

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/388317984>

Le vocabulaire du français parlé Rapport de recherche

Technical Report · December 2025

CITATIONS

0

READS

46

3 authors, including:



Edward Arnold

Trinity College Dublin

2 PUBLICATIONS 16 CITATIONS

[SEE PROFILE](#)



Dominique Labbé

Institut d'études politiques de Grenoble

257 PUBLICATIONS 1,656 CITATIONS

[SEE PROFILE](#)

Arnold Edouard
French Department/Centre for European Studies
School of Languages, Literatures & Cultural Studies
Trinity College Dublin, the University of Dublin
ejarnold@tcd.ie

Labbé Cyril
Equipe Sigma
Laboratoire d'Informatique de Grenoble
Université Grenoble-Alpes
cyril.labbe@imag.fr

Labbé Dominique
Université Grenoble-Alpes
dominique.labbe@umrpacte.fr

Le vocabulaire du français parlé

Rapport de recherche

Version 01
Février 2025

Résumé

A l'aide d'un grand corpus de français parlé, le rapport décrit les principales caractéristiques lexicales de cette langue parlée. La comparaison avec l'écrit est faite avec un corpus de romans contemporains. Le français parlé est dominé par un petit nombre de verbes – surtout conjugués au présent de l'indicatif –, de pronoms personnels, d'adverbes et de conjonctions de subordination. Les noms et les adjectifs usuels sont assez peu nombreux et couvrent chacun un large champ sémantique ("chose" et "gens" en sont les archétypes), les prépositions et les conjonctions de coordination sont plus rares. Les principaux syntagmes verbaux et nominaux sont présentés. Au total, environ 2500 vocables usuels suffisent pour une conversation courante. Les résultats pourront être utiles aux enseignants de français langue étrangère et ils peuvent aussi combler en partie l'absence de corpus représentatifs des usages écrits et parlés de la langue française.

Abstract

Using a large corpus of spoken French, the report describes the main lexical characteristics of this language. The comparison with the written language is made using a corpus of contemporary novels. Spoken French is dominated by a small number of verbs – mainly conjugated in the present indicative –, personal pronouns, adverbs and subordinating conjunctions. Common nouns and adjectives are quite few in number and each covers a wide semantic field ("chose" and "gens" being the archetypes), prepositions and coordinating conjunctions are rarer. The main verbal and noun phrases are reviewed. In total, around 2500 common words are sufficient for an average conversation. The results may be useful to teachers of French as a foreign language and they can also partly fill the lack of corpora regarding the written and spoken uses of the French language.

Sommaire

Partie I. Corpus et méthodes

Chapitre 1 Corpus et normes de dépouillement

- 1.1 Deux corpus
- 1.2 Les normes de dépouillement

Chapitre 2 Statistiques

- 2.1 Effectifs, rangs, fréquences et vocables caractéristiques
- 2.2 Une collection d'événements rares et inégalement répartis

Chapitre 3. Les parties du discours

- 3.1 Poids des catégories grammaticales dans les deux genres
- 3.2 Calcul des fréquences par catégories grammaticales
- 3.3 Calcul des vocables caractéristiques

Partie II. Le groupe du verbe

Chapitre 4. Les verbes caractéristiques du français parlé

- 4.1 Recensement des verbes usuels
- 4.2 Les verbes caractéristiques de la conversation

Chapitre 5. L'auxiliaire

- 5.1 Les deux types de combinaisons des verbes
- 5.2 L'auxiliaire "avoir"
- 5.3 L'auxiliaire "être"
- 5.4 La singularité de "avoir été" dans le français parlé
- 5.5 Les principales combinaisons verbe + verbe infinitif

Chapitre 6. Les satellites du verbe

- 6.1 Les pronoms
- 6.2 Les adverbes
- 6.3 Les conjonctions de subordination

Partie III. Le groupe du nom

Chapitre 7. Les noms communs

- 7.1 Les noms communs du français parlé
- 7.2 Les principaux thèmes

7.3 L'intégration du nom dans le groupe verbal

Chapitre 8. Les adjectifs

8.1 Les adjectifs usuels du français parlé

8.2 Devant ou derrière le nom ?

Chapitre 9. Les mots outils dans le groupe du nom

9.1 Les déterminants

9.2 Les prépositions

9.3 Les prépositions dans les groupes nominaux complexes

9.4 Les conjonctions de coordination

Bibliographie

Annexes

1. Le corpus du français parlé

2. Le corpus du français écrit

3. Les 1420 noms les plus utilisés dans le français parlé

4. Les 753 adjectifs les plus utilisés dans le français parlé (classement alphabétique)

Table des tableaux

Tableau 1. Index hiérarchique des vocables les plus fréquents dans le français parlé comparé à l'écrit

Tableau 2. Densités comparées des catégories grammaticales (fréquence pour mille mots) et indice de spécificité.

Tableau 3. Densités comparées des deux principaux groupes de catégories grammaticales dans les français écrit et parlé (pour mille mots)

Tableau 4. Les temps et modes des verbes dans les français parlé et écrit (fréquence en pour mille verbes)

Tableau 5. Les principaux verbes utilisés dans les français écrit et parlé (fréquence pour mille mots)

Tableau 6. Densité des deux types de constructions "verbe + verbe" dans les français parlé et écrit (fréquence en pour mille mots)

Tableau 7. Les taux de couverture du texte par les combinaisons les plus fréquentes formées avec les auxiliaires avoir et être

Tableau 8. Les temps et modes de l'auxiliaire avoir devant un participe passé dans le français parlé

Tableau 9. Les verbes les plus utilisés au participe passé avec l'auxiliaire "avoir"

- Tableau 10. Les temps et modes de l'auxiliaire être devant un participe passé
- Tableau 11. Les principaux verbes au participe passé avec l'auxiliaire "être" dans le français parlé
- Tableau 12. Les syntagmes verbaux les plus fréquents dans les français parlé et écrit
- Tableau 13. Les auxiliaires modaux de la conversation comparés à ceux de l'écrit
- Tableau 14. Les pronoms personnels et impersonnels dans les français écrit et parlé (fréquences pour mille mots)
- Tableau 15. Les adverbes usuels en français parlé et écrit (fréquences pour mille mots)
- Tableau 16. Les conjonctions de subordination dans les français parlé et écrit
- Tableau 17. Les principaux noms du français parlé comparés à leur utilisation à l'écrit (fréquence pour mille mots)
- Tableau 18. Les principaux syntagmes verbaux intégrant un nom sans déterminant dans les français parlé et écrit (classement par effectifs décroissants)
- Tableau 19. Les principaux verbes pivots dans les syntagmes "verbe+nom"
- Tableau 20. Les principaux adjectifs du français parlé comparés à leur utilisation à l'écrit (fréquence pour mille mots)
- Tableau 21. Les cent syntagmes (avec adjectif antéposé au nom) les plus fréquents dans le français parlé avec leurs effectifs
- Tableau 22. Les adjectifs antéposés apparaissant au moins dans une combinaison répétée au moins cinq fois (classement hiérarchique sur les effectifs)
- Tableau 23. Les déterminants dans le français parlé comparé à l'écrit (classement hiérarchique, en pourcentage du total des occurrences des déterminants)
- Tableau 24. Les prépositions dans le français parlé comparé à l'écrit (fréquence en pour mille prépositions)
- Tableau 25. Les principaux syntagmes "nom+préposition+nom" dans le français parlé (classement par effectifs décroissant)
- Tableau 26. Les conjonctions de coordination dans le français parlé comparé à l'écrit (classement hiérarchique, fréquences en pour mille du total des occurrences des conjonctions de coordination)

Depuis une vingtaine d'années, l'étude du français parlé s'est développée - spécialement pour l'enseignement du français langue étrangère (FLE). Un certain nombre de colloques, de thèses, d'articles, d'ouvrages et de numéros de revues lui ont été consacrés. Pour les ouvrages, outre celui fondamental de Cuq & Gruca (2002), Weber (2013), Ravazzolo & Al. (2015), Tracy & Al. (2022). Tous ces ouvrages insistent sur l'importance de l'oral, ils souhaitent que les documents authentiques prennent le pas sur les exemples fabriqués (ou littéraires). Ils signalent également l'importance de la fréquence à propos de la communication et de l'enseignement. En cela, les chercheurs français en didactique des langues ne font que suivre, avec un certain retard, un engouement général pour la linguistique de corpus (Sinclair 1991). Pour le français, outre l'ouvrage pionnier de Habert & Al. (1997), on se reportera aux contributions plus récentes rassemblées dans Williams (2005), dans Frérot & Pecman (2021) et dans Tracy & Al. (2022), spécialement Paternostro (2022) qui établit la différence entre l'oral – qui peut ne pas être spontané - et la langue parlée.

