

Comment les données de corpus pourraient renouveler les manuels de grammaire? Illustrations à partir de quelques pronoms et de la forme même

Mireille Bilger, Paul Cappeau

► To cite this version:

Mireille Bilger, Paul Cappeau. Comment les données de corpus pourraient renouveler les manuels de grammaire? Illustrations à partir de quelques pronoms et de la forme même. *Linx*: bulletin du Centre de recherches linguistiques de Paris X Nanterre, Nanterre: Centre de recherches linguistiques de Paris 10, 2013, Corpus et apprentissage du français, p. 177-199. <<http://journals.openedition.org/linx/1526>>. <10.4000/linx.1526>. <hal-01221168>

HAL Id: hal-01221168

<https://hal-univ-perp.archives-ouvertes.fr/hal-01221168>

Submitted on 14 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Linx

Revue des linguistes de l'université Paris X Nanterre

68-69 | 2013

Corpus et apprentissage du français

Comment les données de corpus pourraient renouveler les manuels de grammaire ?

Illustrations à partir de quelques pronoms et de la forme même

Mireille Bilger et Paul Cappeau



Édition électronique

URL : <http://linx.revues.org/1526>

DOI : 10.4000/linx.1526

ISSN : 2118-9692

Éditeur

Université Paris Ouest – département
Sciences du langage

Édition imprimée

Date de publication : 19 novembre 2013

Pagination : 177-199

ISSN : 0246-8743

Référence électronique

Mireille Bilger et Paul Cappeau, « Comment les données de corpus pourraient renouveler les manuels de grammaire ? », *Linx* [En ligne], 68-69 | 2013, mis en ligne le 19 novembre 2015, consulté le 02 octobre 2016. URL : <http://linx.revues.org/1526> ; DOI : 10.4000/linx.1526

Ce document est un fac-similé de l'édition imprimée.

© Tous droits réservés

***Comment les données de corpus
pourraient renouveler les manuels
de grammaire ?
Illustrations à partir de quelques pronoms
et de la forme même***

Mireille Bilger, Université de Perpignan Via Domitia, CRESEM

Paul Cappeau, Université de Poitiers, FoReLL

1. Introduction

L'exploitation de grands corpus – de nature et de « genres » différents (oral/écrit, littérature, presse, oral de parole privée, de parole publique, etc.) – a permis de jeter un éclairage neuf sur de nombreux faits de langue et d'en renouveler à la fois l'analyse et la présentation. Ce constat est indéniable notamment pour l'anglais via l'ouvrage de Biber et al. (1999) ou encore, dans le domaine de la lexicographie, le projet Cobuild initié par Sinclair (Sinclair, 1991) ou plus récemment le projet de création de manuels, *Touchstone* (McCarthy, 2004). En revanche, en ce qui concerne la langue française, ces types de travaux sont encore peu nombreux ou bien trop parcellaires. Pourtant, comme le signalent Arnavielle et Siouffi (2012 : 6) :

La réflexion générale autour de la grammaire est active, à l'heure actuelle. Plusieurs colloques, séminaires, journées d'étude, rencontres ont porté, récemment, sur la difficile question de l'adéquation entre outils scientifiques de description et visées (par exemple pédagogiques au sens large) de la grammaire.

ou encore Cadet et Guerin (2012 : 4) :

On peut s'interroger sur ce qui est effectivement en jeu dans les différentes didactiques. L'enseignement d'une langue, quelle que soit la perspective (langue

maternelle, étrangère ou seconde) pose d'emblée le problème de la sélection des formes sur lesquelles repose le savoir transmis : il ne peut s'agir de la langue dans toute l'étendue de son hétérogénéité et de sa dynamique.

Ces citations récentes tirées de deux numéros de revues (*Langue Française* et *Le français aujourd'hui*) illustrent bien l'impact que l'exploitation des corpus pourrait avoir aujourd'hui, en France, sur les réflexions touchant à l'enseignement de la grammaire et de la langue maternelle LM ou étrangère LE.

En ce qui concerne les manuels, c'est essentiellement dans le domaine du FLE que l'on trouve des exemples intéressants de choix originaux dans l'approche grammaticale, et certains ouvrages, déjà anciens, tels que Job, Mis et Pissavy (1986), Bérard et Lavenne (1991), De Salins (1996), ou plus récents, comme celui de Lafon et Zeggah-Wuyts (2009), font preuve d'une singularité qui ne trouve pas de réel équivalent dans le domaine du FLM.

Ce décalage s'explique sans aucun doute, comme le montre Vigner (2003), par l'histoire, les traditions et les différentes fonctions que l'on assigne à la langue enseignée qui, de fait, ne sont pas de même nature. Blanche-Benveniste et Jeanjean (1986) ont également montré comment, en voulant répondre aux préoccupations de l'enseignement du FLE (donner une version réaliste de la langue), débutent, dès les années 1950, le recueil de données orales attestées et leur analyse. Au départ, ces dernières sont utilisées pour une meilleure approche du lexique ; par la suite, elles vont timidement être exploitées en vue d'une présentation renouvelée des formes grammaticales (Chambers, 2009).

Ce rapide panorama et ces tendances de fond ne doivent pourtant pas masquer l'essentiel : la grammaire présentée traditionnellement comme celle de la langue est, généralement dans la tradition française, la grammaire d'un usage particulier : la langue écrite littéraire.

La linguistique sur corpus, qui permet à la fois une étude qualitative et quantitative des formes étudiées (Halliday, 1991), révèle que très régulièrement les emplois de telle ou telle forme ne sont pas uniformes mais se distribuent de manière différenciée selon les types et les « genres » de productions (Bilger, 2008 ; Bilger & Cappeau, 2004). En ce sens, cette approche méthodologique peut conduire à un réaménagement dans la liste des formes présentées qui tient mieux compte de la fréquence (cf. l'exemple des relatifs composés). Elle peut déplacer la frontière entre faits de lexique et faits de grammaire en montrant que certaines associations de clitiques compléments sont réservées pour l'essentiel à quelques verbes (cf. l'exemple sur l'ordre des pronoms compléments). Elle peut aussi conduire à enrichir le contenu grammatical lui-même en faisant apparaître des structures peu présentes à l'écrit et pour cela ignorées des grammaires alors qu'elles sont fortement présentes à l'oral (cf. l'exemple sur la forme *même*).

L'objectif de cette contribution est de montrer, à travers ces divers exemples, comment la prise en compte de la diversité des usages permettrait de modifier de

façon plus ou moins sensible la description de certains faits de langue et au-delà la représentation de la grammaire dans les manuels (Conrad, 2000).

2. La situation des corpus oraux en France

La situation des corpus en France est quelque peu singulière et atypique. L'élaboration de Frantext (dans les années 1970) constituait une avancée majeure et plutôt novatrice dont la portée a toutefois été fortement limitée par deux facteurs : l'orientation littéraire de plus de 80% des documents rassemblés et la non prise en compte de documents oraux. Des tentatives originales très anciennes ont pourtant, à différentes périodes, été amorcées : les Archives de la Parole (dans les années 1910), le français fondamental (années 1950), le corpus d'Orléans (années 1970), le corpus du GARS (à partir des années 1980). Il existait donc à côté de l'ambitieux projet sur l'écrit un matériau oral (ainsi qu'une réflexion et un savoir-faire) qui a été quelque peu négligé. C'est probablement de ce rendez-vous manqué avec la langue parlée (et de son corollaire : une surreprésentation de l'écrit littéraire) que l'on peut dater l'important retard pris par la France dans la constitution de corpus représentatifs de la langue en usage. Retard qui n'a pas touché aussi fortement d'autres langues européennes telles que l'anglais, l'italien, le portugais ou l'espagnol, car l'emprise de la langue écrite littéraire était certainement moins marquée.

