, à travers ses emplois en contexte (étude de *là* – Forget, 1989 –, de *enfin*

² On peut tout de même citer l'étude de Östman (1990) sur le rapport entre particules et prosodie. Plus récemment, Simon et Grobet (2002) ont cherché à mettre en relation emplois de *mais* et de *parce que* et prosodie.

- Cadiot *et al.*, 1985 –, de *ben* – Bruxelles & Traverso, 2001 –, *etc.*) ;
- une approche portant sur une relation entre un ensemble de formes et une fonction considérée comme sémantique (marqueurs de « conséquence » – Hybertie, 1996 – , de « concession » – Morel, 1996 – de « cause » – Nazarenko, 2000 –).

La question qui se pose alors est la suivante : le terme « marqueurs discursifs » désigne-t-il préférentiellement des formes, ou des fonctions résultant d'une interprétation ? La difficulté que recouvre cette question tient précisément au fait que la relation entre une forme et une « fonction » donnée n'est jamais bi-univoque, et que la littérature sur la question ne traite pas de catégories établies formellement indépendamment de toute considération pragmatique. Un problème immédiat se pose donc sur le plan du repérage même des MD, qui ne peuvent être recensés par leur seule forme.

Il apparaît en effet que bon nombre de formes ne constituent pas des MD dans tous leurs emplois, soit parce qu'elles subsument des fonctionnements morphologiques et/ou syntaxiques très différents (le *bon* adjectif de *c'est bon* et le *bon* particule de *bon enfin bref*), soit parce que leurs fonctions discursives sont très différentes (comparer les formes *enfin* dans *il est enfin arrivé* et dans *il a réussi ses examens, enfin je crois*). La classe des MD peut donc difficilement être décrite en termes extensionnels, ce qui pose problème pour la reconnaissance des MD dans les gros corpus.

2. Les formes susceptibles de fonctionner comme MD

Les outils dont on dispose actuellement pour repérer les MD dans les gros corpus se ramènent à deux grands types :

- des concordanciers permettant d'extraire des formes ou des lemmes,
- des outils d'annotation de corpus aux niveaux morphologique et syntaxique.

L'application de concordanciers permet de recueillir les occurrences de formes susceptibles de fonctionner comme des marqueurs en discours. Etant donné qu'une forme donnée n'est pas toujours un marqueur discursif dans tous ses emplois, elle produit donc du bruit. L'idéal serait alors de

disposer d'un moyen d'annoter les corpus en amont, en attribuant une catégorie spécifique aux MD dès le repérage des formes.

2.1. Catégorisations automatiques

Au niveau de l'annotation de ces formes, un problème se pose donc, qui est celui de leur éventuelle catégorisation, et de la pertinence de cette catégorisation. On sait que les « connecteurs » et « particules » sont considérés par les grammaires comme relevant de différentes catégories grammaticales : conjonctions de subordination, conjonction de coordination, prépositions, adverbes, voire interjections. Les analyseurs morphologiques, élaborés pour fonctionner sur des textes écrits pour lesquels la catégorisation de ces types d'unités se pose moins crucialement, sont démunis face à des corpus oraux. A titre d'illustration, le tableau ci-dessous indique les catégories morphologiques qui ont été attribuées par Cordial 7 à certains items issus d'une interaction du « Corpus de Référence du Français Parlé » :

ah bon	I I	Interjection + Interjection
alors	Rgp	Adverbe (autre que comparatif ou de négation)
aussi	Rgp	Adverbe
ben	Rgp	Adverbe
	Ncm	Nom commun (masc, nombre indéterminé)
bon	Rgp	Adverbe
	Afpms	Adjectif qualificatif non comparatif (masc sing)
c'est-à-dire	Rgp	Adverbe
d'ailleurs	Sp Rgp	Préposition + Adverbe
	Sp Dt-fs-	
de toute façon	Ncfs	Préposition + Adjectif indéfini + Nom commun
donc	Cc	Conjonction de coordination
	Da-ms-d	Article défini (masc sing) +
du coup	Ncms	Nom commun (masc sing)
en fait	Sp Ncms	Préposition + Nom commun (masc sing)
enfin	Rgp	Adverbe
là	Rgp	Adverbe
mais	Cc	Conjonction de coordination
même	Rgp	Adverbe
par contre	Sp Ncms	Préposition + Nom commun (masc sing)
	Sp Sp	Préposition + Préposition
par exemple	Sp Ncms	Préposition + Nom commun (masc sing)
parce que	Cs	Conjonction de subordination
puis	Rgp	Adverbe
putain	Ncfs	Nom commun (féminin singulier)
quoi	Pr-..-	Pronom relatif (genre et nombre indéterminés)
	Pr-..a	Pronom relatif objet direct
	Pr-..d	Pronom relatif objet indirect
sinon	Cs	Conjonction de subordination
surtout	Rgp	Adverbe
voilà	Sp	Préposition
vraiment	Rgp	Adverbe

Tableau 1. Étiquetage des marqueurs discursifs de l'oral par Cordial 7 « Corpus de référence du français parlé », PCR-R00PRI001.

2.2. Problèmes

Ces catégorisations sont problématiques à plusieurs titres : elles peuvent ne pas tenir compte du fait que certains MD sont graphiquement complexes, ou distinguer des catégories pour un fonctionnement unique, ou encore, à l'inverse, ranger sous une même étiquette des fonctionnements très différents.