De nombreuses langues sont étudiées sous leur forme parlée à l'aide de vastes corpus et d'interfaces de consultation plus ou moins ergonomiques. Les modèles demeurent le British National Corpus (Burnard 1995) et le corpus de l'américain parlé en ligne à l'université de Santa-Barbara (<https://www.linguistics.ucsb.edu/research/santa-barbara-corpus/>). A l'heure où ces lignes sont écrites, il n'existe rien de tel pour le français malgré des études pionnières comme celles de Blanche-Benveniste & Jeanjean (1986) et Gadet (1989). Ces dernières années, quelques ressources électroniques ont été mises en ligne mais, pour l'instant, les enseignants et les apprenants en FLE ne disposent pas de véritables corpus de grandes dimensions pour le français parlé. La plupart des tentatives sont dominées par l'aspect phonétique (par exemple Durand 2009) et font peu de place au lexique proprement dit. La tentative la plus récente et la plus aboutie est le projet Orfeo (<https://orfeo.ortolang.fr/>) dont Parisse & Al. (2021) présente la méthode et certaines particularités sans en dissimuler les limites. Pour les bases de données constituées dans un objectif FLE proprement dit, on peut citer le projet Clapi-FLE (Ravazzolo & Al. 2015) et la base FLORALE, mise en ligne par l'Université de Lausanne (Surcouf 2021). Comme le soulignent ces auteurs, les principales difficultés résident dans l'alignement de l'oral et de l'écrit, la maîtrise de l'outil informatique et l'élaboration d'une interface utilisateur aussi conviviale que possible afin de ne pas rebuter l'utilisateur qu'il soit apprenant, enseignant ou chercheur.

Sans prétendre combler ces manques ni élaborer une théorie de la langue, nous proposons dans cet article, une comparaison du français parlé et du français écrit à l'aide de deux corpus extraits d'une bibliothèque électronique qui sera présentée avant d'analyser les principales caractéristiques lexicales du français parlé.

Ces corpus n'ont pas été constitués pour l'étude de la langue mais pour des recherches sociologiques, politologiques ou littéraires. Grâce à leur volume important et à leur grande diversité, ces corpus peuvent donner une représentation approchée de la langue parlée, même si nous ne prétendons pas en faire des "échantillons représentatifs" du français. D'ailleurs, on signalera à plusieurs reprises dans ce rapport que les sujets de recherche pèsent plus ou moins lourd sur le vocabulaire de ces corpus.

De plus, cette recherche ne porte pas sur l'ensemble du français parlé mais sur le seul lexique à l'exclusion de la phonétique et de la syntaxe. Enfin, le corpus ne porte pas sur les régionalismes ou le français parlé dans le reste de la francophonie.

Nous espérons simplement fournir aux chercheurs et aux enseignants de FLE quelques données utiles. Ce document pourra peut-être aussi servir aux enseignants de classes spécialisées pour les élèves connaissant des difficultés d'apprentissage du français ou aux orthophonistes et autres professionnels de la langue.

Partie I.

Corpus et méthodes

A défaut d'échantillons représentatifs des deux genres, nous proposons d'utiliser des "corpus" - c'est-à-dire de vastes collections de textes, constituées pour d'autres usages et rassemblées dans la bibliothèque électronique du français moderne qui a été développée depuis plus de trente-cinq ans par C. et D. Labbé (<http://lexicometrie.univ-grenoble-alpes.fr/>). Cette bibliothèque comporte actuellement plus de 80 millions de mots étiquetés. Les principales sections concernent la littérature - 32 millions de mots dont 12,9 pour la littérature contemporaine - et la politique : 29 millions de mots (présentation dans Arnold & Al. 2016). Plusieurs sections ne sont pas en ligne dont celle portant sur le français parlé (annexe 1).

La présente étude se base sur une comparaison de ce français parlé avec la littérature contemporaine (annexe 2). Tous les textes ont été dépouillés à l'aide des outils d'indexation, de concordance et de statistiques qui seront présentés avant de donner les principales caractéristiques lexicales du français parlé. Cette uniformité de traitement ouvre la possibilité d'une comparaison systématique, comme si tous les mots étaient passés sous la même toise.

Chapitre 1

Corpus et normes de dépouillement

Pour établir des statistiques fiables sur les phénomènes langagiers, il faut rassembler un grand nombre d'observations dans des collections systématiques (corpus) qui doivent être constitués en suivant exactement les mêmes normes de dépouillement. Sinon, le chercheur risque de considérer comme des différences de fond des accidents provenant d'un manque de rigueur dans les observations ou dans l'établissement des données.

1.1 Deux corpus

Le corpus du français parlé comporte essentiellement des transcriptions d'entretiens sociologiques dont l'objectif premier n'était pas d'étudier la langue mais diverses questions de société ou de politique. Ces transcriptions ont été réalisées par des sténographes qui ont suivi les règles classiques en la matière. Par exemple, les allitérations – très courantes à l'oral - sont effacées : "v'là" ou "c'te" ont été transcrits "voilà" et "cette". Ce corpus ne permet donc pas d'étudier la phonétique du français parlé qui est une grosse part de la pédagogie du FLE. En contrepartie, cette normalisation rend possible une comparaison systématique avec l'écrit et, au-delà, il ouvre à l'étude du lexique de la langue parlée, c'est-à-dire le français "ordinaire", pour reprendre l'expression de F. Gadet, qui est l'objet essentiel du FLE. Autrement dit, cette recherche ne porte pas sur les sons du français *parlé* mais sur son lexique.

Certes, ce corpus n'est pas constitué d'échantillons aléatoires comme le British National Corpus de telle sorte que sa représentativité peut être discutée. Cependant, ces textes sont très divers et proviennent d'un grand nombre de locuteurs (plusieurs milliers si l'on compte les questions ouvertes des sondages). Cette diversité et cette étendue permettent de considérer qu'il n'y a pas de biais majeurs concernant les principales caractéristiques lexicales du français parlé.

Pour le corpus du français écrit qui sert d'étalon de comparaison, puisque les transcriptions de français parlé ont été effectuées durant les quarante dernières années, il est logique de limiter le corpus écrit à l'époque contemporaine.

Pour cette étude exploratoire, nous avons renoncé à inclure le discours politique. Ce corpus comporte une majorité de transcriptions de l'oral mais ce ne sont pas des propos spontanés. Même dans les entretiens radio-télévisés, qui se rapprochent le plus de l'oral, il y a une proportion considérable de "français soutenu" car les hommes politiques ont l'habitude de

s'exprimer en public dans une langue très éloignée de la vie quotidienne (Arnold 2019 ; Arnold & Labbé 2016).

L'expérience porte donc sur deux grands ensembles de textes :

- A Le français parlé : 433 transcriptions d'entretiens et des réponses à des questions ouvertes dans quatre sondages (annexe 1).
- B : Littérature contemporaine : 134 livres par 51 auteurs contemporains comportant 12,9 millions de mots. Ce corpus sert d'échantillon du français écrit (annexe 2).

Quelles sont les différences entre A et B ?

La réponse à cette question dépend de plusieurs conditions. Il faut d'abord que les deux ensembles soient comparables, c'est-à-dire que les textes aient été dépouillés exactement de la même manière.

1.2 Normes de dépouillement

L'originalité essentielle de ces corpus réside dans la normalisation des graphies, leur balisage et leur étiquetage en suivant les principes posés par Muller en 1963 et 1977. Les normes utilisées pour constituer la BEFM sont présentées dans Labbé (1990). Ces traitements permettent une comparaison sans biais entre les textes.

Ces opérations préalables sont de trois types.

Premièrement, dans les entretiens, on trouve des balises comme : <Q> (pour question) et <R> (pour réponse) afin d'isoler la parole de l'enquêté de celle de l'enquêteur qui l'interroge et pour que l'ordinateur ne confonde pas les deux. Dans la présente étude, seuls les propos des enquêtés ont été conservés (car les questions sont souvent préparées et sont proches de l'écrit alors que les réponses sont spontanées).

Deuxièmement, la graphie des mots est standardisée. En plus des élisions mentionnées plus haut, cette standardisation porte tout spécialement sur les mots à majuscule initiale (toponymes, patronymes, sigles...). En effet, la plupart des mots sont susceptibles d'avoir plusieurs graphies : majuscules ou minuscules, élisions, abréviations... Par exemple, dans certains entretiens, "monsieur" est écrit en toutes lettres, ailleurs M (suivi ou non d'un point), dans d'autres encore : Mr... Les élisions, les chiffres et les dates sont d'une infinie variété, parfois en lettres, parfois en chiffres, avec comme séparateur de millier : une virgule, un blanc ou rien du tout ; pour les décimales : un point ou une virgule, etc. Enfin, il faut compter avec les fautes d'orthographe qui émaillent ce genre de documents, même lorsqu'ils ont été saisis par des secrétaires

expérimentées, ce qui était le cas pour la plupart des entretiens. De plus, des milliers de mots français sont susceptibles d'avoir plusieurs orthographes comme "puis" et "peux".