Il faut dès lors attendre le début du 21^e siècle pour que plusieurs projets donnent accès de façon simplifiée à des données orales et à des transcriptions. Depuis les années 2000, le paysage est, de fait, en pleine transformation et plusieurs grands projets, élaborés dans des perspectives différentes, ont émergé et continuent à se développer. On peut citer notamment¹ :

- ESLO (Enquêtes sociolinguistiques à Orléans) qui a nettoyé le célèbre corpus d'Orléans (ESLO 1) et a prolongé ce travail (ESLO 2, Baude & Dugua, 2011) ;
- PFC (Phonologie du français contemporain) qui a développé des enquêtes (lectures, entretiens) à travers le territoire français et l'a élargi à la francophonie (Detey et al., 2010) ;
- CFPP2000 (Corpus de français parlé parisien des années 2000) qui est constitué d'entretiens longs avec des locuteurs parisiens (Branca-Rosoff et al., 2011) ;
- CRFP (Corpus de référence du français parlé) – dont on attend toujours la disponibilité – collecte de situations (privée, publique et professionnelle) dans les grandes villes de l'hexagone (Equipe DELIC, 2004) ;
- CIEL-F (Corpus d'interactions écologiques de la langue française) constitué de situations non provoquées, recueillies dans la francophonie (Gadet et al., 2009).

Ce rapide panorama montre qu'il est aujourd'hui envisageable, pour un large public de spécialistes mais aussi d'utilisateurs avertis, de disposer d'échantillons larges d'oral attesté, produit dans des situations variées, et qu'il est également possible de

¹ Un inventaire détaillé des corpus oraux figure sur le site de l'IRCOM (<http://ircom.corpus-ir.fr/site/corpus.php>). Voir aussi Delahaie, ce volume.

recupérer des exemples (transcription + son) de multiples structures ou de valider des intuitions sur certains usages (comme le montreront les paragraphes suivants).

3. La présentation des relatifs composés

S'il est un secteur où les choix de présentation effectués dans les grammaires et les manuels sont peu convaincants, c'est bien celui de l'ordre de succession des relatifs. Pour les formes simples, c'est l'ordre latin qui, depuis des siècles, est repris... Ce qui ne peut manquer de surprendre. En ce qui concerne les formes composées, il est plus difficile de déterminer ce qui guide l'organisation.

Ainsi, Abry et Chalaron (2000) dans la *Grammaire des premiers temps* proposent un tableau qui recense les associations entre des prépositions (fournies dans la première colonne) et les formes en *lequel*. L'organisation de ce tableau (Tableau 1) semble obéir à deux logiques distinctes :

Prépositions	<i>lequel, laquelle, lesquels, lesquelles</i>
après	Il trouvera peut-être un jour le bonheur après lequel il court depuis toujours
avec	
chez	
contre	
[...]	
selon	
sur	
à	
grâce à	
face à	
au cours de	
au-delà de	
au milieu de	
auprès de	

Tableau 1 :
Extrait de la présentation des relatifs composés (*Grammaire des premiers temps*, 2000)

Dans une première partie de la liste (de *après* jusqu'à *sur*), c'est l'ordre alphabétique qui est adopté. Ce choix, s'il peut répondre à un souci pratique – la facilité pour retrouver une forme – reste néanmoins atypique (car les faits de grammaire répondent rarement à une logique alphabétique) et n'est peut-être pas le plus indiqué quand la liste comporte peu d'unités. Ce mode de présentation fournit en tout cas peu d'indications sur l'emploi grammatical (les propriétés, les conditions d'emploi, la fréquence des assemblages, etc.).

La deuxième partie de la liste – qui n'est pas formellement séparée de la première – vient brouiller l'apparente logique d'organisation du début puisque la

Comment les données de corpus pourraient renouveler les manuels de grammaire ?

succession repart de *à* et contient des locutions prépositives qui ne respectent pas l'ordre alphabétique.

L'observation de deux corpus diversifiés² – afin de ne pas proposer un seul usage de la langue aux apprenants – fournit les résultats globaux présentés dans le Tableau 2. Ces résultats mettent à mal le tableau précédent sur plusieurs points et permettent de dégager un certain nombre d'observations, qui confortent et parfois complètent les observations de Bilger (2008) :

	ECRIT				TOT	ORAL				TOT	Rang
	<i>lequel</i>	<i>laquelle</i>	<i>lesquels</i>	<i>lesquelles</i>		<i>lequel</i>	<i>laquelle</i>	<i>lesquels</i>	<i>lesquelles</i>		
dans	42	42	20	9	113	42	20	7	5	74	2
sur	47	26	15	7	95	18	0	4	11	33	4
avec	26	37	7	3	73	9	3	6	2	20	5
pour	10	8	5	7	30	5	11	7	10	33	7
contre	2	3	3	3	11	1	1	0	0	2	13
devant	5	8	2	1	16	1	1	1	1	4	8
par	25	20	14	5	64	1	1	0	1	3	6
pendant	2	6	3	4	15	1	1	0	0	2	9
sous	7	4	1	0	12	1	1	0	0	2	12
à	<i>106</i>	85	<i>38</i>	<i>32</i>	261	<i>20</i>	19	<i>9</i>	<i>6</i>	54	1
(...) de	<i>45</i>	57	<i>24</i>	<i>16</i>	142	<i>3</i>	6	<i>1</i>	<i>0</i>	10	3
selon	0	0	0	0	0	0	1	0	0	1	20
vers	4	7	1	0	12	0	1	0	0	1	13
parmi	0	0	6	3	9	0	0	6	1	7	10
chez	6	3	3	3	15	0	0	1	0	1	10
derrière	4	2	2	3	11	0	0	0	0	0	15
entre	0	0	2	1	3	0	0	0	0	0	19
sans	0	5	1	0	6	0	0	0	0	0	17
après	3	1	0	0	4	0	0	0	0	0	18
par-dessus	1	0	0	0	1	0	0	0	0	0	19
à travers	1	5	3	0	9	0	0	0	0	0	16
	902					247					

Tableau 2 : Distribution des prépositions dans deux corpus de 1 million de mots

Le phénomène observé relève bien plus de la langue écrite que de l'oral puisque les exemples recueillis sont trois fois plus nombreux dans le corpus écrit. D'autre part,

² Chaque corpus comprend 1 million de mots. Le corpus oral est constitué de transcriptions de thématiques différentes établies à Clermont-Ferrand, Nancy, Grenoble et Poitiers, cela afin de disposer d'un échantillon de l'usage courant. Le corpus écrit est constitué de textes littéraires 19^e siècle pour l'essentiel, dans le souci d'observer une langue proche de l'usage décrit par les grammaires.

la diversité des prépositions y est aussi plus importante (les six dernières prépositions du tableau n'ont pas été trouvées dans le corpus oral)³.

Le choix des prépositions simples retenues dans le tableau du manuel surprend quelque peu puisque certaines des formes présentées (*durant, par-dessous*) n'ont pas été observées dans nos productions ; inversement, des prépositions qui sont attestées en nombre (*devant, pendant, etc.* – voir les exemples 1 et 2) ne sont pas fournies dans le document proposé aux apprenants.

1. Après environ une demi-heure, **pendant laquelle** mademoiselle de Verneuil et d'Orgemont se regardèrent [...] (Écrit, Balzac)
2. en face de cette situation **devant laquelle** euh il se trouvait tout à coup de devoir répondre à cette question (Oral, P98)

L'arbitraire du choix, c'est du moins ainsi que l'on interprétera ces faits, ne peut que faire ressortir l'intérêt des corpus qui « encadrent » ou « orientent » les données qu'il est utile et pertinent de proposer à des apprenants.