2.2.1. MD graphiquement complexes

Certains marqueurs discursifs dont la forme s'étale sur plusieurs mots graphiques sont traités comme des successions d'items. Si ce choix pourrait à la rigueur être pertinent pour la combinaison *ah bon*, il l'est beaucoup moins pour *d'ailleurs*, *de toute façon*, *du coup*, *en fait*, *par contre* et *par exemple*, qui se voient attribuer des chaînes de catégories morphologiques, et non pas une catégorie unique.

2.2.2. Cas des multi-catégorisations non pertinentes

Dans le cas où une même forme peut se voir attribuer plusieurs catégories exclusives les unes des autres (i.e. dans le cas où la forme est considérée comme ambiguë sur le plan morphologique), cette multi-catégorisation n'est pas toujours cohérente, dans la mesure où elle ne rend pas toujours compte d'une ambiguïté homonymique. Ainsi, *bon*, en dehors de son statut d'adjectif, est considéré comme un adverbe, mais il est étiqueté « interjection » dans la combinaison *ah bon*, sans qu'on sache ici quels arguments pourraient permettre de distinguer adverbe et interjection. Par ailleurs, la forme *ben* est étiquetée « adverbe » lorsqu'elle se situe en début d'énoncé :

- (1) L2 je sais pas combien ça fait mais sinon qu'est-ce que enfin - toi tu sais un petit peu ce que tu aurais tu sais déjà ce que tu vas avoir envie de faire - après plus tard avec - donc passer passer un concours déjà tu sais quelle école tu veux avoir
L1 **ben** je - je vais présenter - les Beaux-Arts certainement Cergy - Art-Déco mais je sens que je vais me faire jeter mais minablement quoi -

et (plus curieusement) « nom commun » ailleurs, comme dans les exemples suivants³ :

³ Le Littré (1873) répertorie effectivement un « ben » substantif masculin, dont la définition est la suivante : « Nom donné au moringa oléifère de Lamarck, arbre dont le fruit, appelé noix de ben, fournit une huile employée principalement dans la parfumerie ». Le Robert (1989) ne le répertorie pas. Que ce terme soit toujours en usage ou pas, on peut se demander s'il est pertinent de le prévoir dans des étiqueteurs morphologiques devant s'appliquer à des textes autres que de spécialité.

- (2) L2 et tu as déjà créé un monté un petit film là
L1 ouais ouais je suis en train là <je suis en train de le finir
L2 c'est vrai> - et vous étiez en équipe
L1 euh **ben** moi j'ai j'ai tourné tout seul - mais j'ai participé au film
d'une copine - euh pour le le tournage quoi - euh sinon moi j'ai j'ai fait
ça tout seul avec des enfin j'ai c'est moi qui filmais
- (3) L2 et vous y allez en rollers
L1 ah non il y a il y a ça ça par contre c'est tendu hein - parce que il y
a de l'eau
L2 d'accord
L1 et - ah **ben** non c'est impossible de descendre en rollers - il y a des
mecs qui descendent avec leur vélo mais je sais pas comment ils font
quoi

2.2.3. Attribution non pertinente d'une catégorie

Enfin, le problème le plus ardu posé par cet étiquetage réside dans le fait que l'attribution d'une catégorie donnée peut être non pertinente en discours ou sujette à caution. Ce peut être le cas notamment pour *bon*, *enfin*, *là*, et *quoi*, pour ne citer que ceux-là.

2.2.3.1. Le cas de *bon*

Bon peut en effet être étiqueté « adjectif » ou « adverbe » selon le contexte, et il se trouve que la valeur d'adjectif n'est pas toujours justifiée, comme l'atteste (4), où *bon* se voit attribuer la valeur "Afpms" (adjectif qualificatif non comparatif), alors qu'il ne se comporte plus ici comme un adjectif :

- (4) L2 c'était quoi c'était un documentaire une petite fiction
L1 bah on avait un sujet c'était faire faire faire - c'est-à-dire filmer une
action ou euh - c'était enfin c'est assez compliqué comme sujet et moi
mon action c'était jeter la pierre et en fait je suis allé - j'ai fait de la
pratique et - je suis allé jeter la pierre et je suis allé en parler avec
enfin j'ai parlé avec des gens X j'ai interrogé des gens - et euh
L2 je- jeter la pierre mais parce que **bon** enfin tu tu le prends dans
<dans le sens
L1 dans le sens euh>
L2 du l- du langage aussi <jeter la pierre à quelqu'un là
L1 ouais ouais> ouais ouais jeter la pierre dans le sens quand on - on
met la faute sur quelqu'un d'autre aussi quoi

2.2.3.2. Le cas de *enfin*

Enfin est systématiquement étiqueté « adverbe ». Si elle peut être justifiée sur le plan morphologique (notamment par l'invariabilité de la forme), l'attribution automatique de cette catégorie unique ne permet pas de différencier des comportements aussi distincts que ceux de (5), (6) et (7) :