Troisièmement, des mots différents peuvent s'écrire de la même manière. Par exemple, "suis" (suivre ou être ?), "est", "été", "être" (noms ou verbes ?). Si le dépouillement porte sur ces formes graphiques, il devient impossible d'étudier les verbes. En effet, dans tout texte français, ces "homographies" touchent plus du tiers des mots et ce sont des mots très fréquents, comme "être" qui est toujours le verbe le plus utilisé ou "est" qui est la flexion la plus fréquente du verbe "être"... Pour surmonter cette difficulté, chaque mot reçoit une étiquette (ou lemme) qui correspond à l'entrée du dictionnaire à laquelle appartient la forme graphique présente dans le texte, suivant les principes posés par C. Muller. Par exemple pour un verbe, l'étiquette comporte, outre la graphie standardisée, l'infinitif, le temps et le mode (quand il est possible de le déterminer). Ces étiquettes correspondent à la nomenclature des dictionnaires de langue, de telle sorte que l'utilisateur puisse accéder aux vocables en suivant les conventions usuelles et sans avoir à modifier ses habitudes.

Dans la suite de ce rapport, l'analyse portera donc sur les vocables et non sur les formes graphiques. Toutefois, celles-ci ne sont pas perdues. Le lecteur pourra les retrouver dans les index alphabétiques des deux corpus qui sont mis en ligne avec ce rapport.

Chapitre 2.

Statistiques

Les corpus utilisés pour cette recherche sont de grande taille de telle sorte que leur observation requiert des dépouillements informatiques – notamment l’indexation des mots classés de différentes manière - et des outils statistiques spécifiques.

2.1 Effectifs, rangs, fréquences et vocables caractéristiques

Dans la suite de ce rapport, pour comparer les deux corpus, trois outils statistiques seront privilégiés (rang, fréquence et vocabulaire caractéristiques). Ils permettront de mettre en lumière les différences considérables existant entre les deux registres de langue.

Rang et fréquence

Les index décomptent les effectifs de chacun des vocables de chaque corpus. Par exemple, le verbe le plus utilisé, à l’écrit comme dans la conversation, est "être". Il apparaît 321 493 fois dans le corpus écrit et 161 670 dans celui de la conversation. Comme les deux corpus ne sont pas de même longueur, ces effectifs (absolus) ne sont pas directement comparables.

Pour comparer deux vocabulaires, le rang est l’outil le plus simple. Il exprime l’ordre de préférence pour certains vocables mais aussi les contraintes de la langue. Ainsi dans tout texte en français, long au moins d’au moins quelques centaines de mots, l’article "le" arrive en tête, suivi de la préposition "de", ou encore les trois verbes les plus utilisés sont toujours "être", "avoir" et "faire" (dans cet ordre). Ces exemples permettent de comprendre que les préférences individuelles, tout comme les genres, doivent nécessairement se couler dans le système de la langue.

Le deuxième outil est la fréquence relative qui consiste à ramener les chiffres absolus au même dénominateur (en statistique lexicale : pour mille mots, noté ‰). Le corpus du français écrit est long de 12 769 494 mots et celui de la conversation de 3 634 737. Dans le français écrit, la fréquence du verbe "être" est :

$$F_{écrit\text{être}} = \frac{321\,493}{12\,769\,494} * 1000 = 25,2 \text{ ‰}$$

En moyenne, pour mille mots dans un texte français, on rencontre 25,2 fois le verbe "être".

Le même calcul sur le français de la conversation donne pour ce verbe une fréquence de 44,5%. On voit que la fréquence de ce verbe est de 77 % plus élevée dans la conversation que dans le français écrit contemporain.

Vocabulaire caractéristique

Pour savoir si les écarts constatés ne sont pas simplement dus au hasard, on utilise un test statistique issu de la loi hypergéométrique. Ce calcul a été présenté par Lafon 1980 et 1984. Il est discuté et adapté dans Labbé & Labbé 1994 (voir également Monière & Labbé 2018). Dans l'exemple donné ci-dessus, cet indice indique qu'il y a moins d'une chance sur mille de se tromper en rejetant l'hypothèse d'un phénomène unique simplement affecté par des fluctuations aléatoires et en affirmant que le verbe être est significativement plus employé dans la conversation par rapport à l'écrit.

Cet indice varie uniformément entre 0 et 1. Plus il est proche de zéro, plus la fréquence du mot est significativement faible dans le français parlé. Inversement, plus l'indice est proche de 1, plus le mot est significativement surutilisé dans le français parlé par rapport à l'écrit. Dans la suite de ce rapport, le seuil utilisé est un risque inférieur à une chance sur cent de se tromper en considérant comme vraie l'hypothèse que le mot considéré est une caractéristique du français parlé. En cas de suremploi, les tableaux comportent le signe + et – dans le cas inverse.

L'analyse des deux vocabulaires aboutit à un premier constat concernant le langage en général : tout message – d'une longueur minimale de quelques centaines de mots – est une collection d'événements rares dont les plus fréquents ne sont pas porteurs de sens, du moins lorsqu'ils sont isolés du contexte de leur émission.

2.2 Une collection d'événements rares et inégalement répartis

Les outils statistiques permettent d'abord d'examiner l'idée commune selon laquelle les vocables les plus utilisés donnent des indications concernant les préférences des locuteurs et, en réunissant une grande population de locuteurs suffisamment divers, cette collection offrira une image du vocabulaire usuel.

L'index hiérarchique permet de tester cette idée commune. Il classe les vocables selon leur fréquence. Le tableau 1 présente la tête de liste en suivant le classement du corpus parlé.

Tableau 1. Index hiérarchique des vocables les plus fréquents dans le français parlé comparé à l'écrit

Lemme et Catégorie grammaticale	A Parlé		B Ecrit		(A-B)/B %	Indice
	Rang	Fréquence (‰)	Rang	Fréquence (‰)		
le (dét)	1	70,7	1	89,2	- 21	-
de (pré)	2	48,4	2	64,5	- 25	-
être (v)	3	44,5	4	25,2	+ 77	+
avoir (v)	4	35,5	8	21,9	+ 62	+
je (pro)	5	33,2	6	23,9	+ 39	+
ce (pro)	6	27,1	15	9,6	+ 183	+
on (pro)	7	21,2	27	4,9	+ 336	+
pas (adv)	8	20,6	14	9,8	+ 111	+
à (pré)	9	19,5	5	25,2	- 22	-
que (cj)	10	19,3	11	12,9	+ 50	+
il (pro)	11	18,9	3	27,1	-30	-
un (det)	12	18,4	7	23,4	-21	-
ça (pro)	13	15,8	41	3,0	425	+
et (cj)	14	14,6	9	18,3	-20	-
ne (adv)	15	12,2	10	13,8	-11	-
qui (pro)	16	10,9	19	7,6	43	+
faire (v)	17	9,9	24	7,5	32	+
y (pro)	18	9,6	36	3,4	186	+
mais (cj)	19	8,6	28	4,6	86	+
ils (pro)	20	8,2	33	3,8	117	+

Lecture : le vocable le plus utilisé dans le français parlé comme à l'écrit (l'article "le") a une densité d'utilisation de 70,7 pour mille mots dans le français parlé et de 89,2 pour mille mots à l'écrit. Les deux dernières colonnes indiquent que la fréquence de "le" dans le français parlé est de 21 % inférieure à celle de l'écrit et que cet écart a moins d'une chance sur cent de se produire au hasard. Le sous-emploi du déterminant "le" est donc une caractéristique du français parlé.

Ces vingt vocables couvrent à eux seuls près de la moitié des propos émis à l'oral. En dehors des verbes "être", "avoir" et "faire", ce sont des mots outils, c'est-à-dire que leur signification dépend des autres mots employés avec eux mais aussi de la qualité du locuteur, du thème qu'il aborde et de la situation dans laquelle il parle. Faute de ces informations, la liste du tableau 1 semble de peu d'utilité. Cependant, elle amène plusieurs remarques.

Certes les deux listes comportent douze vocables en commun (*le, de, être, avoir, je, à, que, il, un, et, ne, qui*) mais, en dehors des deux premiers, le classement est différent. Par exemple, le pronom "il" - qui est le troisième vocable le plus employé à l'écrit - est absent des dix premiers dans le français parlé ; de même que l'article "un", qui est le 7^e vocable le plus utilisé à l'écrit, passe au douzième rang dans la conversation. En résumé, la comparaison des rangs indique que, par rapport à l'écrit, le français parlé semble privilégier :

- les verbes : *être, avoir* et *faire*

- les pronoms personnels : *je, on* et *ils* (qui est aussi le lemme de "elles") ; les pronoms démonstratifs : *ce* et *ça* ; les relatifs *qui* et *y*.

- l'adverbe "pas" (contrairement à "ne") ;
- les conjonctions : *que* et *mais*.