Si l'on s'en tient aux seules données orales, une hiérarchie se dégage très fortement et réduit drastiquement les formes qui méritent d'être présentées. Seules quatre prépositions semblent réellement productives : *dans, sur, pour, avec*. On peut aussi ajouter *à* (et dans une bien moindre mesure *de*) qui intervient dans la fabrication de formes synthétiques (*auquel, auxquelles...*).

L'appui sur des données de corpus déplace ainsi le travail grammatical qui de l'exploitation d'une mécanique quelque peu artificielle de la combinaison avec une préposition (piochée dans une liste très vaste) s'oriente vers la maîtrise d'une construction en relation avec des usages attestés.

4. L'ordre des pronoms compléments

Les règles d'apparition des pronoms compléments quand ils sont cumulés comme dans (3)-(4) font partie des faits de langue qui trouvent place dans la plupart des grammaires.

3. Pierre **le lui** donne
4. Marie **lui en** a parlé

Les explications proposées dans les ouvrages (grammaires et manuels) relèvent de deux types distincts. Les unes, plus anciennes, prennent appui sur les fonctions que remplissent ces pronoms. Elles donnent lieu à des schémas souvent complexes qui

³ Bilger (2008) identifie à partir d'autres corpus écrits (un corpus de presse notamment) certaines collocations non observées dans le corpus littéraire. On peut citer en particulier : *au cours de laquelle/duquel* ou *selon lequel / laquelle* et les assemblages tels que *condition(s) dans laquelle / lesquelles* et *raison(s) pour laquelle / lesquelles* qui peuvent être vus comme des clichés de l'écriture journalistique. Inversement la préposition *avec* était peu observée.

peuvent même apparaître comme confus. On en trouve par exemple dans Bonnard (1981) ou dans Arrivé et al. (1986), d'où est extrait le document ci-après (Figure 1) :

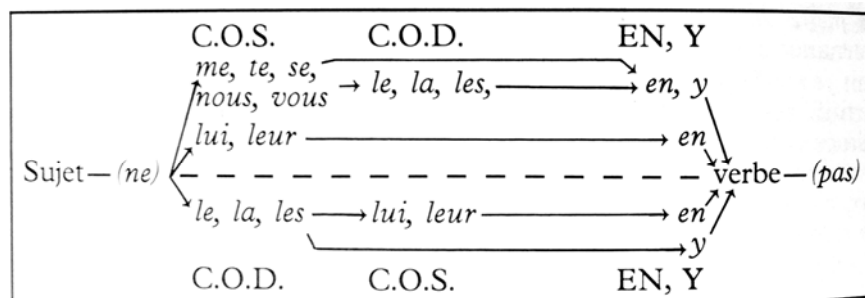


Figure 1 : L'ordre des pronoms compléments dans Arrivé, Gadet & Galmiche (1986)

Les secondes s'appuient sur les séries des formes elles-mêmes (c'est-à-dire les paradigmes de type *le, la, les* vs *lui, leur*). Cette orientation s'est progressivement substituée à l'approche précédente. On la trouve par exemple dans McBride (1997) et dans divers manuels comme *Grammaire pratique du Français* (2000). Voici le résumé partiel de la combinatoire que propose ce dernier ouvrage (Figure 2) :

Quand il y a deux pronoms compléments

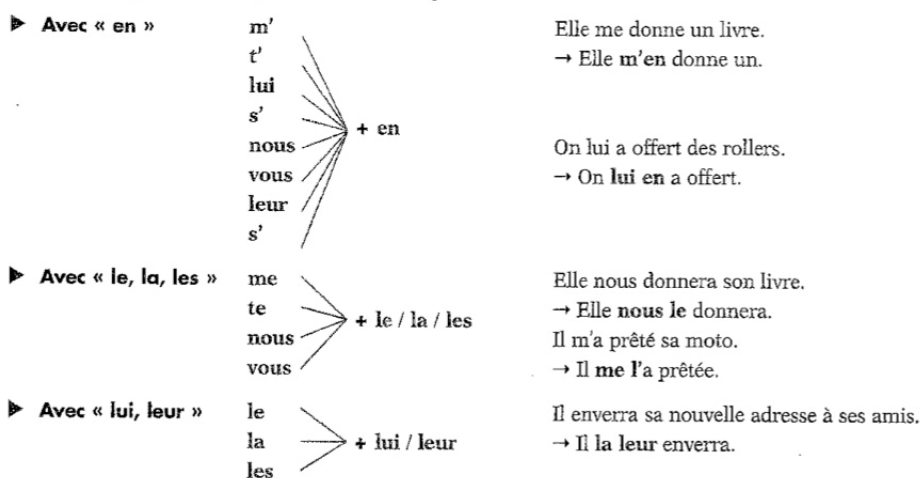


Figure 2. L'ordre des pronoms compléments dans *Grammaire pratique du français*.

D'un point de vue grammatical, l'exhaustivité qu'apporte cette présentation peut sembler, à première vue, satisfaisante et parfaitement adaptée aux besoins des apprenants. Toutes les suites potentielles sont indiquées et mises à disposition du lecteur.

Une autre vision du problème de l'ordre des pronoms compléments peut être fournie par un ouvrage comme celui de Blanche-Benveniste (2010) qui élabore sa description à partir des distributions observées dans des corpus. Plusieurs points en ressortent et conduisent à adopter un regard plus critique sur la présentation précédente. Premièrement, il existe un fort déséquilibre entre les diverses séquences virtuelles de deux clitiques. Seules quelques-unes sont réellement exploitées par les locuteurs : *y + en*, *se + en*, *me + en*. D'autres suites, souvent mises en avant dans les grammaires ou les manuels comme *le + lui* ou *le + leur*, sont quasi exceptionnelles. Ainsi les constructions avec un réfléchi sont privilégiées par les usagers mais pas nécessairement par l'enseignement.

Deuxièmement, les données de corpus sont éclairantes sur la vision quelque peu déformée propagée par les grammaires. Le Tableau 3 permet de vérifier que le nombre d'occurrences des suites *le lui* et *s'en* est très déséquilibré (en faveur de *s'en*, ce que les grammaires ne signalent pas) et que ces suites sont massivement utilisées à l'écrit mais peu à l'oral (en particulier pour *le lui*) :

	<i>le lui</i>	<i>s'en</i>
Écrit (Litt 19 ^e) ⁴	57	525
Oral (divers)	7	115

Tableau 3 : Nombre d'occurrences des suites de clitiques

En outre, pour chaque suite, un verbe tient une place prépondérante. Ainsi à l'écrit, la suite *le lui* se rencontre dans 28% des cas avec le verbe *dire* (5) et *s'en* dans 35 % des cas avec le verbe *aller* (6) :

5. Je lui ai pardonné, vous pouvez **le lui** dire. (Écrit, Zola)
6. Et elle **s'en** alla, précédée du vendeur [...] (Écrit, Zola)

Troisièmement, Blanche-Benveniste (2010) a proposé à divers locuteurs des suites de clitiques. Il en ressort que certaines suites (comme *le lui* en particulier) sont acceptées plus facilement avec quelques verbes particuliers comme *donner*, *prêter* et *faire* et que les locuteurs sont plus réticents face à ces assemblages avec d'autres verbes comme *envoyer* ou *emprunter*. Cela montre qu'une présentation mécanique de ce fait grammatical (qui ne tient pas compte des séquences de clitiques et des verbes impliqués) ne peut convenir.

Ces rapides observations aboutissent à s'interroger sur la frontière entre lexique et syntaxe que les corpus incitent à reconsidérer (Sinclair, 1991). Un point, comme la suite de deux clitiques compléments, habituellement traité en grammaire (d'où le souci de l'exhaustivité dans les combinaisons présentées) semble plutôt relever d'un fait de lexique (ce sont quelques combinaisons clitiques + verbes qui méritent d'être apprises et retenues). La présentation dans les manuels pourrait tirer profit de ces informations

⁴ Il s'agit des corpus présentés plus haut.

pour ne présenter qu'un nombre très limité de combinaisons et renvoyer à des annexes les tableaux des combinaisons potentielles si peu attestées.