- (5) il a **enfin** terminé sa thèse (exemple construit)
- (6) L1 ça c'est - et puis ensuite il y a l'intersaison - où se mélangent cette clientèle de l'hiver plus une clientèle touristique - et **enfin** l'été on peut dire les mois de juillet et août - où là il n' y a pratiquement qu'une clientèle touristique (SAI-PRO001, 13)
- (7) L2 je connais pas moi Rome
L1 ah c'est pff - je sais pas c'est - c'est indescriptible quoi **enfin** c'est pas ça ressemble à Paris et en même temps c'est complètement différent

Deux problèmes se posent ici :

- d'une part, la catégorie « adverbe », comme on le sait, regroupe des unités aux fonctionnements syntaxiques très différents selon leurs emplois. Ainsi, en (5), *enfin* est intégré au constituant verbal et joue le rôle de modifieur du verbe, alors qu'en (6) et (7), il échappe aux réseaux de dépendances syntaxiques ;
- d'autre part, on a ici affaire à des emplois discursifs différents, dans la mesure où seul le *enfin* présent en (7) serait considéré comme une particule (cf. 3.2.5).

2.2.3.3. Le cas de *là*

Le cas de *là*, également catégorisé « adverbe » de manière systématique par Cordial, est plus complexe, dans la mesure où il est assez difficile de trouver des critères qui permettraient de poser des catégories distinctes pour cette forme.

D. Forget (1989) souligne d'ailleurs une certaine « continuité » entre un *là* purement déictique spatial ou temporel (illustré par (8) et (9)), et un *là* (très fréquent en québécois) qu'elle appelle « discursif », et qu'elle considère comme une particule (10) :

- (8) mets ça **là**, sur la table. (< Forget, 1989 : 58)
- (9) - Quand est-ce qu'on y va ?
- On y va **là**, tout de suite. Prépare-toi. (< Forget, 1989 : 58)

- (10) une fois que t'as ajouté tes trois œufs **là**, bon ben tu mélanges pendant cinq minutes (< Forget, 1989 : 65)

Elle signale cependant que les deux *là* se différencieraient sur le plan prosodique.

Or, un logiciel comme Cordial étiquette *là* « adverbe » dans tous les cas. Ainsi, il met sur le même plan les emplois de *là* dans (11) et (12) :

- (11) L2 mais tu fais les circuits - enfin très très cachés quoi parce que normalement <c'est interdit quoi

L1 ah ouais ouais je suis je suis jamais allé> par Denfert-Rochereau **là** où les touristes ils rentrent quoi je sais même pas quelle tronche ça a par là quoi

- (12) L1 j'ai fait de la pratique et - je suis allé jeter la pierre et je suis allé en parler avec enfin j'ai parlé avec des gens X j'ai interrogé des gens - et euh

L2 je- jeter la pierre mais parce que bon enfin tu tu le prends dans <dans le sens

L1 dans le sens euh>

L2 du l- du langage aussi <jeter la pierre à quelqu'un **là**

L1 ouais ouais> ouais ouais jeter la pierre dans le sens quand on - on met la faute sur quelqu'un d'autre aussi quoi

2.2.3.4. Le cas de *quoi*

Enfin, la forme *quoi* est systématiquement catégorisée pronom relatif dans l'interaction passée au crible de Cordial, ce qui ne manque pas de poser un certain nombre de problèmes. Ainsi, *quoi* est considéré comme un pronom relatif par le logiciel, sans autre choix possible, dans les emplois suivants :

- (13) il y a tellement de trucs en fait c'est c'est c'est c'est c'est infini **quoi** les possibilités de cette ville

- (14) L2 il y a des carrières c'est impressionnant **quoi** - c'est - c'est genre - ah ouais c'est c'est c'est c'est il y a quinze mètres de plafond c'est des cathédrales sous terre **quoi** c'est

L1 ah ouais

L2 ouais c'est un truc de fou **quoi** - sérieusement **quoi**

Or, il est évident qu'ici, *quoi* ne possède pas les propriétés d'un pronom, dans la mesure où il ne commute pas avec un autre pronom.

2.2.4. Bilan

Il semble que la recherche des MD à partir de leurs formes, et donc à partir des catégorisations morphologiques de ces formes, pose deux types de problèmes.

- Il existe tout d'abord des cas où le marqueur discursif ne peut plus être considéré comme relevant de la catégorie morphologique qui lui est attribuée : c'est le cas notamment pour *quoi*, qui perd ses propriétés de pronom lorsqu'il devient particule. Dans des cas de ce type, la catégorie morphologique de la forme lorsqu'elle ne constitue pas un MD est en quelque sorte incompatible avec l'emploi de cette forme comme MD : si *quoi* est pronom, alors il n'est pas un MD, et inversement. Il pourrait être utile, pour des cas de ce type, de prévoir dès l'analyse morphologique, le fait que la forme puisse être employée en discours avec un statut tel qu'on ne peut plus lui attribuer les catégories grammaticales habituellement prévues.
- Dans d'autres cas, la catégorie morphologique n'est pas en elle-même problématique : elle occulte simplement le fait que la forme peut également, tout en relevant toujours de la classe qui lui est attribuée, fonctionner comme un MD en discours. L'exemple de *enfin* illustre ce type de cas : il n'est pas incompatible que *enfin* soit catégorisé comme un adverbe sur le plan morphologique et soit employé en tant que particule dans le discours. Le seul problème qui se pose est de savoir dans quelles proportions, ou avec quelle fréquence, l'adverbe en question va fonctionner comme une particule.