De plus, les écarts de densité entre les deux corpus sont considérables et aucun de ces écarts ne peut s'expliquer par le hasard. Ainsi l'utilisation de l'article "le" est d'un cinquième plus faible en moyenne dans le français parlé qu'à l'écrit ; pour la préposition "de", c'est même une baisse d'un quart. En revanche, le parlé utilise beaucoup plus les verbes "être" et "avoir" dont la majorité des occurrences sont des auxiliaires, ce qui justifie leur classement dans les mots outils, comme l'avait fait l'équipe du français fondamental (Gougenheim & Al. 1956). Les outils préférés du français parlé semblent être le pronom démonstratif "ce" (dont la fréquence triple presque par rapport au français écrit), le pronom personnel "on" (quadruplement) et l'adverbe "pas" (doublement).

Autre surprise, le tableau 1 ne contient que des mots outils, sauf *faire*) qui se trouve au 24^e rang à l'écrit mais au 17^e dans le français parlé. Dans la conversation, le substantif le plus utilisé est "chose" - en 64^e position avec une fréquence de 2,8 pour mille mots, soit trente fois moins que l'article "le". A l'écrit, le substantif le plus utilisé est "homme" au 66^e rang.

En fait, dans toute langue – que ce soit à l'oral spontané ou à l'écrit - les mots les plus fréquents ont un rôle syntaxique et n'acquièrent un contenu qu'avec le contexte (et la situation d'énonciation). C'est particulièrement vrai pour les pronoms : certes, à l'oral, le "je" est la personne qui parle – sauf énoncé rapporté ("vous avez dit "Je..."), tournure qui ne se trouve pas dans les entretiens. En revanche, à l'écrit, ce même pronom peut incarner une multitude de personnes, de telle sorte que sa seule définition possible est "celui qui parle à ce point précis du texte" (sur cette question de l'énonciation du sujet dans le langage : Benveniste (1959) et Kerbrat-Orrecchioni 1981).

On ajoutera que, dans le français parlé comme à l'écrit, les dix vocables les plus fréquents couvrent plus du tiers du texte à eux seuls. A l'écrit, l'ensemble des mots outils occupent en moyenne 57 % de ce texte ; en situation d'oral spontané cette proportion atteint 62 %. Autrement dit, pendant plus de la moitié du temps le lecteur d'un texte en français, ou l'auditeur, lisent ou écoutent des mots qui tirent leurs sens du contexte et de la situation d'énonciation, sens évidemment perdu dans les listes de mots (index) comme celle du tableau 1.

Dans la conversation, les mots les plus usuels reviennent donc plus souvent qu'à l'écrit, ce qui signifie nécessairement que le vocabulaire mobilisé est moins étendu. Le flux normal des échanges limite le choix des mots et rend les répétitions moins sensibles. En contrepartie, un corpus suffisamment étendu permet d'isoler le vocabulaire réellement indispensable pour échanger dans cette langue.

On en conclue également que le vocabulaire d'un corpus est une grande collection d'événements rares (les occurrences d'un vocable ou d'une catégorie grammaticale) très inégalement distribués, dont les plus fréquents ne sont pas les plus facilement interprétables. Ce phénomène d'inégale distribution des fréquences a été mis en lumière par Zipf 1935 (voir également Mandelbrot, 1957).

Enfin, le tableau 1 contient huit pronoms dont sept sont plus utilisés dans la conversation par rapport à l'écrit, de même pour les trois verbes les plus usuels. A l'inverse, les deux déterminants et les deux prépositions, présentes dans cette courte liste (le, un, de, à), sont nettement sous-employées dans le corpus du français parlé. Peut-on généraliser ce constat et affirmer que le français parlé marque une préférence pour les verbes ou les pronoms et une défiance envers les déterminants et les prépositions ? Il faut pour cela examiner les mouvements affectant les parties du discours dans ces deux corpus. Ce sera l'occasion de présenter cet outil essentiel dans l'analyse lexicométrique.

Chapitre 3

Les parties du discours

L'étiquetage des mots, décrit dans le premier chapitre de ce rapport, rattache chacun d'eux à sa catégorie grammaticale ("parties du discours"). Les index permettent donc de calculer les effectifs de chacune de ces catégories, de les convertir en fréquence et de leur appliquer le test statistique qui met en lumière les différences significatives entre les deux corpus.

3.1 Poids des catégories grammaticales dans les deux genres

Quand ils parlent, les Français utilisent-ils la même densité de verbes, de pronoms, de noms, qu'à l'écrit ? Le tableau 2 donne une réponse négative et permet de caractériser les différences entre les deux genres. Ces caractéristiques obligent à reconsidérer la comparaison entre ces deux corpus.

Sauf pour l'infinitif du verbe, tous les écarts sont considérables et il y a moins d'une chance sur cent de se tromper en affirmant qu'ils ne peuvent s'expliquer par le hasard.

Le français parlé manifeste une préférence significative pour les verbes, utilisés massivement au présent, qui est la plupart du temps un présent de narration puisque les entretiens portaient, en grande partie, sur des événements passés que les enquêteurs demandaient de raconter.

Autre constat : à l'écrit, outre le présent, l'imparfait et le passé simple sont les temps privilégiés ; les temps composés, avec un auxiliaire et un participe passé, sont moins importants. Dans le français parlé, le passé simple disparaît et la densité de l'imparfait recule de plus de moitié. Outre le présent de narration, chez la plupart des locuteurs, le passé se réduit au passé composé. Même le futur est significativement sous-employé. Nous verrons plus loin qu'il est en partie remplacé par "aller" (au présent) suivi d'un infinitif ("demain, je vais faire...").

Enfin, dernière remarque à propos des verbes en français parlé : la quasi-disparition du participe présent. Cressot (1963) a signalé que ce participe est la forme verbale la plus proche de l'adjectif (d'où l'expression "adjectif verbal") qui a pris depuis le XIX^e siècle un développement considérable. Selon cet auteur, il permet d'attribuer aux choses une vie et une dynamique tout en atténuant l'éclat des adjectifs en "-eur" auxquels ces participes présents se substituent souvent. Manifestement, ces nuances littéraires n'ont pas leur place à l'oral.

Les enseignants de FLE passent beaucoup de temps à faire apprendre les modes et les conjugaisons alors que, en dehors du présent de l'indicatif et secondairement du passé composé, les autres temps et modes sont peu ou plus guère utilisés dans la vie quotidienne. Mieux : à

l'oral, certains – comme le passé simple ou le subjonctif - ne doivent pas l'être sous peine de préciosité. Naturellement, une fois franchi le stade de la première initiation, la maîtrise du français passe nécessairement par l'approfondissement de ce système qui au cœur de cette langue mais il n'est pas nécessaire pour la conversation courante.

Tableau 2. Densités comparées des catégories grammaticales (fréquence pour mille mots) et indice de spécificité.

Catégories	A ‰ (Parlé)	B ‰ (Ecrit)	(A-B)/B ‰
Verbes	191.4	173.8	+ 10.1
Futurs	2.1	2.4	- 10.4
Conditionnels	4.0	3.8	+ 6.6
Présents	112.2	56.2	+ 99.8
Imparfais	17.6	37.6	- 53.3
Passés simple	0.0	16.8	- 100.0
Participes passés	24.3	22.8	+ 6.3
Participes présents	1.1	4.3	- 75.6
Infinitifs	30.1	30.0	0.1
Noms propres	8.2	27.7	- 70.4
Noms communs	130.0	176.1	- 26.1
Adjectifs	33.8	52.5	- 35.6
Adj. participe passé	2.7	10.3	- 73.7
Pronoms	196.5	144.8	+ 35.7
Pronoms personnels	104.3	92.4	+ 12.9
Déterminants	124.4	158.7	- 21.6
Articles	89.1	112.6	- 20.8
Nombres	14.7	10.0	+ 47.4
Possessifs	7.2	20.5	- 64.7
Démonstratifs	3.7	7.0	- 47.4
Indéfinis	9.6	8.6	+ 11.7
Adverbes	120.8	76.4	+ 58.1
Prépositions	108.9	138.2	- 21.2
Coordinations	31.0	26.4	+ 17.2
Subordination	39.9	22.3	+ 78.7
Mots étrangers	0.1	0.7	- 87.8

Lecture : dans les textes écrits, on rencontre en moyenne 173,8 verbes pour mille mots contre 191,4 dans le français parlé, soit 10,1 % de plus. Une telle différence a moins d'une chance sur mille de se produire au hasard (dernière colonne, signe +).

Le tableau 2 montre que, fondamentalement, les parties du discours sont liées entre elles et que leur emploi dépend d'un choix entre le verbe et le nom. Par exemple, un locuteur qui utilise beaucoup de verbes emploie nécessairement plus de pronoms, d'adverbes et de conjonctions de subordination par rapport à celui qui privilégie le nom et ses satellites (adjectifs, déterminants et prépositions).

Naturellement, il ne s'agit pas d'une délimitation stricte : certains pronoms relatifs ou certains adverbes sont employés aussi avec des noms ; de même, certains verbes peuvent être précédés ou suivis d'une préposition, etc.