5. La description de *même*

Les grammaires de FLM fournissent assez peu d'indications unifiées sur de nombreux morphèmes, ce qui est effectivement le cas pour la forme *même*. Le fait que cette dernière appartienne à plusieurs catégories et possède divers sémantismes favorise sans aucun doute l'éparpillement de l'analyse.

Ainsi, l'index de la nouvelle édition de la *Grammaire méthodique du français* (Riegel et al., 2009) renvoie le lecteur à de multiples paragraphes et le conduit à consulter 10 pages différentes pour connaître les divers emplois de la forme *même*. C'est donc une image éclatée de cet élément qui semble s'imposer.

En ce qui concerne les ouvrages de FLE, la forme *même* est rarement mentionnée – si ce n'est dans *La nouvelle grammaire du français* (Delatour et al., 2004) – et quand elle l'est, c'est dans des emplois qui se révèlent marginaux⁵ (cf. par exemple, *quand bien même* dans *La grammaire des premiers temps* (Abry et al., 1999)), alors que, comme nous le verrons, cette forme est bien représentée dans les productions, aussi bien orales qu'écrites. Elle occupe d'ailleurs la 53^e place dans la liste des 200 mots les plus utilisés en français parlé (cf. l'inventaire fourni par Blanche-Benveniste, 2010), d'où l'intérêt de s'appuyer sur des corpus pour en mieux cerner les emplois.

Si les descriptions disponibles s'appuient sur des catégories bien installées (adverbe, pronom et adjectif) et bifurquent rapidement vers des commentaires sémantiques⁶, les données de corpus vont nous permettre de fournir de nouvelles pistes d'analyse et de proposer une description à la fois plus homogène et plus détaillée.

5.1. Rapide présentation du cadre d'analyse

5.1.1. Les corpus utilisés

Quatre corpus de taille identique (200 000 mots chacun) mais de nature différente ont été utilisés pour mener à bien cette étude :

- le corpus oral Parole Privée (oral Privé) est constitué d'une partie du Corpus de Référence de Français Parlé (DELIC, 2004) ;
- le corpus oral Parole Publique/Professionnelle (oral Pro/oral Public) est également constitué d'une partie du CRFP à laquelle s'ajoutent des corpus⁷ de paroles professionnelles (avocats, hommes politiques, etc.) ;

⁵ En tout cas du point de vue quantitatif (voir plus loin).

⁶ Le TLF, par exemple, propose un classement qui entrecroise deux types de classement : une première partie repose sur la valeur sémantique alors qu'une deuxième se fonde sur la position.

⁷ Ces corpus complémentaires ont été édités par Paul Cappeau.

- le corpus *Littérature* (Litt) comprend des œuvres romanesques classiques du XIX^e siècle (entre autres, Zola, Balzac, Hugo, Flaubert) ;
- le corpus de *Presse* (Presse) est constitué d'articles de certains quotidiens et hebdomadaires (*Le Monde*, *Libération* ou *Le Nouvel Observateur*).

Les résultats bruts obtenus figurent dans le Tableau 4 :

	ORAL		ECRIT	
	Privé	Public / Pro	Littéraire	Presse
Nombre occurrences	802	885	440	393

Tableau 4 : Les occurrences de *même(s)*

Dans les corpus oraux, la forme *même* est deux fois plus employée ; nous verrons par la suite que ce phénomène est étroitement lié à la forte proportion de la locution *quand même*.

5.1.2. Paramètres d'analyse

Pour rendre compte du fonctionnement de *même*, nous avons eu recours à divers paramètres :

- la variation morphologique, la distribution (c'est-à-dire la position de *même* et la nature du groupe dans lequel il prend place) et la sémantique ;
- le travail sur corpus nous a également permis de dégager un certain nombre de collocations et d'observer leur inégale répartition selon les corpus utilisés.

La morphologie constitue le premier critère qui permet de subdiviser en deux l'ensemble des formes retenues, selon la répartition suivante (Tableau 5) :

<i>même</i> forme variable	<i>même</i> forme invariable
- vous aurez les mêmes produits (oral Pub) - on voit toujours les mêmes (Litt)	- des gens dangereux ou même méprisables (Litt) - ils ont même une très grande marque de bière (oral Pri)

Tableau 5 : Classement des deux formes de *même*

La colonne de droite correspond à ce que l'on appelle traditionnellement l'emploi *adverbial*, qui sera détaillé dans le 5.3.

La colonne de gauche regroupe à la fois ce que l'on désigne comme l'emploi *adjectival* et *pronominal*. La distribution dans les corpus est alors la suivante (Tableau 6) :

	ORAL		ECRIT	
	Privé	Public / Pro	Littéraire	Presse
<i>même</i> variable	174 (21,6%)	321 (36%)	249 (56,60%)	188 (48%)
<i>même</i> invariable	628 (78,4%)	564 (64%)	191 (43,40%)	206 (52%)

Tableau 6 : Répartition des deux formes dans les différents corpus

Ce simple relevé esquisse déjà une certaine singularité du corpus littéraire, puisque c'est le seul qui présente la forme variable comme majoritaire. En ce sens, il n'est pas surprenant que les dictionnaires et en particulier le TLF qui s'appuie sur *Frantext* (composé à 80 % de textes littéraires) commencent régulièrement leur article sur *même* par la description de l'adjectif.

Notre analyse portera d'abord sur la forme variable (plus rapide à décrire) puis sur la forme invariable qui oblige à regarder diverses locutions.

5.2. Même : forme variable

Dans ce cas, il est nécessaire de tenir compte de la position qu'occupe *même*.

5.2.1. Même postposé

Postposée, cette forme variable se rencontre principalement dans 2 contextes : soit après un pronom personnel disjoint, soit après un nom. La réalisation de *même* après un pronom tonique s'effectue différemment selon les personnes et les types de corpus. L'écrit littéraire présente le total le plus élevé (3 fois plus d'occurrences que dans l'oral privé) par l'emploi massif de *lui/elle-même*, alors que dans l'oral privé, sans surprise, la forme *moi-même* se rencontre fréquemment :

7. Fontanet **lui-même** en fut étonné (écrit Litt)

8. puis j'ai ensuite **moi-même** fait l'école hôtelière à Toulouse (oral Pro)

La forme *même* peut être réalisée après le nom tête du groupe nominal :

9. il faut que ça fasse effet **le jour même** (oral Privé)

10. c'était dès **cette nuit même** qu'il fallait ... (écrit Litt)

11. **les processus mêmes** du développement se sont ralentis (Presse)

Ces emplois postposés peuvent être rapprochés des précédents et, comme nous pouvons le voir (Tableau 7), ils sont nettement plus utilisés dans l'écrit littéraire.