La catégorie « marqueurs discursifs » s'avère donc être une catégorie discursive, dont la constitution résulte de l'interprétation en discours de certaines formes. Si l'on veut continuer à parler des MD en terme de « classe », il semble que l'on ait tout intérêt à différencier des **niveaux** de catégorisation ou d'étiquetage, entre lesquels on n'observe pas nécessairement d'isomorphisme : le terme « marqueurs discursifs » peut s'appliquer à certains éléments seulement de la catégorie habituelle des adverbes, des conjonctions, voire des pronoms, dans certains seulement de leurs emplois, où les catégories morphologiques « habituelles » de ces formes peuvent ne plus s'avérer pertinentes.

Je parlerai, dans ce qui suit, de **statut** d'une occurrence pour désigner la catégorie, quel que soit son « niveau » (morphologique / discursif), qui s'avère pertinente pour rendre compte du fonctionnement de cette occurrence donnée.

En l'état actuel des connaissances sur les MD, il n'existe pas d'autre moyen, pour analyser ceux-ci, que d'effectuer un repérage forme par forme, pour ensuite inventorier et classer les statuts de ces formes en discours, et ne retenir que ceux qui relèvent du marquage textuel des activités discursives.

3. Formes et statuts : aspects quantitatifs

3.1. Fréquence des formes

Un premier repérage sur gros corpus des formes susceptibles de fonctionner comme marqueurs discursifs permet de dégager des phénomènes intéressants concernant la fréquence de ces formes en français parlé. Il apparaît en effet que certaines formes pourtant très fréquentes ont été très peu étudiées dans le champ linguistique (comme *bon*, *là*, *quoi*, ou *ben*), alors que d'autres, qui apparaissent plus rarement à l'oral, font l'objet d'une littérature abondante (*car*, *à cause de*, *puisque*, par exemple).

Le tableau ci-dessous fait apparaître le nombre d'occurrences obtenues sur deux corpus de français parlé : CORPAIX, (1 050 000 mots), et le « Corpus de Référence du Français Parlé » (CRFP, 450 000 mots). L'ensemble des deux corpus totalise donc 1 500 000 mots. La somme des occurrences pour chaque corpus permet de classer les formes par ordre de fréquence. En se fondant sur un débit moyen de 200 mots par minute, on donne, dans la dernière colonne, la fréquence avec laquelle ces formes apparaissent dans la parole, en jours, heures, minutes et secondes. Il est bien évident que ces chiffres indiquent des moyennes qui supposent que l'on parle 24 heures sur 24 et qui ne tiennent absolument pas compte de variations sociolinguistiques ou idiolectales⁴.

⁴ Ces fréquences ne tiennent pas non plus compte de la répartition des occurrences au sein d'un même discours, selon les tâches (narration, argumentation, ...) qui y sont effectuées.

Fréquence des marqueurs discursifs en français parlé

	Forme	CORPAIX	CRFP	TOTAL	Soit 1 tou(te)s les
1	mais	7520	3075	10595	42s
2	donc	4268	3807	8075	55s
3	alors	4476	1683	6159	1mn 13s
4	bon	3564	2144	5708	1mn 18s
5	là	3585	1536	5121	1mn 27s
6	bien	3337	1550	4887	1mn 32s
7	parce que	3155	1568	4723	1mn 35s
8	quoi	2366	1497	3863	1mn 56s
9	ben	2754	1001	3755	1mn 59s
10	puis	2542	1153	3695	2mn 01s
11	enfin	2429	1145	3574	2mn 05s
12	aussi	1888	1045	2933	2mn 33s
13	voilà	1674	928	2602	2mn 52s
14	après	1525	755	2280	3mn 16s
15	quand même	1254	702	1956	3mn 48s
16	en fait	1035	719	1754	4mn 13s
17	par exemple	943	425	1368	5mn 27s
18	c'est-à-dire	584	424	1008	7mn 22s
19	puisque	474	274	748	9mn 58s
20	surtout	488	168	656	11mn 23s
21	justement	438	186	624	11mn 57s
22	plutôt	372	142	514	14mn 31s
23	d'ailleurs	304	200	504	14mn 48s
24	ensuite	339	165	504	14mn 48s
25	disons	298	130	428	17mn 26s
26	sinon	263	128	391	19mn 05s
27	en plus	238	142	380	19mn 38s
28	finalement	203	93	296	25mn 13s
29	par contre	180	95	275	27mn 08s
30	seulement	209	60	269	27mn 44s
31	de toute façon	164	62	226	33mn 01s
32	car	160	41	201	37mn 08s
33	pourtant	112	33	145	51mn 28s
34	or	61	81	142	52mn 33s
35	de plus	93	37	130	57mn 24s
36	en tout cas	95	35	130	57mn 24s
37	notamment	60	57	117	1h 04mn
38	tandis que	72	42	114	1h 5mn 30s
39	malgré	46	44	90	1h 23mn
40	au fond	66	23	89	1h 24mn
41	ainsi	52	26	78	1h 46mn
42	à cause de	60	14	74	1h 41mn