Ces réserves admises, le décompte permet de vérifier que le français parlé privilégie le verbe (et ses satellites) au détriment du nom (et des éléments qui y sont associés). Le tableau 3 ci-dessous résume ces mouvements.

Tableau 3. Densités comparées des deux principaux groupes de catégories grammaticales dans les français écrit et parlé (pour mille mots)

Catégories	A (parlé)	B (Ecrit)	(A-B)/B %
Groupe du verbe	548.6	417.4	+31.4
Groupe du nom	436.3	579.6	-24.7

Les écarts sont considérables. Il n'y a aucune chance que ces différences puissent être le fruit du hasard. On peut donc conclure que :

- le français écrit est dominé par le substantif (associé aux adjectifs, déterminants et prépositions) ;
- le français parlé privilégie le verbe (associé aux pronoms, adverbes et conjonctions de subordination).

Halliday (1985) a fait le même constat pour l'anglais. Pour le français, certains chercheurs comme Blanche-Benveniste (2010) en ont eu l'intuition mais ne disposaient pas des corpus nécessaires pour une vérification.

Ce constat entraîne deux modifications dans les calculs présentés ci-dessus (fréquences des vocables et identification des vocables caractéristiques).

3.2 Calcul des fréquences par catégories grammaticales

Puisque le poids des catégories grammaticales diffère sensiblement entre l'oral spontané et l'écrit, il est normal que la plupart des verbes ou des pronoms – considérés individuellement – soient suremployés dans le premier registre alors que beaucoup de noms ou de déterminants paraîtront sous-employés. Pour connaître les verbes caractéristiques de l'un ou de l'autre de ces deux genres, il faut donc neutraliser ces propensions différentes à utiliser telle ou telle catégorie grammaticale.

La première solution consiste à placer, au dénominateur du calcul des fréquences, l'effectif total de la catégorie grammaticale et non plus le nombre total des mots. Par exemple, dans le corpus

du français parlé, il y a 695 598 occurrences d'un verbe. C'est la surface de texte couverte par cette catégorie grammaticale.

Dans le corpus du français parlé la fréquence du verbe "être" devient :

$$F_{\text{VerbesParléêtre}} = \frac{161670}{695\,598} = 232,4\%$$

Sur mille verbes employés dans la conversation, le verbe être en fournit 232, soit pas loin du quart – alors qu'à l'écrit, cette fréquence est de : 144,8‰, soit une augmentation de 60 % que l'on peut lire ainsi : dans le français parlé, par rapport à l'écrit, l'augmentation nette de l'utilisation du verbe être est de 60 %. L'adjectif "net" signifie que l'influence des variations de densité des catégories grammaticale a été neutralisée ou encore que l'on raisonne à structure grammaticale constante.

Rappelons que, en considérant la totalité des mots dans les deux corpus, ces fréquences sont respectivement de 44,5 ‰ et 25,2 ‰, soit + 77 % qui est la variation brute (produite par l'augmentation nette et par la préférence pour le verbe propre à la conversation).

Dans les tableaux qui vont suivre, c'est la variation nette qui sera utilisée.

Le test statistique présenté plus haut, pour déterminer les vocables caractéristiques, sera donc aménagé pour ne porter que sur la variation nette de densité entre les deux corpus.

3.3 Calcul des vocables caractéristiques

Deux méthodes sont possibles.

Premièrement, dans la suite de ce qui a été présenté dans le paragraphe précédent à propos de l'écart net, le test statistique porte non plus sur la fréquence calculée avec l'ensemble des mots, mais sur celle calculée pour chaque catégorie grammaticale.

Deuxièmement, on peut pondérer la fréquence globale du corpus de référence (ici le français écrit) par la structure grammaticale du corpus parlé. Par exemple, la densité moyenne des verbes est de 10,1 % supérieure dans le français parlé par rapport à l'écrit. Donc, si, dans la conversation, la fréquence d'un verbe est de 10,1 % plus forte qu'à l'écrit, cela signifie que son emploi est semblable dans les deux registres (compte tenu de cette propension différente à utiliser les verbes).

La quatrième colonne du tableau 2 ci-dessus fournit les coefficients de pondération. Ainsi, pour les verbes, ce coefficient est égal à 1.101. On multiplie, par ce coefficient, la fréquence de

chaque verbe, dans le français écrit, pour obtenir la fréquence théorique que devrait avoir, dans le français parlé, ce verbe sous l'hypothèse d'une absence de différence entre les deux registres.

Pour un verbe i , l'écart entre les deux fréquences sera :

$$\frac{F_{\text{parlé}_i} - (F_{\text{écrit}_i} * 1.101)}{F_{\text{écrit}_i} * 1.101}$$

On remarque que les deux procédés aboutissent au même résultat final.

Le test statistique permettant de déterminer le vocabulaire caractéristique est modifié selon le même principe (pour le détail du calcul : Monière & Labbé 2018).

Grâce à ces calculs, il est possible d'identifier quels sont les principaux vocables caractéristiques du français parlé et écrit. On examinera successivement le groupe du verbe puis celui du nom tels qu'ils ont été délimités plus haut. Les listes de mots sont arrêtées quand le taux de couverture dépasse les 90 %. Autrement dit, elles fournissent l'essentiel du vocabulaire nécessaire pour tenir une conversation en français "ordinaire" (Gadet).

Partie II.

Le groupe du verbe

En français, le verbe est identifié aisément : il est le seul à varier selon la personne, le temps et le mode. Il est normalement un élément indispensable de la phrase, même si l'on dénomme parfois "phrase" un mot ou un groupe de mots compris entre deux ponctuations fortes tout en étant dépourvus de verbe. Par exemple, les titres de chapitres ou de tableaux dans ce rapport. En dehors de ces "phrases nominales", le verbe est le pivot indispensable de la phrase.

Après avoir décrit les verbes les plus fréquents et les plus caractéristiques de la conversation, on examinera les principaux vocables associés aux verbes : pronoms, adverbes et conjonctions du subordination.

Enfin, les verbes peuvent se combiner entre eux. Les combinaisons les plus fréquentes seront présentées à l'aide de la notion d'auxiliaire qui rend compte de ce phénomène et de ses principales règles.

Chapitre 4.

Les verbes caractéristiques du français parlé

Si les principaux verbes restent les mêmes dans les deux genres, leur emploi varie de façon assez importante. Il s'agit d'abord des temps et des modes. Comme on l'a vu précédemment à propos des parties du discours, le verbe est nettement suremployé en français parlé. Par rapport à l'écrit, sa densité augmente de plus de 10 %. De plus cette propension plus forte à employer le verbe s'accompagne de changements au niveau des temps et de certains modes.

Dans ce chapitre, les caractéristiques du français parlé seront mesurées de trois manières : le rang parmi les verbes, la fréquence du vocable rapportée à la surface totale du texte puis à la surface occupée par les verbes (Fréquence et ‰ Verbes dans les prochains tableaux). Enfin, l'écart net entre les deux corpus est calculé pour chacun de ces verbes. Rappelons que ce calcul consiste à neutraliser l'impact de la propension plus forte à utiliser les verbes dans la conversation qu'à l'écrit pour mettre en lumière les préférences portant sur chacun d'eux. Le classement des vocables est fait par fréquences décroissantes dans le corpus parlé et descend jusqu'à comprendre au moins les neuf dixièmes de la surface totale occupée par les verbes dans le corpus parlé. Nous répéterons cette opération pour toutes les catégories grammaticales, de telle sorte que ces listes donneront le vocabulaire du français fondamental de la conversation.

4.1 Prépondérance du présent

On a vu plus haut que le français parlé privilégie le présent et sous-emploie les autres temps, spécialement le passé. Mais ce calcul portait sur les variations brutes. Le tableau 4 présente les fréquences pour la catégorie des verbes (‰ Verbes) sur lesquelles sont calculées les variations nettes (dernière colonne).

A propension égale, la densité des participes passés est sensiblement la même dans les deux registres (nous revenons plus en détail sur ce point dans le prochain chapitre). Toutes les autres lignes indiquent des écarts nets importants. Ils sont tous négatifs, sauf pour le présent. Dans le français parlé, presque six verbes sur 10 (586‰) sont utilisés à ce temps contre à peine plus de trois sur dix (323‰) à l'écrit. Il s'agit souvent d'un présent de narration puisque la majorité des entretiens portent sur des événements passés que les enquêteurs demandaient de raconter.