	ORAL		ECRIT	
	Privé	Public / Pro	Littéraire	Presse
Nom + même	13	11	56 (22%)	11

Tableau 7 : Les emplois de *même* postposé à un nom

Quand il précède le nom (voir 5.2.2), *même* marque l'identité ou mieux la similitude. Quand il est placé après il est censé indiquer que le nom possède une qualité par excellence, mais les noms attestés dans les corpus ne rendent pas cette notion de qualité très explicite⁸ :

⁸ Vaugelas semble analyser *même* comme adverbe dans cette configuration. Il propose d'ailleurs de désaccorder le nom et l'adverbe *même* pour souligner sa valeur : *les choses même que je vous ai dites* vs *la chose mêmes que je vous ai dite*. La surprise vient de Brunot et Bruneau (1956) qui semblent valider cette analyse. Damourette et Pichon consacrent un paragraphe à cette question (§ 2894) « les auteurs se sont souvent comportés d'une façon assez capricieuse quant à la variation dans ce cas ». Pour eux, les

12. il faut que ça fasse effet **le même jour**
 13. il faut que ça fasse effet **le jour même** (oral Privé)

5.2.2. Même antéposé

L'emploi, de loin le plus massif, excepté dans le corpus Littéraire, est celui dans lequel *même* est placé avant le nom, le plus souvent entre le déterminant et le nom :

14. as-tu l'impression d'avoir **les mêmes relations** avec tous les clients (oral Pro)
 15. vous savez quand on est deux pour défendre **un même client** (oral Pro)

On observe une très forte prédominance de l'article défini dans ce contexte (cf. Tableau 8) et des collocations différentes selon les productions. Dans les productions orales, un tiers des occurrences de la suite *la même* + N se réalisent sous la forme *la même chose*. Dans l'écrit de la Presse, pour la séquence *du même* + N, on trouve *du même coup* dans 40% des emplois et les séquences *au même* + N utilisent pour 60% des cas, deux lexèmes particuliers : *moment* et *titre* (*au même moment* / *au même titre*) :

	ORAL		ECRIT	
	Privé (103)	Public / Pro (185)	Littéraire (67)	Presse (106)
<i>article défini</i>	98	171	56	96
<i>article indéfini</i>	2	10	9	9
<i>démonstratif</i>	3	4	2	1

Tableau 8 : Les déterminants utilisés devant *même* + Nom

Ce que l'on désigne traditionnellement comme pronom indéfini (*le même* et ses variantes) peut être traité comme un cas particulier de la structure plus générale *le même* [+ N] :

16. dans une cave la famille de fromages est **la même** (oral Pro)
 17. parce que l'objet n'est plus **le même** (oral Privé)

Le Tableau 9 fournit la faible distribution de cet emploi dans les corpus :

	ORAL		ECRIT	
	Privé	Public / Pro	Littéraire	Presse
<i>même</i> pronom =	12	33	5	9
<i>la même</i>	5	7	1	2
<i>le même</i>	4	10	1	3
<i>les mêmes</i>	3	16	3	4

Tableau 9 : Les emplois de *même* « pronom »

deux emplois (adverbe / adjectif) sont possibles dans ce contexte (*les consuls même viendront* et *les consuls mêmes viendront*) un critère permet la distinction : le caractère déplaçable de *même* lorsqu'il est adverbe.

Cette rubrique « pronom » apparaît toutefois avoir un intérêt réduit dans la description et son rapprochement avec l'« adjectif » semble nettement plus pertinent⁹. De fait, il semble bien que ce statut de forme pronominale pose problème et mérite qu'on s'y attarde un peu. Parmi les particularités de cette forme, on peut rappeler que : premièrement, *même* doit être accompagné nécessairement d'un article défini :

18a. les enduits sont toujours **les mêmes** (oral Pro)

18b. * les enduits sont toujours mêmes / des mêmes

Deuxièmement, l'interprétation de ce groupe nécessite un calcul anaphorique qui permet de récupérer un antécédent. Flaux (2008) qui classe *le même* dans les pronoms précise que « son fonctionnement anaphorique est dû au prédicat bivalent *même* qui suppose nécessairement la mention préalable d'un N désignant un référent ». Contrairement à d'autres formes comme *certains* ou *beaucoup* (19), aucun emploi nominal avec le trait /+ humain/ n'est attesté en l'absence du référent :

19a. on craint des risques de plans sociaux massifs et **certains** parlent d'un troisième tour social (O-HP3¹⁰)

19b. je crois que **beaucoup** disaient que et m'ont fait savoir et pas seulement à gauche m'ont fait savoir que s'ils avaient pu signer (O-HP3)

En ce sens, un exemple comme (20) ne peut s'interpréter que si l'antécédent humain a été donné antérieurement :

20. je crois que **les mêmes** disaient qu'ils n'auraient pas signé

Troisièmement, l'ensemble *le même* et variantes rappelle le fonctionnement qu'on trouve dans des séquences comme *le rouge*, *le dernier* (21), pour lesquelles on peut parler de syntagmes nominaux sans lexème nominal (cf. Blanche-Benveniste, 1990, 2010) :

21a. Cette nuit-là, Simon rêva qu'on pouvait reconnaître les couleurs par le toucher. **Le rouge** était proche de la liquidité, une consistance un peu sirupeuse. (*Frantext*, Beck, p. 25)

21b. J'ai viré les petits gars d'une pichenette et **le dernier** a poussé un hurlement terrible en basculant dans le vide, comme si je lui avais arraché un bras. (*Frantext*, Djian, p. 137)

Dans les séquences « le/la/les même(s) », le lexème n'est pas réalisé et *même* est antéposé à ce lexique non réalisé :

21c. il a acheté une voiture

j'ai acheté la même voiture

j'ai acheté la même 0

⁹ Il est d'ailleurs intéressant de noter que dans les grammaires (cf. entre autres Grevisse, 1980), dans le chapitre concernant les pronoms indéfinis, si *même* est régulièrement signalé, il ne donne pas pour autant lieu à une entrée particulière comme c'est le cas pour *certains*, *quelqu'un*, etc., et la plupart des grammaires FLE ne le signalent pas dans la liste des pronoms indéfinis. De même, certains dictionnaires (TLF) présentent *même* essentiellement comme étant un adjectif ou un adverbe, tout en signalant que les séquences « le/la/les même(s) » sont des groupes pronominaux.

¹⁰ O-HP3 est un corpus d'interviews d'hommes politiques.

Dans cet emploi où la forme *même* est antéposée au lexème (réalisé ou pas), celle-ci marque l'identité ou la similitude selon les contextes :

22. ils avaient droit à une fouille à corps tous les jours fouille à corps ensemble dans une **même cellule** (oral Pro)
23. est-ce que pour les **mêmes produits** vous aurez les mêmes produits de vente dans vos différents établissements (oral Pro)
24. mon vote sera **le même** (oral Pro)

On peut donc s'étonner de l'étiquette classique qui traite *le même* comme étant un pronom indéfini. En effet, comme on vient de le rappeler, le contexte antérieur ne peut être rempli que par l'article défini, le fonctionnement souligné par Flaux (2008) insiste sur le lien anaphorique obligatoire avec un SN exprimé, enfin le sens de *même* peut difficilement renvoyer à une valeur de « mal défini ».

Le Tableau 10 présente une synthèse concernant la distribution de l'emploi variable de *même* dans les divers corpus et les collocations observées :

<i>même</i> forme variable	
<i>même</i> postposé	<i>même</i> antéposé
<ul style="list-style-type: none"> • à un pronom <ul style="list-style-type: none"> - <i>lui</i> dans l'écrit littéraire (<i>lui-même</i>) - <i>moi</i> dans l'oral privé (<i>moi-même</i>) • à un nom <ul style="list-style-type: none"> - surtout dans l'écrit littéraire (<i>cette nuit même</i>) 	<ul style="list-style-type: none"> • à un nom <ul style="list-style-type: none"> - <i>chose</i> est le nom le plus utilisé à l'oral : <i>même chose</i> - fréquent dans la presse : <i>du même coup</i>, <i>au même moment</i>, <i>au même titre</i> • à un nom vide, non réalisé <ul style="list-style-type: none"> - emploi le plus fréquent dans l'oral professionnel (<i>ce sont les mêmes</i>)
-> marque le caractère en propre du terme antérieur	-> <i>même</i> possède une valeur assez affaiblie, il pourrait souvent être remplacé par un démonstratif : <i>au même moment</i> / <i>à ce moment</i> , <i>au même titre</i> / <i>à ce titre</i>

Tableau 10 : Synthèse des emplois de *même* variable

Le corpus Littéraire se caractérise par un emploi massif de la forme *même* variable et une répartition assez équilibrée entre 3 emplois : postposé au pronom et au nom et antéposé au nom. Les corpus oraux se caractérisent par un emploi minoritaire de la forme variable et par le fait que les deux tiers des emplois sont antéposés au nom. Le corpus Presse se situe entre les deux autres types de corpus : proche du Littéraire puisque la forme variable est bien représentée, mais aussi proche de l'oral par l'emploi massif de *même* antéposé au nom.