43	en effet	45	20	65	1h 55 mn
44	à propos	47	9	56	2h 13mn
45	du moins	34	20	54	2h 18mn
46	en ce qui concerne	40	13	53	2h 21mn
47	du coup	19	29	48	2h 35mn
48	bref	31	14	45	2h 46mn
49	malgré tout	20	24	44	2h 50mn
50	précisément	30	9	39	3h 11 mn
51	étant donné que	25	10	35	3h 33mn
52	en fin de compte	18	15	33	3h 46mn
53	après tout	23	10	33	3h 46mn
54	par conséquent	27	6	33	3h 46mn
55	tout de même	28	4	32	3h 53mn
56	par ailleurs	16	12	28	4h 46mn
57	concernant	17	7	24	5h 11mn
58	au contraire	0	23	23	5h 24mn
59	quant à	15	7	22	5h 39mn
60	si bien que	14	7	21	5h 55mn
61	en l'occurrence	8	12	20	6h 13mn
62	autrement dit	6	13	19	6h 33mn
63	au fait	11	7	18	6h 55mn
64	en réalité	14	4	18	6h 55mn
65	cependant	15	3	18	6h 55mn
66	de toute manière	8	9	17	7h 19mn
67	certes	8	8	16	7h 46mn
68	en revanche	13	3	16	7h 46mn
69	en tous les cas	11	4	15	8h 17mn
70	néanmoins	6	7	13	9h 34mn
71	inversement	11	1	12	10h 22mn
72	au total	6	5	11	11h 18mn
73	en somme	10	1	11	11h 18mn
74	n'empêche	7	3	10	12h 26mn
75	de ce fait	4	4	8	15h 33mn
76	au bout du compte	3	3	6	20h 44mn
77	en définitive	6	0	6	20h 44mn
78	de fait	6	0	6	20h 44mn
79	à l'inverse	0	5	5	1j 52mn
80	somme toute	1	3	4	1j 7h 06mn
81	de sorte que	1	2	3	1j 17h 27mn
82	à tel point que		2	2	2j 14h 11mn
83	toutefois	1	1	2	2j 14h 11mn
84	en outre	1	1	2	2j 14h 11mn
85	tout compte fait	1	0	1	5j 4h 22mn

3.2. Quelques commentaires à propos de ces fréquences

- La liste des marqueurs recherchés n'est pas exhaustive : elle a été établie a priori en recueillant les formes examinées par la littérature, y compris dans des dictionnaires pour le Français Langue Etrangère comme celui de Grieve (1996). Elle a ensuite été enrichie en fonction des corpus.
- Les formes les plus fréquentes (apparaissant au moins toutes les 2mn dans la conversation) sont essentiellement monosyllabiques, comme on le vérifiera aisément sur les dix premiers items qui, si l'on excepte *alors* et *parce que*, ne comportent qu'une seule syllabe. Notons d'ailleurs que *alors*, *parce que* et aussi *enfin*, qui intervient en 11e position, peuvent se réaliser également sous les formes /lOR/, /pask/ et /fe~/ respectivement⁵, c'est-à-dire comme monosyllabiques. Les caractéristiques phonologiques des items ne suffisent pas, bien sûr, pour expliquer leurs fréquences : *car*, *or*, et *bref* sont monosyllabiques et pourtant plus rares.
- Parmi les formes les plus fréquentes figurent des items généralement considérés comme connecteurs (*mais*, *donc*, *alors*, *parce que*), mais aussi, de façon peut-être moins attendue, des formes à fonctionnement de particule (*bon*, *là*, *quoi*, *ben*). Or, la littérature sur l'enseignement du français (langue maternelle ou langue seconde) met exclusivement l'accent sur les connecteurs : les particules ne sont jamais traitées et sont généralement catégorisées comme des « disfluences » en quelque sorte parasites, alors qu'elles sont manifestement très employées.

Ces données brutes ne permettent cependant pas d'évaluer l'importance quantitative de l'emploi des MD à l'oral, dans la mesure où, dans beaucoup de cas, une même forme peut recouvrir différents statuts. Il convient donc, pour chaque forme, de répertorier les différents statuts que celle-ci peut revêtir en discours.

⁵ Notation phonétique SAMPA.

3.3. Différencier des statuts possibles pour les formes

3.3.1. Objectifs

Parmi les occurrences repérées dans les transcriptions, toutes ne sont pas exploitables : certaines ont en effet été perçues avec plus ou moins de certitude par le(s) transcrip-teur(s) et figurent dans ce que nous appelons une multi-transcription, du type de celle mise en gras en (15) :

- (15) non parce que ça ça tu dis trois cents kilomètres mais du coup /**enfin**,
ben/ il y a des rats

Par mesure de précaution, ces occurrences ont donc été écartées de celles à analyser.

La recherche de différents « statuts » potentiels pour une même forme a été effectuée sur toutes les occurrences exploitables des 11 items les plus fréquents dans le Corpus de Référence du Français Parlé susceptibles de constituer des MD, de manière manuelle et totalement intuitive. L'objectif à terme de ce premier classement est d'essayer de dégager des critères rigoureux qui permettraient d'établir (ou de modifier, voire de contester) ces catégories intuitives. Seules les transcriptions ont été examinées : les statuts ont donc été établis en l'absence de toute indication prosodique.

Parmi ces 11 formes, 5 peuvent être ambiguës concernant leur statut, c'est-à-dire recevoir d'autres statuts en discours que celui de connecteur et/ou celui de particule : *bon*, *là*, *bien*, *quoi*, et *enfin*. Le cas de *là* ne sera pas examiné ci-dessous, en raison du fait que la prosodie est la plupart du temps nécessaire pour décider si cette forme a un statut de MD ou pas.