Tableau 4. Les temps et modes des verbes dans les français parlé et écrit (fréquence en pour mille verbes)

	Parlé		Ecrit		Ecart net
	Fréquence	‰Verbes	Fréquence	‰ Verbes	
Indicatif et subjonctif					
futurs	2,1	11,2	2,4	13,7	-19
présents	112,2	586,4	56,2	323,0	+ 82
imparfaits	17,6	91,8	37,6	216,3	-58
passés simples	0,0	0,2	16,8	96,4	-100
participes passés	24,3	126,9	22,9	131,4	-3
participes présents	1,1	5,5	4,3	24,8	-78
infinitifs	30,1	157,0	30,0	172,7	-9
conditionnel	4,0	20,9	3,8	21,6	-3
Total verbes	191,4	1000,0	173,8	1000,0	

Lecture : Dans le français parlé, pour mille mots, il y a 24,3 verbes au participe passé. Sur mille verbes 126,9 sont des participes passés. Dans le français écrit, la fréquence de ces participes passés est 22,9 pour 1000 mots et ils représentent 131,4 pour mille verbes. Dans le français parlé il y a 3 % de participes passés en moins par rapport à l'écrit (une fois neutralisé l'effet de la propension à utiliser les verbes plus forte dans la conversation).

Autre constat : à l'écrit, outre le présent, l'imparfait et le passé simple sont les temps privilégiés ; les temps composés, avec un auxiliaire et un participe passé, sont moins importants. Dans le français parlé, le passé simple disparaît, alors qu'il représente près d'un emploi sur dix à l'écrit (c'est un temps très utilisé pour le récit). La densité nette de l'imparfait recule de plus de moitié. Outre le présent de narration, chez la plupart des locuteurs, le passé se réduit au passé composé dont nous verrons plus bas qu'il s'agit souvent d'un présent. Même le futur est significativement sous-employé, en partie remplacé par "aller" (au présent) suivi d'un infinitif ("demain, je vais faire...")¹.

4.2 Recensement des verbes usuels du français parlé

Dans le tableau 5, en cumulant les chiffres de la quatrième colonne (‰Verbes), on constate que les quatre premiers couvrent à eux seuls plus de la moitié de toutes les occurrences des verbes, que la proportion des trois quarts est atteinte au 29^e (*essayer*) et celle des 90 % au 163^e (*supporter*). Ces verbes peuvent être considérés comme la base de la conversation en français usuel.

Tableau 5. Les principaux verbes utilisés dans les français écrit et parlé (fréquence pour mille mots)

Vocables	Parlé (A)			Ecrit (B)			(A-B)/B %
	Rang	Fréquence	‰Verbes	Rang	Fréquence	‰Verbes	

¹ Voir dans le prochain chapitre, le paragraphe consacré aux auxiliaires modaux suivis d'un infinitif.

être	1	44,48	232,4	1	25,18	144,8	+ 60
avoir	2	35,50	185,5	2	21,89	125,9	+ 47
faire	3	9,88	51,6	3	5,47	31,5	+ 64
dire	4	7,72	40,3	4	4,62	26,6	+ 52
aller	5	5,62	29,3	5	3,18	18,3	+ 60
pouvoir	6	4,25	22,2	6	2,98	17,1	+ 30
savoir	7	4,18	21,8	8	2,23	12,8	+ 70
falloir	8	3,40	17,7	14	1,21	7,0	+ 155
voir	9	3,25	17,0	7	2,45	14,1	+ 20
vouloir	10	3,03	15,9	9	1,82	10,5	+ 51
penser	11	2,53	13,2	22	0,92	5,3	+ 150
mettre	12	2,01	10,5	18	1,10	6,3	+ 66
trouver	13	1,60	8,4	17	1,12	6,4	+ 30
passer	14	1,48	7,7	13	1,29	7,4	+ 4
croire	15	1,37	7,1	19	1,06	6,1	+ 18
prendre	16	1,34	7,0	11	1,54	8,9	- 21
parler	17	1,31	6,8	15	1,13	6,5	+ 6
arriver	18	1,30	6,8	31	0,75	4,3	+ 57
devoir	19	1,27	6,6	12	1,43	8,3	-20
venir	20	1,21	6,3	10	1,61	9,3	-32
donner	21	1,10	5,7	23	0,91	5,3	+ 9
travailler	22	1,07	5,6	105	0,21	1,2	+ 365
payer	23	0,79	4,1	130	0,17	1,0	+ 332
partir	24	0,78	4,1	41	0,49	2,8	+ 43
changer	25	0,77	4,0	82	0,26	1,5	+ 168
voter	26	0,76	4,0	1134	0,01	0,1	-
connaître	27	0,74	3,9	30	0,75	4,3	- 10
demander	28	0,67	3,5	16	1,12	6,4	- 46
essayer	29	0,66	3,5	58	0,38	2,2	+ 57
vivre	30	0,60	3,1	51	0,41	2,4	+ 32
rester	31	0,58	3,0	24	0,85	4,9	- 38
comprendre	32	0,56	2,9	27	0,80	4,6	- 37
entendre	33	0,54	2,8	28	0,78	4,5	-38
regarder	34	0,52	2,7	20	1,04	6,0	- 54
aimer	35	0,51	2,7	33	0,64	3,7	- 27
marcher	36	0,48	2,5	63	0,35	2,0	+ 23
chauffer	37	0,47	2,5	778	0,02	0,1	-
appeler	38	0,44	2,3	37	0,54	3,1	- 26
rentrer	39	0,44	2,3	80	0,27	1,6	+ 46
revenir	40	0,43	2,2	39	0,51	3,0	- 25
sortir	41	0,42	2,2	32	0,69	4,0	- 44
laisser	42	0,42	2,2	21	0,92	5,3	- 59
commencer	43	0,40	2,1	42	0,49	2,8	- 26
attendre	44	0,39	2,1	26	0,82	4,7	- 56
expliquer	45	0,38	2,0	72	0,29	1,6	+ 22
servir	46	0,37	1,9	74	0,28	1,6	+ 17
poser	47	0,37	1,9	57	0,39	2,2	- 14
sentir	48	0,36	1,9	29	0,77	4,4	- 58
tenir	49	0,35	1,8	25	0,83	4,8	- 62
régler	50	0,35	1,8	356	0,06	0,4	+ 407
arrêter	51	0,34	1,8	45	0,47	2,7	- 35
permettre	52	0,32	1,7	107	0,21	1,2	+ 42
obliger	53	0,32	1,7	208	0,11	0,6	+ 169
intéresser	54	0,32	1,7	178	0,13	0,7	+ 127
rendre	55	0,31	1,6	40	0,51	2,9	- 45
occuper	56	0,31	1,6	96	0,22	1,3	+ 25
toucher	57	0,30	1,6	114	0,20	1,2	+ 37
acheter	58	0,30	1,6	183	0,12	0,7	+ 120

chercher	59	0,29	1,5	34	0,57	3,3	- 54
suivre	60	0,28	1,5	54	0,39	2,3	- 34
écouter	61	0,28	1,5	64	0,34	2,0	- 27
ouvrir	61	0,28	1,5	38	0,54	3,1	- 53
utiliser	63	0,27	1,4	332	0,07	0,4	+ 255
importer	64	0,25	1,3	119	0,19	1,1	+ 24
gagner	65	0,25	1,3	142	0,15	0,9	+ 47
habiter	66	0,24	1,3	204	0,11	0,6	+ 95
aider	67	0,24	1,2	127	0,17	1,0	+ 25
retrouver	68	0,23	1,2	55	0,39	2,2	- 47
monter	69	0,23	1,2	55	0,39	2,2	- 47
installer	70	0,22	1,1	145	0,15	0,9	+ 32
discuter	71	0,22	1,1	330	0,07	0,4	+ 190
sembler	72	0,22	1,1	36	0,56	3,2	- 64
fonctionner	73	0,21	1,1	547	0,04	0,2	+ 444
proposer	74	0,21	1,1	172	0,13	0,8	+ 46
devenir	75	0,21	1,1	43	0,49	2,8	- 60
répondre	76	0,21	1,1	35	0,56	3,2	- 66
remettre	77	0,21	1,1	113	0,20	1,2	- 5
tomber	78	0,21	1,1	49	0,44	2,5	- 57
tourner	79	0,20	1,1	50	0,42	2,4	- 56
continuer	80	0,20	1,1	61	0,36	2,1	- 49
exister	80	0,20	1,1	152	0,15	0,8	+ 26
fermer	82	0,20	1,0	102	0,21	1,2	- 14
perdre	83	0,20	1,0	60	0,37	2,1	- 51
rencontrer	84	0,20	1,0	118	0,19	1,1	- 5
couper	85	0,20	1,0	169	0,14	0,8	+ 32
amener	86	0,20	1,0	250	0,09	0,5	+ 95
coûter	87	0,19	1,0	508	0,04	0,2	+ 344
dépendre	88	0,19	1,0	508	0,04	0,2	+ 332
finir	88	0,19	1,0	56	0,39	2,2	- 56
apporter	90	0,18	1,0	167	0,14	0,8	+ 23
consommer	91	0,18	1,0	1298	0,01	0,0	+ 1960
vendre	92	0,18	1,0	298	0,08	0,4	+ 118
concerner	93	0,18	0,9	310	0,07	0,4	+ 124
présenter	93	0,18	0,9	162	0,14	0,8	+ 16
lire	95	0,18	0,9	69	0,32	1,8	- 49
allumer	96	0,18	0,9	159	0,14	0,8	+ 15
recevoir	97	0,17	0,9	93	0,24	1,4	- 33
jouer	98	0,17	0,9	71	0,30	1,7	- 49
apprendre	99	0,17	0,9	87	0,25	1,5	- 38
créer	100	0,17	0,9	371	0,06	0,3	+ 163
garder	101	0,17	0,9	91	0,24	1,4	-37
choisir	102	0,16	0,9	157	0,14	0,8	6
accepter	103	0,16	0,9	154	0,15	0,8	2
paraître	104	0,16	0,8	46	0,46	2,7	-68
décider	105	0,16	0,8	104	0,21	1,2	-32
laver	105	0,16	0,8	438	0,05	0,3	206
rappeler	107	0,15	0,8	112	0,20	1,2	-32
augmenter	108	0,15	0,8	757	0,02	0,1	528
préférer	109	0,15	0,8	147	0,15	0,9	-9
valoir	110	0,15	0,8	128	0,17	1,0	-21
gérer	111	0,15	0,8	1251	0,01	0,0	1472
plaire	112	0,14	0,8	177	0,13	0,7	2
baisser	113	0,14	0,7	216	0,11	0,6	22
foutre	114	0,14	0,7	98	0,22	1,3	-41
manger	115	0,14	0,7	99	0,22	1,3	-41
envoyer	116	0,14	0,7	122	0,18	1,0	-30