5.3. Même : forme invariable

Le simple décompte des occurrences de *même* dans des corpus de taille comparable (cf. le tableau 4) pourrait laisser penser que cette forme se rencontre bien plus fréquemment à l'oral qu'à l'écrit (environ deux fois plus). Les tendances globales

repérées dans les différents corpus sont en effet très convergentes et semblent neutraliser la distinction selon les types de production (privé, public, littéraire, presse) au profit de la seule différence entre oral et écrit.

Cependant, cette première impression est grandement faussée par le poids de la locution *quand même* qui représente de 30 à 45 % des emplois à l'oral contre moins de 2 % à l'écrit. Si l'on écarte cette locution de l'inventaire, les données chiffrées apparaissent bien plus équilibrées (Tableau 11) :

	ORAL		ECRIT	
	Privé	Public / Pro	Littéraire	Presse
<i>quand même</i>	361 (57%)	272 (48%)	7 (3,60%)	7 (3,40%)
autres emplois	267	292	184	199

Tableau 11 : Séparation de *quand même* et des autres emplois adverbiaux

5.3.1. La locution *quand même*

À l'oral, cette locution adverbiale *quand même* correspond à 53% des emplois invariables pour seulement 3,5% à l'écrit. Nous ne ferons pas une étude approfondie de cette locution mais nous indiquerons quelques faits que les corpus permettent d'observer.

Dans les productions étudiées, cette locution s'interprète essentiellement comme un marqueur « facilitant la mise en acceptabilité d'une contradiction », pour reprendre les termes de Moeschler et Spengler (1981), ou pour « atténuer une proposition qui peut paraître trop forte » (Beeching, 2005).

Dans cet emploi, *quand même* va pouvoir s'insérer dans des emplacements divers (Tableau 12) :

	Début de la CV	A l'intérieur de la CV	Fin de la CV	Autres contextes
Oral privé [255]	22	131 (51 %)	83	19
Oral pro [256]	12	160 (62 %)	44	40

Tableau 12 : Place occupée par *quand même* dans les corpus oraux

a) Début de la construction verbale :

25a. voilà + et puis bon peut-être **quand même** pour moi alors ça c'est une indication pour moi (oral Pro)

En fait, *quand même* est rarement en position initiale, mais suit fréquemment un autre introducteur (*mais, bon, voilà*, etc.), ce qui laisse penser que, dans l'usage, la valeur d'atténuation exprimée par cette locution vient en renfort d'un autre terme :

25b. mais + **quand même** la plupart du temps on se dit moi je fais ce que je veux comme je veux (oral Pro)

25c. voilà mais bon **quand même** il y a des inconvénients (oral Privé)

b) À l'intérieur de la construction verbale :

C'est la position majoritairement observée dans nos données :

26a. je pense **quand même** à la mort (oral Privé)

26b. je crois que j'ai **quand même** très souvent un a priori (oral Pro)

La locution se rencontre assez régulièrement entre l'auxiliaire et le participe et quelquefois aussi entre le sujet et le verbe :

27a. j'avais **quand même** fait un essai (oral Privé)

27b. sachant que le daim **quand même** blanchit (oral Privé)

c) Fin de la construction verbale :

28a. mais il y a eu quatre ans d'apprentissage **quand même** (oral Privé)

28b. allez + il fait pas froid + on est mieux en Martinique **quand même** hein (oral Pro)

d) Autres contextes :

29a. je suis dans un cadre **quand même** bien précis (oral Privé)

29b. puisque le propre **quand même** de l'image en bande dessinée c'est (oral Pro)

Cette locution se rencontre parfois lors de l'ajustement d'une formulation, *quand même* est réalisé immédiatement avant une reformulation :

30a. il y a **quand même** + il y a une espèce d'amalgame hein au niveau de cette jeunesse (oral Pro)

30b. mais + bon il y a **quand même** il y a un chemin qui s'est fait (oral Pro)

Pour conclure sur cette locution *quand même*, celle-ci est bien spécifique de l'oral, et comme le signale Beeching (2005, 2007), elle est utilisée par tous les locuteurs, quelle que soit la situation de parole et semble stable (tout autant produite en 1968 qu'aujourd'hui).

Dans les manuels, notamment de FLE, des équivalents moins marqués comme relevant de l'oral sont souvent proposés ; c'est le cas pour *tout de même*, et dans un registre plus littéraire, pour *néanmoins* et *cependant*.

À noter cependant que ces termes ont souvent des possibilités distributionnelles moins développées, par exemple : *cependant* est peu envisageable en fin de construction, etc. Le choix de ces substituts, en position initiale, est assez conforme à ce qu'indiquent nos corpus (Tableau 13).

	ORAL		ECRIT	
	Privé	Public / Pro	Littéraire	Presse
<i>cependant</i>	3	2	47	22
<i>néanmoins</i>	2	5	10	18
<i>tout de même</i>	4	8	5	9

Tableau 13 : Les alternatives à *quand même* dans les corpus

Il existe un emploi de cette locution *quand bien même* qui peut exprimer la concession adversative, comme dans l'exemple suivant produit à l'oral, par un locuteur professionnel (il s'agit en l'occurrence d'un avocat qui en fait le pivot d'une liste) :

- (31) c'est la stricte vérité elle est très simple extrêmement délicate **quand bien même** je serais avocat **quand bien même** je parlerais bien **quand bien même** j'aurais les moyens de me défendre parce que l'on va me dire + peut-être vrai peut-être faux (oral Public)

C'est toutefois le seul exemple de cette locution qui figure dans nos données. On ne peut alors qu'être dubitatif quant à la place importante qu'occupe cette tournure dans les grammaires et au décalage que cela illustre entre l'usage et les descriptions de la langue proposées.

5.3.2. Même adverbe modalisateur

Dans cet emploi adverbial, la forme *même* est traditionnellement analysée comme un outil d'enchérissement (ou de gradation), (cf. Bat-Zeev, 2001 ; Mehdaoui, 2000), qui va pouvoir porter sur des éléments grammaticaux très différents : infinitifs (32), participes passés (33), ou autres formes :

32. les machineurs devaient **même** doubler la vitesse (écrit Litt)
33. j'ai **même** été assez loin (oral Pro)

Voici quelques emplois intéressants à commenter :

a) Même suivi d'une préposition :

La forme invariable *même* peut se réaliser avant une préposition comme dans :

34. Tout est légèreté, **même dans** l'agression (écrit Presse)

À l'oral, la présence de *même* s'accompagne fréquemment d'une antéposition de la séquence prépositionnelle régie. On peut considérer que cette forme devient alors un marqueur de focalisation qui signale la valeur informative du constituant placé avant le verbe :

- 35a. **même pour** mes petits fils je dois faire ça (oral Privé)
35b. **même à** un Bordelais je le dis (oral Privé)
35c. dans le métier de fleuriste **même au bout de 10 ans** on n'a pas tout appris (oral Privé)

À l'écrit, ce procédé est bien moins habituel mais pas absent :

36. D'ailleurs, **même en** Australie, on se met à fabriquer des 'fromages traditionnels' (écrit Presse)

b) Même suivi de la particule négative pas (même pas) :