3.3.2. Le cas de *bon*

La forme *bon* peut a priori revêtir trois statuts différents au moins dans le discours : adjectif (*c'est un bon élève*), particule (*bon ben voilà*), ou membre d'une locution. Les cas qui ont été considérés comme « membre d'une locution » concernent *ah bon*, *pour de bon*, *à quoi bon*. La grammaire traditionnelle répertorie également un statut supplémentaire d'adverbe, dans des constructions du type *ça sent bon*, ou *il fait bon*, où *bon* peut en fait être vu comme un modifieur du verbe.

Si l'on excepte les occurrences ayant fait l'objet d'une multi-transcription (qui concernent 36 cas sur 2144, et qui laissent à 2108 le nombre d'occurrences exploitables), les statuts de la forme *bon* dans le Corpus de Référence du Français parlé se répartissent de la façon suivante :

Occurrences exploitables	Particules	Adjectif	dans locution	Ambigu	Adverbe
2108	1824	137	105	40	2
100%	86,54%	6,50%	4,98%	1,89%	0,09%

On constate donc que le statut d' « adverbe » est quasiment négligeable dans le corpus, alors que celui de particule y est très nettement prédominant (plus de 86%).

3.3.3. Le cas de *bien*

La forme *bien* est sans doute l'une des plus polyvalentes, dans la mesure où elle revêt des statuts multiples dans les corpus oraux :

- elle peut d'abord faire partie de ce que les grammairistes considèrent comme une conjonction de subordination, telle que *bien que* ou *si bien que*
- elle peut également, dans *ou bien*, être membre d'une « conjonction de coordination »
- elle peut être un substantif (*avoir des biens, faire du bien*)
- elle entre ensuite dans des locutions plus ou moins figées telles que *bien sûr, bien entendu, voire bien évidemment*
- elle peut jouer un rôle d'adverbe et/ou de quantificateur, comme en (16) et (17) :
 - (16) L2 [...]je n'ai rien contre les chiens parce que moi j'ai eu des chiens hein et je les aime **bien** mais
 - (17) alors beaucoup d'entre vous ne savent pas très **bien** dessiner
- elle peut faire partie de la particule *eh bien*
- elle peut constituer une particule à part entière, comme en (18) :
 - (18) je voudrais vous dire donc en français - aujourd'hui quelques mots - relatifs - au thème - de la femme - dans le théâtre lopesque - ce qui nous conduit à tout naturellement bien sûr - à étudier ultérieurement - le personnage féminin de notre comedia{mot étranger} je parle ça va je parle pas trop fort - **bien** - donc si vous voulez - euh aujourd'hui laissons de côté le dramatis personae{mots étrangers} - nous y reviendrons ultérieurement - je vais donc vous parler - du thème de la femme - dans la production euh dramatique lopesque

Ces différents statuts sont ceux qui ont été observés sur la totalité des occurrences dans le Corpus de Référence du Français Parlé, après avoir éliminé dix cas de multi-transcriptions. Ils se répartissent comme suit :

Occ expl	Adverbe / quantif	dans locution	eh bien	conj de subord	conj de coord	particule	ambigu	nom
1550	1019	300	138	26	20	19	15	3
100%	66%	19%	9%	1,7%	1,3%	1,2%	0,9%	0,2%

A l'inverse de ce qui se passe pour les autres formes susceptibles de constituer des MD, le statut de particule est très minoritaire à l'oral : si l'on additionne les occurrences de *bien* particule et de *eh bien*, on obtient un total de 10,2% des emplois où *bien* fonctionne comme une particule.

Ce faible score pourrait sans doute s'expliquer par le fait que *eh bien* est concurrencé à l'oral par *eh ben*, et que les cas où *bien* constitue à lui seul une particule se rencontrent surtout dans les discours pré-planifiés : la moitié des *bien* particules du corpus figurent dans des interactions publiques.

3.3.4. Le cas de *quoi*

Une précédente étude sur CORPAIX (Chanet 2001) avait montré que *quoi* se rencontre sous trois statuts différents dans les corpus de français parlé :

- proforme (ou pronom)
- membre d'une locution plus ou moins figée : *n'importe quoi, comme quoi* (sous certaines réserves)
- particule énonciative.

Après avoir éliminé 11 cas de multi-transcriptions (sur 1497 formes), voici comment se répartissent ces différents statuts dans le Corpus de Référence du Français Parlé :

Occurrences exploitables	Particules	Pronom	dans locution	Ambigu
1486	1134	246	70	36
100%	76,31%	16,56%	4,71%	2,42%

Il est intéressant de constater que la proportion de particules est quasiment similaire à celle qui avait été établie pour CORPAIX (70,5%).

On peut donc considérer qu'à l'oral, les trois quarts environ des *quoi* prononcés sont des particules.