apercevoir	117	0,14	0,7	73	0,28	1,6	-56
manquer	117	0,14	0,7	108	0,21	1,2	-40
espérer	119	0,13	0,7	141	0,15	0,9	-20
représenter	119	0,13	0,7	328	0,07	0,4	77
construire	121	0,13	0,7	528	0,04	0,2	219
reprendre	122	0,13	0,7	59	0,38	2,2	-69
montrer	123	0,13	0,7	81	0,26	1,5	-56
défendre	124	0,13	0,7	243	0,09	0,5	24
négociier	125	0,13	0,7	981	0,01	0,1	745
programmer	126	0,12	0,6	1667	0,00	0,0	2235
réussir	127	0,12	0,6	171	0,13	0,8	-18
bouger	128	0,12	0,6	125	0,18	1,0	-39
exprimer	129	0,12	0,6	294	0,08	0,4	38
refaire	130	0,12	0,6	501	0,04	0,2	165
évoluer	131	0,12	0,6	883	0,02	0,1	543
souvenir	132	0,11	0,6	86	0,25	1,5	-59
avancer	133	0,11	0,6	103	0,21	1,2	-51
oublier	134	0,11	0,6	84	0,26	1,5	-61
suffire	135	0,11	0,6	131	0,16	0,9	-40
améliorer	136	0,11	0,6	1352	0,01	0,0	1196
tirer	137	0,11	0,6	52	0,40	2,3	-76
respecter	138	0,10	0,5	575	0,03	0,2	181
enlever	139	0,10	0,5	278	0,08	0,5	16
battre	140	0,10	0,5	149	0,15	0,9	-37
boire	141	0,10	0,5	90	0,24	1,4	-62
éteindre	142	0,10	0,5	259	0,09	0,5	5
reconnaître	143	0,10	0,5	78	0,27	1,6	-67
prévoir	144	0,10	0,5	353	0,06	0,4	44
gêner	145	0,10	0,5	376	0,06	0,3	55
éviter	146	0,10	0,5	181	0,13	0,7	-29
compter	147	0,10	0,5	100	0,21	1,2	-59
descendre	148	0,10	0,5	70	0,31	1,8	-72
participer	149	0,09	0,5	640	0,03	0,2	199
développer	150	0,09	0,5	669	0,03	0,2	226
marquer	151	0,09	0,5	279	0,08	0,5	2
tromper	152	0,09	0,5	188	0,12	0,7	-31
téléphoner	153	0,09	0,5	265	0,09	0,5	-7
durer	154	0,09	0,5	247	0,09	0,5	-13
ressentir	154	0,09	0,5	283	0,08	0,5	-2
estimer	156	0,09	0,5	420	0,05	0,3	58
réfléchir	156	0,09	0,5	140	0,15	0,9	-49
élire	158	0,09	0,4	1089	0,01	0,1	598
porter	159	0,09	0,4	47	0,44	2,5	-82
intervenir	160	0,09	0,4	408	0,05	0,3	50
satisfaire	161	0,09	0,4	621	0,03	0,2	158
placer	162	0,08	0,4	260	0,09	0,5	-12
dépenser	163	0,08	0,4	900	0,02	0,1	375
supporter	163	0,08	0,4	44	0,07	0,4	15

Lecture : dans le français parlé, le verbe le plus utilisé est le verbe *être* avec une fréquence de 44,5 pour mille mots, ce qui représente 232,2 pour mille emplois d'un verbe. Dans le français écrit, c'est aussi le verbe le plus employé avec une fréquence de 25,2 pour mille mots soit 144,8 pour mille emplois d'un verbe. La fréquence de ce verbe est, en moyenne, de 77 % plus élevée à l'oral qu'à l'écrit (dernière colonne).

Pour deux verbes, au moins, la comparaison parlé/écrit est impossible car leur densité dans le corpus "oral" s'explique par le contenu d'un certain nombre d'entretiens : "voter" dans le cadre de deux enquêtes et deux sondages sur les comportements politiques ; "chauffer" dans les

recherches sur le confort de la maison. Ce problème du thème peut concerner deux ou trois autres verbes comme "consommer", "allumer", etc.

Dans le français parlé, "être" et "avoir" fournissent 42 % des verbes et avec les deux suivants (faire et dire), ils représentent plus de la moitié des apparitions d'un verbe. En y ajoutant "aller" et "pouvoir", les six premiers sont communs aux deux registres, avec les mêmes rangs. C'est dire combien ces six-là sont au cœur de la langue française. Tout apprentissage du français commence nécessairement par eux.

4.3 Les verbes caractéristiques de la conversation

En considérant l'écart net (dernière colonne) - c'est-à-dire après que soit neutralisée la propension plus forte à utiliser des verbes dans la conversation qu'à l'écrit -, il apparaît que dix verbes ont des densités d'utilisation non significativement différentes dans les deux corpus, soit dans l'ordre du tableau : parler, passer, remettre, rencontrer, choisir, accepter, plaire, éteindre, marquer, ressentir. Pour tous les autres, les écarts sont significatifs au seuil choisi (une chance d'erreur sur cent). Le signe en dernière colonne permet donc de conclure au suremploi (+) ou au sous-emploi caractéristique dans le français parlé (-).

En première analyse, quatre constatations s'imposent.

Premièrement, les six premiers verbes - présents dans les deux corpus et dans le même ordre - sont nettement plus mobilisés dans la conversation qu'à l'écrit. *Etre, avoir, faire, dire, aller et pouvoir* - dans cet ordre - sont au cœur du français parlé.

Deuxièmement, en considérant les cent premiers, une soixantaine sont suremployés dans la conversation (signe + en dernière colonne du tableau) contre 36 sous-employés, quatre étant sensiblement aussi présents dans les deux genres. Nous avons déjà signalé que le vocabulaire de la conversation est réduit par rapport à celui de l'écrit. Cette réduction se fait donc d'abord au profit des verbes les plus usuels, ce qui est logique puisque l'écrit donne le temps au remord, à la chasse aux répétitions et à la précision du vocabulaire, ce que ne permet pas le flux des échanges oraux.

Troisièmement, les deux premiers sont aussi des auxiliaires. Pour "avoir", cet auxiliaire représente environ les deux tiers des emplois, ce qui est logique puisque le passé composé est, après le présent de narration, la manière usuelle de parler du passé. Quant à "être", outre le réfléchi – à l'écrit comme dans la conversation – il a essentiellement une fonction de copule, l'attribut pouvant être un adjectif, un nom (comme on le verra en troisième partie) ou un

participe passé. Toutefois, le comportement de ces auxiliaires est assez différent dans les deux registres comme expliqué plus bas.

Quatrièmement, le verbe "faire" arrive toujours en troisième position mais en première si l'on admet que "être" comme "avoir" sont surtout des outils (en position d'auxiliaires) et non des verbes au sens plein. Ainsi se vérifie la formule selon laquelle « le verbe *faire* est la bonne à tout faire de la langue française ». En effet, il peut se substituer à beaucoup de verbes d'action (par exemple : "faire une conquête" permet d'éviter le redoutable "conquérir" tout comme "faire une découverte" pour "découvrir"). D'où son utilisation nettement plus importante dans le français parlé qu'à l'écrit.

Du point de vue des suremplois, "falloir" est le verbe fréquent qui connaît la plus forte augmentation entre l'oral spontané et l'écrit (multiplication par 2,55 par rapport à l'écrit). Il exprime un impératif impersonnel (voir le pronom "il" discuté plus bas) et une obligation imprécise (ou sans destinataire) contrairement à *devoir* (en recul) qui la plupart du temps vise une obligation légale ou morale (et secondairement une forte probabilité), raisons pour lesquelles la conversation s'en méfie.