La séquence *même pas* à l'oral est fréquente (près de 12% des occurrences, 2 fois plus que dans les autres types de corpus) :

37. ça fait **même pas** deux ans (oral Privé)

Cette modalisation de la particule négative permet de créer ce que l'on peut appeler un noyau non verbal (cf. Blanche-Benveniste, 1990, 2010), comme dans les exemples suivants :

38a. **même pas** en rêve (Bachelot entendu à la TV)

38b. **même pas** peur

c) Même précédé de la particule négative pas (pas même) :

L'adverbe *même* peut être précédé de *pas*, comme dans les exemples suivants :

39. Personne, **pas même** Mme Grandet, n'avait la permission de venir. (écrit, Litt)

40. et rien ne pouvait l'en empêcher – **pas même** elle. (écrit, Litt)

Dans ce contexte, la forme *pas* est très proche de *ni*. Il est alors intéressant de noter que cet emploi semble être spécifique de l'écrit Littéraire, où il représente près de 11% des occurrences. En revanche, on n'en relève pas à l'oral et très peu dans la Presse. En effet, l'oral semblant plutôt privilégier l'ordre inverse, **même pas** plutôt que **pas même** :

41. je me rappelle de rien **même pas** où j'étais (oral Privé)

d) Même suivi d'une conjonction :

On rencontre des utilisations de *même* en renforcement de certaines conjonctions : *si*, *que* et *quand* (Grevisse, 1980) semble considérer que la liste peut être étendue, il cite (§ 1076) *même avant que*, suivi de points de suspension). Dans nos corpus, seules trois associations sont attestées (Tableau 14) :

	ORAL		ECRIT	
	Privé	Public / Pro	Littéraire	Presse
<i>même si</i>	33	39	2	48
<i>même que</i>	5	1	0	7
<i>même quand</i>	2	5	2	0

Tableau 14 : *Même* suivi d'une conjonction

On peut remarquer que les « conjonctions » renforcées par *même* n'introduisent pas toujours de réelles subordonnées (c'est-à-dire des constructions régies) mais parfois aussi des séquences associées, qui ne peuvent pas, par exemple, être extraites entre *c'est... que*.

e) même si :

Ces constructions semblent plutôt fonctionner comme des parenthèses (ce que la ponctuation souligne à l'écrit) :

42a. et et puis **même si** on accepte de les changer c'est très dur que le changement devienne pérenne (oral Pro)

42b. * c'est même si on accepte de changer que c'est très dur que le changement devienne pérenne

43. Et pourtant, d'aller au château des Carpathes, **même si** l'on rémunérerait son déplacement, cela ne pouvait lui convenir en aucune façon. (écrit Litt.)

L'oral mais aussi l'écrit de Presse favorise la locution *même si* (près de 24% des occurrences pour la Presse) qui est quasi absente de l'écrit Littéraire.

f) même quand :

On peut fournir des commentaires comparables pour *même quand* :

- 44a. eh bien écoutez je n'aliène pas ma liberté **même quand** on me dit la main sur le cœur que j'en garderais tous les avantages (oral HP)
- 44b. * c'est même quand on me dit que j'en garderais tous les avantages que je n'aliène pas ma liberté
45. Avec cet instinct, cette finesse de la femme qui a de l'esprit en toute chose, **même quand** elle console, Eugénie voulait tromper la douleur de son cousin en l'occupant de lui-même. (écrit Litt.)

On peut trouver, à l'oral, des suites susceptibles d'être ambiguës, comme dans (46) où *quand* peut s'interpréter comme une anticipation du second *quand* temporel ou comme la locution *quand même* suivie d'un *quand* temporel :

46. il faut réfléchir **quand même quand** on fait un bouquet (oral Privé)

il faut réfléchir		quand
	même	quand on fait un bouquet

il faut réfléchir	quand même	quand on fait un bouquet
-------------------	------------	--------------------------

g) même que :

Nous n'avons relevé aucun exemple où *même que* s'interprète comme formant un seul bloc, comme dans les exemples du type :

47. il a été grossier **même qu'**il m'a traité de ... (cf. Le Goffic, 1993)
48. **Même que** le Colonel Maçon vient d'être tué (cf. TLF1)

Cette locution est pourtant souvent présentée comme typique de l'oral ou du langage familier ou « populaire »¹¹. La séquence *même que* est certes attestée dans nos corpus mais elle s'analyse différemment en deux items séparés, comme dans (49) :

49. il faut absolument construire l'Europe + je pense **même que** les changements nécessaires en France ne peuvent s'inscrire que dans une perspective européenne (oral Public)

Dans ces exemple, *même* s'insère entre *je pense* et sa valence complément réalisée sous la forme d'une *que-P* : *que les changements...*

¹¹ Ball (2000) mentionne la forme *même* seulement dans ce type d'emploi, mais les exemples fournis (deux sur trois) sont issus d'un roman *La dérobade* de J. Cordelier (1976) et non de productions orales.

5.3.3. Récapitulatif des différents emplois de *même* invariable (Tableau 15)

<i>même</i> forme invariable		
<i>quand même</i>	<i>même</i> modalisateur	<i>même</i> + conjonction
utilisation très importante à l'oral, rare à l'écrit	devant une préposition : s'accompagne à l'oral d'une antéposition fréquente du Syntagme Prépositionnel (<i>même à un Bordelais je le dis</i>)	<i>même si</i> : absent de l'écrit littéraire, bien présent à l'oral et dans la presse
placé plutôt à l'intérieur de la construction	<i>pas même</i> : dans l'écrit littéraire	<i>même quand</i> : assez rare
<i>quand bien même</i> : souvent décrit, peu attesté	<i>même pas</i> : à l'oral	<i>même que</i> : absent, doit souvent être réanalysé

Tableau 15 : Les principaux emplois de *même* invariable

5.5. Bilan

Il nous a paru utile de détailler l'étude de cette forme *même*, tant elle illustre bien l'intérêt de travailler à partir de corpus diversifiés. Le résultat de ce travail n'est ni novateur ni complet, il permet toutefois de mieux apprécier les contextes les plus attestés. Cette analyse souligne parfois l'existence d'une discordance entre l'usage d'une forme et les développements qu'elle a suscités dans les grammaires ou les manuels, et elle a permis de mettre en lumière des associations parfois négligées.

Au terme de cette étude, on est tenté de suggérer que les différentes interprétations sémantiques (égalité/similitude, identité, modalité de renchérissement) et les différentes propriétés à la fois catégorielles et distributionnelles de cette forme devraient être hiérarchisées différemment dans les manuels. Des exemples comme (50)-(52) dans lesquels la forme *même* apparaît plusieurs fois signalent bien la hiérarchie que l'on devrait opérer. L'adverbe de modalité avec effet de renchérissement est de toute évidence l'emploi majoritaire dans les productions contemporaines :

- 50. j'ai été **moi-même** une fois **même** invité par l'évêque d'Amiens (oral Privé)
- 51. il y a un état d'esprit commun quoi **même si quand même** j'imagine qu'il y a différents personnages oral (Privé)
- 52. parce que j'ai eu ce décalage **quand même même** en étant petite (oral Privé)

5. Conclusion générale

Le principal objectif de cet article était d'illustrer l'intérêt de fonder les descriptions grammaticales sur des corpus de nature et de « genres » différents. Pour certains points de syntaxe (comme les pronoms relatifs ou la succession des pronoms personnels), les analyses semblent peu susceptibles de modification, néanmoins l'appui sur des corpus, variés et équilibrés, s'avère utile puisqu'il permet d'apporter un éclairage nouveau, en faisant, par exemple, ressortir des écarts de distribution importants selon les formes étudiées et en mettant en lumière des associations lexicales fortes. La palette des observations sera d'autant plus riche que les descriptions

s'appuieront effectivement sur différents types de corpus écrits ou oraux (Conrad & Biber, 2009), comme le montre ici l'analyse plus détaillée de *même*. Le fait que l'on puisse aujourd'hui travailler sur un plus grand nombre d'usages offre de multiples pistes, que les manuels pourraient explorer davantageusement ; cela leur permettrait de donner une représentation de la langue plus conforme à la réalité, du moins qui tiendrait compte un peu plus de son hétérogénéité.