3.3.5. Le cas de *enfin*

A ma connaissance (Chanet 2003), *enfin* est considéré dans la littérature comme pouvant recevoir quatre statuts différents :

- adverbe ne jouant pas le rôle d'un MD, considéré comme de sémantisme « temporel », et comme véhiculant des connotations de « soulagement » du locuteur :

il a **enfin** terminé sa thèse

- connecteur temporel, reliant chronologiquement des événements auxquels il est fait référence (c'est-à-dire se situant dans le monde construit par le discours) :

il a d'abord fait ça, puis ça, et **enfin** ça

- organisateur du texte (Adam, 1990), reliant chronologiquement des opérations discursives effectuées par le locuteur (c'est-à-dire se situant au niveau de l'activité discursive même), souvent dans une activité d'énumération :

à la base, d'abord, c'était un carré de coton bleu [...] ; puis se tenait au second étage un donjon de gâteau de Savoie [...] ; et **enfin**, sur la plate-forme supérieure, on voyait un petit Amour [...] (< Adam 1990 : 154)

- particule énonciative (Fernandez 1994)

il me semble que les gens au fond éprouvent ce genre de chose s'ils sont un peu an- peut-être un peu **enfin** anormaux en quelque sorte mais bon (< Fernandez 1994 : 184)

Il se trouve que ces quatre statuts ne sont pas représentés à l'oral : on ne trouve pas de « connecteur temporel » dans les conversations en face à face, du moins dans le Corpus de Référence du Français Parlé. Par ailleurs, comme le montre le tableau ci-dessous, il est assez frappant de constater que l'oral comporte presque exclusivement des particules :

Occurrences exploitables	Particules	Ambigu	Organisateur	Adverbe régi	Connecteur
1131	1109	19	2	1	0
100%	98,06%	1,68%	0,18%	0,08%	/

Ces chiffres laissent imaginer la gymnastique mentale que doivent effectuer les enfants passant d'une perception orale de la langue à un apprentissage de l'écrit, où *enfin* n'est particule que dans les simulations d'oralité, et où la majorité des emplois sont ceux d'organisateur du texte, et d'adverbe ne jouant pas un rôle de MD⁶.

Par ailleurs, les organisateurs textuels sont également des marqueurs discursifs. Au total, il n'existe qu'une seule occurrence de *enfin* (sur 1131), dans le Corpus de Référence du Français Parlé, dont on est sûr qu'elle n'a pas un statut de MD.

3.4. Fréquences des occurrences de MD

Les chiffres précédents montrent, s'il en était besoin, qu'il ne suffit pas de repérer une forme donnée pour repérer un MD. Bien plus, la fréquence du statut de MD pour certaines formes semble dépendre des genres discursifs et donc, à l'oral, des types d'interaction (*cf.* le cas de *bien*).

En ce qui concerne les fréquences de MD tous types d'interactions confondus, le tableau donné en 3.1. peut donc être revu, en éliminant les cas de multi-transcriptions et en ne retenant que les statuts de MD des formes.

Les fréquences obtenues dans le Corpus de Référence du Français Parlé pour le statut de MD de *bon*, *bien*, *quoi* et *enfin* sont résumées ci-dessous :

forme	Nb formes	Occurrences exploitables	formes à statut de MD	soit, en % des occurrences	soit 1 tou(te)s les
bon	2144	2108	1824	86,54%	1mn 13s
bien	1550	1550	157	10,2%	14mn 10s
quoi	1497	1486	1134	76,31%	1mn 57s
enfin	1145	1131	1011	98,24%	2mn 12s

On le voit, les fréquences restent élevées et ne sont pratiquement pas modifiées, excepté pour *bien*. On peut donc en conclure que l'étude des marqueurs discursifs de l'oral reste une priorité dans la description du

⁶ Un repérage dans la presse (Le Monde diplomatique) sur 150 occurrences avait permis d'établir que dans 50% des cas, *enfin* n'était pas un MD, et que dans 40% des cas, il fonctionnait comme organisateur textuel (Chanet 2003).

français parlé et pour l'enseignement du Français Langue Etrangère, étant donné la fréquence d'occurrence des MD à l'oral.

4. Conclusions

Les fréquences des MD en français parlé montrent que l'on a tout intérêt, si l'on veut traiter de grands corpus, à pouvoir repérer ceux-ci de la manière la plus précoce possible. Les informations quantitatives, qui permettent de prédire la probabilité qu'une forme donnée soit employée comme un MD, peuvent être importantes : il est par exemple intéressant de constater que dans le Corpus de Référence du Français Parlé, plus de 98% des *enfin* sont employés comme particules. Mais ces chiffres demandent d'une part à être confirmés par l'étude d'autres corpus, et d'autre part, à être établis sur des bases autres qu'intuitives.

Le travail qui reste à établir dans l'optique d'un éventuel repérage automatique des MD est considérable, car il nécessite un inventaire minutieux des propriétés « distributionnelles » des MD. Il apparaît en effet que ces propriétés sont hétérogènes, et que les particules, à la différence des connecteurs, sont toujours hors des dépendances syntaxiques, ce qui rend leur approche assez complexe. Pourtant, les particules n'apparaissent pas toutes aux mêmes endroits dans la chaîne parlée, et certainement pas de manière anarchique (Teston 2002). On a affaire ici à des contraintes de positionnement qui dépassent (dans le cas des particules au moins) celles de la syntaxe, et qui restent à mettre au jour de manière systématique. En définitive, la description des marqueurs discursifs dans de grands corpus conduit à étudier ces unités dans tous leurs aspects (morphologiques, micro-syntaxiques, macro-syntaxiques, et discursifs), et, en constituant un lieu d'étude privilégié des pratiques de l'oral, questionne les méthodologies usuelles en pragmatique.