Bibliographie

- ABRY, D. & CHALARON, M-L. (2000), *La grammaire des premiers temps* (Volumes 1 & 2), Grenoble, PUG.
- ARNAVIELLE, T. & G. SIOUFFI (2012), « Ecrire une grammaire du français aujourd'hui », *Langue française*, 176, p. 3-10.
- ARRIVÉ, M., GADET, F., GALMICHE, M. (1986), *La grammaire d'aujourd'hui*, Paris, Flammarion.
- BALL, R. (2000), *Colloquial French Grammar: A practical Guide*, Oxford, Blackwell.
- BAT-ZEEV SHYLDKROT, H. (2001), « Grammaticalisation et valeur sémantique : à propos d'un emploi particulier de *même* », C. Buridant, G. Kleiber & J.-C. Pellat (éds.), *Itinéraires linguistiques et grammaticaux*, Louvain/Paris, Peeters, p. 1-13.
- BAUDE, O. & DUGUA, C. (2011), « (Re)faire le corpus d'Orléans quarante ans après : quoi de neuf, linguiste ? », *Corpus*, 10, p. 99-118.
- BEECHING, K. (2005), « Politeness-induced semantic change: the case of *quand même* », *Language Variation and Change*, 17, p. 155-186.
- BEECHING, K. (2007), « La co-variation des marqueurs discursifs *bon, c'est-à-dire, enfin, hein, quand même, quoi et si vous voulez* : une question d'identité ? », *Langue française*, 154, p. 78-93.
- BÉRARD, E. & LAVENNE, C. (1991), *Modes d'emploi. Grammaire utile du français*, Paris, Hatier/Didier.
- BIBER, D., JOHANSSON, S., LEECH, G., CONRAD, S. & FINEGAN, E. (1999), *Longman Grammar of Spoken and Written English*, Londres/New York, Longman.
- BILGER, M. (1998), « Le statut micro et macro-syntaxique de ET », dans M. Bilger, F. Gadet & K. van den Eynde (éds), *Analyse linguistique et approches de l'oral. Hommage à Claire Blanche-Benveniste*, Louvain/Paris, Peeters, p. 91-102.
- BILGER, M. (2000), *Corpus : méthodologie et applications linguistiques*, Paris, Honoré Champion.
- BILGER, M. (2008), « De l'intérêt des corpus diversifiés pour les descriptions en (morpho)syntaxe. Réflexions et illustration avec le pronom relatif *lequel* », *Verbum*, 30/4, p. 275-286.
- BILGER, M. & CAPPEAU, P. (2004), « L'oral ou la multiplication des styles », *Langage et société*, 109, p. 13-30
- BLANCHE-BENVENISTE, C. (2010), *Le français. Usages de la langue parlée*. Louvain/Paris Peeters.

- BONNARD, H. (1981), *Code du français courant*, Paris, Magnard.
- BRANCA-ROSOFF, S, FLEURY, S, LEFEUVRE, F. & PIRES, M. (2011), « Constitution et exploitation d'un corpus de français parlé parisien », *Corpus*, 10, p. 81-98.
- CADET, L. & GUERIN, E. (éds.) (2012), « FLM, FLE, FLS au-delà des catégories », *Le français aujourd'hui*, 176.
- CHAMBERS, A. (2009), « Les corpus oraux en français langue étrangère : authenticité et pédagogie », *Mélanges CRAPEL*, 31, p. 16-33.
- CONRAD, S., & BIBER, D. (2009), *Real Grammar: A Corpus-Based Approach to English*, New York, Pearson.
- DELATOUR, Y., JENNEPIN, D., LÉON-DUFOUR, M. & TEYSSIER B. (2004), *Nouvelle grammaire du français*, Paris Hachette.
- DELIC (équipe) (2004), « Présentation du corpus de référence du français parlé », *Recherches sur le français parlé*, 18, p. 11-42.
- DE SALINS, G.-D. (1996), *Grammaire pour l'enseignement/apprentissage du FLE*, Paris, Didier/Hatier.
- DETEY, S., DURAND, J., LAKS, B. & LYCHE, C., (éds) (2010), *Les variétés du français parlé dans l'espace francophone. Ressources pour l'enseignement*, Paris, Ophrys.
- FLAUX, N. (2008), « Les pronoms indéfinis en français : une classe à (re)définir », *Travaux de linguistique*, 56, p. 7-46.
- GADET, F., LUDWIG, R. & PFÄNDER, S. (2009) : « Francophonie et typologie des situations », *Cahiers de linguistique*, 34/1, p. 143-162.
- GREVISSE, M. (1980), *Le bon usage*, Paris-Gembloux, Duculot.
- HABERT, B. (2000), « Des corpus représentatifs : de quoi, pour quoi, comment ? », dans M. Bilger (éd.), *Linguistique sur corpus : études et réflexions*, Perpignan, Presses Universitaires de Perpignan, p. 11-58.
- HALLIDAY, M. A. K. (1991), « Corpus studies and probabilistic grammar », dans K. Aijmer & B. Altenberg (éds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, Londres/New-York, Longman, p. 30-43.
- JOB, B., MIS, B. & PISSAVY, A-M. (1986), *Comment dire ? Grammaire simplifiée*, Paris, Clé International.
- LAFFON, M. & ZEGGAH-WUYTS, F. (2009), *La grammaire en action*, Paris, Clé International.
- LE GOFFIC, P. (1993), *Grammaire de la langue française*, Paris, Hachette Education.
- MCBRIDE, N. (1997), *Grammaire française*, Paris, Hachette.
- MARTIN, R., (1975), « Sur l'unité du mot même », *Travaux de linguistique et de littérature*, 13/1, p. 227-243.
- MCCARTHY, M. (2004), *Touchstone: From Corpus to Coursebook*, Cambridge, Cambridge University Press.
- MEHDAOUI, M. (2000), *L'adverbe même en français contemporain : concession et/ou enchérissement*, thèse de doctorat, Strasbourg.

Comment les données de corpus pourraient renouveler les manuels de grammaire ?

- MOESCHLER, J. & N. DE SPENGLER (1981), « *Quand même* : de la concession à la réfutation », *Cahiers de Linguistique française*, 2, p. 53-112.
- MILLER, J. & WEINERT, R. (1998), *Spontaneous Spoken Language. Syntax and Discourse*, Oxford, Oxford University Press.
- RASTIER, F. (2005), « Enjeux épistémologiques de la linguistique de corpus » dans G. Williams (éd.), *La linguistique de corpus*, Rennes, Presses Universitaires de Rennes, p. 31-45.
- RIEGEL, M., PELLAT, J.-C. & RIOUL, R. (2009), *Grammaire méthodique du français*, Paris, Presses Universitaires de France.
- SINCLAIR, J. (1991), *Corpus, Concordance, Collocation*, Oxford, Oxford University Press.
- TOGNINI-BONELLI, E. (2001), *Corpus Linguistics at Work*, Amsterdam, John Benjamins.
- VAN PETEGHEM, M. (1997), « Sur un indéfini marginal : *même* exprimant l'identité », *Langue française*, 116, p. 61-80.
- VÉRONIQUE, D., CARLO, C., GRANGET, C., KIM, J-O. (2009), *L'acquisition de la grammaire du français, langue étrangère*, Paris, Didier.