Références

- Adam, J.-M. (1990). *Éléments de linguistique textuelle*. Bruxelles : Mardaga.
- Anscombre, J.C., & Ducrot, O. (1983). *L'argumentation dans la langue*. Bruxelles : Mardaga.
- Auchlin, A. (1981a). *Mais hein, pis bon, ben alors, voilà, quoi!* Marqueurs de structuration et complétude. *Cahiers de Linguistique Française* 2, 141-159.
- Auchlin, A. (1981b). Réflexions sur les marqueurs de structuration de la conversation. *Études de Linguistique Appliquée* 44, 88-103.
- Bruxelles, S., & Traverso V. (2001). *Ben* : apports de la description d'un « petit mot », du discours à l'étude des polylogues. Revue en ligne *Marges Linguistiques* 2, 38-55. (<http://www.marges-linguistiques.com>)
- Cadiot, A., & Ducrot, O., Fradin, B., & Nguyen, T.B. (1985). *Enfin*, marqueur métalinguistique. *Journal of Pragmatics* 9, 199-239.
- Chanet, C. (2001a). Connecteurs, particules, et représentations cognitives de la planification discursive. In E.T. Németh (Ed.), *Cognition in language use : Selected papers from the 7th International Pragmatics Conference, Vol.1*, (pp.44-55). Antwerp : International Pragmatics Association.
- Chanet, C. (2001b). 1700 occurrences de la particule *quoi* en français parlé contemporain : approche de la « distribution » et des fonctions en discours. Revue en ligne *Marges Linguistiques* 2, 56-80. (<http://www.marges-linguistiques.com>).
- Chanet, C. (2003). La forme *enfin* en français parlé contemporain : vers une typologie des statuts et des emplois. *8 Simposio de Comunicación social (Santiago de Cuba, 20-24 janvier 2003), Actas I*, Santiago de Cuba : Centro de Linguística Aplicada, 394-399
- Cosnier, J. (1988). Grands tours et petits tours. In J. Cosnier., N. Gelas, & C. Kerbrat-Orecchioni (Eds.), *Echanges sur la conversation*, (pp.175-184). Lyon : Editions du CNRS.
- De Gaulmyn, M.-M. (1987). Les régulateurs verbaux : le contrôle des récepteurs. In : J. Cosnier., & C. Kerbrat-Orecchioni (Eds.), *Décrire la Conversation*, (pp. 203-223). Lyon : Presses Universitaires de Lyon.
- Ducrot, O. (1980). *Les mots du discours*. Paris : Minuit.
- Fernandez, M. M. J. (1994). *Les particules énonciatives*. Paris : Presses Universitaires de France.
- Forget, D. (1989). *Là*. Un marqueur de pertinence discursive. *Revue Québécoise de Linguistique Vol.18-1*, 57-83.
- Fraser, B. (1999). What are discourse markers ? *Journal of Pragmatics* 31, 931-952.
- Grieve, J. (1996). *Dictionary of contemporary french connectors*. London / New-York : Routledge.
- Gülich, E., & Kotschi, T. (1983). Les marqueurs de reformulation paraphrastique. *Cahiers de Linguistique Française* 5, 305-346.
- Hybertie, C. (1996). *La conséquence en français*. Paris/Gap : Ophrys.
- Morel, M.-A. (1996). *La concession en français*. Paris/Gap : Ophrys.
- Mosegaard-Hansen, M.-B. (1995). Marqueurs métadiscursifs en français parlé : l'exemple de *bon* et de *ben*. *Le français moderne vol. 63-1*, 20-41.

- Mosegaard-Hansen, M.-B. (1998). *The functions of discourse particles*. Amsterdam/Philadelphia : John Benjamins Publishing Company.
- Nazarenko, A. (2000). *La cause et son expression en français*. Paris/Gap : Ophrys.
- Östman, J.O. (1990). Particles and Prosody. On the language internal interactions of prosody and pragmatic particles. In J. Verschueren, (Ed.), *Levels of linguistic adaptation*, (pp. 5-35). Amsterdam : John Benjamins Publishing Company.
- Rossari, C. (1989). Analyse contrastive en français et en italien de l'adverbial *enfin*. In Ch. Rubattel (Ed.), *Modèles du discours en Suisse romande*, (pp. 215-223) Berne : Peter Lang.
- Rossari, C. (1997). *Les opérations de reformulation*. Berne : Peter Lang.
- Roulet, E., Auchlin, R., Schelling, M., Moeschler, J., & Rubattel, C. (1985). *L'articulation du discours en français contemporain*. Berne : Peter Lang.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge : Cambridge University Press.
- Simon, A.-C., & Grobet, A. (2002). Intégration ou autonomisation prosodique des connecteurs. *Proceedings of the 1st International Conference on Speech Prosody, Aix-en-Provence, (Speech Prosody 2002) 11-13 April 2002*, 647-650.
- Teston, S. (2002). *Eclairée par les apports de la pragmatique et de la prosodie, une analyse syntaxique des particules discursives*. Mémoire de D.E.A. de Linguistique Française. Université de Provence.
- Vincent, D. (1993). *Les ponctuels de la langue et autres mots du discours*. Québec : Nuit Blanche Éditeur.