"Penser" est l'autre verbe très utilisé dans le français parlé par rapport à l'écrit. Un examen des contextes d'emploi montre que son sens va de *déduire* à *croire* en passant par *estimer* et *juger*. En quelque sorte, il remplit pour la pensée le rôle de "faire" pour l'action.

Il faut également souligner l'emploi très caractéristique de "aller" qui désigne rarement un déplacement physique ("aller au bureau") mais représente un véritable auxiliaire de temps, exprimant le futur proche, sinon immédiat : "aller faire", "aller dire"... L'intérêt est double : utiliser l'infinitif du verbe en s'épargnant certains futurs irréguliers (la racine du verbe au futur n'est pas son infinitif) et établir une continuité entre la conversation et cette action à venir : "je vais le faire" est plus fort que "je le ferai". On verra plus bas que l'usage de ce verbe – au passé – pose un problème de confusion avec le verbe "être".

Enfin, seuls dans cette tête de liste, *prendre*, *devoir* et *venir* (contrairement à *aller*) sont en recul. L'examen des contextes d'emploi permet de comprendre cette relative désaffection. Par exemple, les deux sens du verbe *devoir* – évoqués plus haut. Quant à "venir", la majorité des emplois servent d'auxiliaires exprimant un passé immédiat par rapport au présent de l'énonciation (*venir de faire*) qui semble être surtout un procédé littéraire au service du récit. La conversation privilégie plutôt le présent de narration ou le passé composé.

Ces emplois peuvent être retrouvés aisément à l'aide du concordancier et peuvent fournir de bons exemples pour la définition de chacun de ces verbes usuels. En particulier, ces contextes ont permis de repérer plusieurs cas particuliers caractéristiques du français parlé dont un

particulièrement remarquable (les passés composés avec être). Ils ont également servi à l'étude des auxiliaires du français parlé.

Chapitre 5.

L'auxiliaire

Depuis Saussure, on sait que le sens d'un mot lui vient de deux types de relations. Sa place dans la langue résulte des relations paradigmatiques qu'il entretient avec d'autres mots (hyperonymies, hyponymies, antonymies, etc.) et des associations avec d'autres vocables que le discours réalise au fur et à mesure qu'il se déroule (syntagmes). Ces relations syntagmatiques peuvent être recensées grâce aux vastes corpus étiquetés.

Naturellement, toute association de plusieurs mots, à la surface d'un texte, est porteuse de sens. Mais pour qu'elle soit statistiquement significative, il faut qu'elle se répète et, plus elle se répète, plus elle caractérisera le discours qui l'utilise (Sur cette notion de syntagme répété, voir, Pibarot & Al. 1995).

On examine dans ce chapitre, la combinaison de deux verbes ensemble. Les combinaisons impliquant des noms seront étudiées dans la prochaine partie. Ces combinaisons de deux verbes (ou plus) peuvent se faire selon deux types principaux.

5.1 Les deux types de combinaisons des verbes

Le premier type est bien connu et se rapporte au système complexe du temps des verbes dans la langue française (Benveniste 1959 et 1965). Pour former un passé composé, un plus-que-parfait, un passé ou un futur antérieurs, il faut placer un auxiliaire (avoir ou être) devant un autre verbe au participe passé. Le choix est dicté par le sens que le locuteur veut communiquer. De plus, si l'auxiliaire "avoir" suivi d'un participe passé est – apparemment – toujours un passé, l'utilisation de l'auxiliaire "être" est plus complexe car s'ajoutent aux passés, les formes passives et pronominales. Rappelons que le français parlé utilise surtout le présent (deux fois plus qu'à l'écrit). Le passé simple disparaît et l'usage de l'imparfait diminue de plus de moitié par rapport à l'écrit. En revanche, les participes passés sont sensiblement autant utilisés dans les deux registres (une fois neutralisée la propension plus forte à utiliser le verbe dans les échanges oraux).

Le second type de syntagme verbal est moins bien connu : en français, il est possible de combiner deux verbes – autre que "avoir" et "être" – en faisant jouer au premier le rôle d'un auxiliaire du second (qui sera placé à l'infinitif et non plus au participe passé). Par exemple "vouloir faire" place dans l'ordre de la volonté l'action alors que "pouvoir faire" la place dans la dimension du possible, etc.

A la suite de Benveniste (1965), nous appelons les verbes - qui peuvent se trouver en première position dans ce type de combinaisons - des "auxiliaires modaux" ou des "verbes modalisateurs" et les seconds, en position de compléments, les verbes "modalisés" (Labbé & Labbé 2013). Enfin, on tiendra compte de ce que certains mots peuvent se glisser entre l'auxiliaire et le verbe complément (avoir *mal* fait, vouloir *bien* faire, vouloir *le* faire, venir *de* faire...).

On demande donc à l'ordinateur de retrouver toutes ces combinaisons, en utilisant les lemmes : toutes les flexions de l'auxiliaire sont rattachées à l'infinitif de ce verbe.

On obtient d'abord une indication sur la fréquence relative du phénomène (tableau 6).

Tableau 6. Densité des deux types de constructions "verbe + verbe" dans les français parlé et écrit (fréquence en pour mille mots)

Combinaisons	Parlé		Ecrit		Ecart
	Fréquence	%Verbes	Fréquence	%Verbes	
Participes passés					
avoir	19,1	99,9	14,8	85,1	+17
être	8,5	44,4	8,1	46,3	-4
Total	27,6	144,3	22,9	131,4	+10
Infinitifs	29,5	154,1	23,7	136,2	+13

Lecture : dans le français parlé, sur 1000 mots, il y a 19,1 verbes au participe passé avec l'auxiliaire avoir, soit pratiquement un dixième des verbes. Dans le français écrit, sur 1000 mots, il y a 14,8 participes passés avec l'auxiliaire avoir, soit 85,1 % verbes. L'écart net entre la densité d'emploi dans les deux genres est de +17 % dans le français parlé.

Ce décompte amène au moins trois constats :

- Il s'agit de phénomènes massifs. Si l'on tient compte de ce que ces combinaisons mobilisent au moins deux verbes (l'auxiliaire et le complément), il faut multiplier par deux les proportions données dans les deux colonnes %Verbes. Il apparaît alors que, dans le français parlé, 60 % des occurrences d'un verbe entrent dans l'un de ces deux types de combinaisons d'un verbe complément précédé d'un auxiliaire. A l'écrit, c'est encore 54 % des occurrences d'un verbe qui sont dans cette situation. Autrement dit, la combinaison de deux verbes – l'un comme auxiliaire et l'autre comme complément - est un outil essentiel du français, parlé comme écrit.
- Pour les participes passés, une fois neutralisée la propension plus forte à utiliser des verbes dans le français parlé, on observe une nette préférence pour les participes passés avec "avoir" et pour les associations d'un verbe "modal" avec un infinitif.
- Dans les deux genres, les combinaisons associant un modalisateur suivi d'un verbe à l'infinitif sont plus nombreuses que les participes passés et couvrent une surface de texte supérieure. Et ceci plus encore dans la conversation que dans le français écrit.

Il est donc surprenant que cette seconde construction si fréquente soit peu mentionnée dans les grammaires et les ouvrages sur le français. Par exemple, Chevalier et Al. (1964) mentionnent le seul verbe "faire" comme pouvant être suivi d'un infinitif (*faire voir*).

Enfin, une caractéristique supplémentaire mérite d'être notée : la grande diversité des combinaisons possibles. L'examen des "taux de couverture" du texte par ces combinaisons, classées par effectifs décroissants, en se plaçant aux bornes utilisées précédemment pour délimiter les verbes usuels (tableau 7). Ici, le calcul utilise comme dénominateur la surface couverte par l'ensemble des participes passés de l'auxiliaire considéré.

Tableau 7. Les taux de couverture du texte par les combinaisons les plus fréquentes formées avec les auxiliaires avoir et être

Couverture	Parlé			Ecrit		
	avoir	être	être/avoir	avoir	être	être/avoir
50 %	13e	41e	3,2	36e	93e	2,6
75 %	61e	172e	2,9	169e	340e	2,0
90 %	208e	481e	2,3	491e	830e	1,7

Lecture : dans le français parlé, les 13 premières combinaisons couvrent plus de 50 % des emplois d'un participe passé avec *avoir* alors qu'il en faut 41 pour atteindre ce seuil avec l'auxiliaire *être* soit 3,2 fois plus. A l'écrit ces proportions sont : 36 avec *avoir* et 93 avec *être* soit 2,6 fois plus.

Rappelons que pour la totalité des verbes (chapitre précédent), la moitié des occurrences est atteinte au 4^e verbe et les neuf-dixièmes au 163^e. Les combinaisons des verbes sont donc beaucoup plus diluées, surtout avec le verbe "être" (en moyenne trois fois plus qu'avec l'auxiliaire *avoir*). Le même phénomène est observé à l'écrit avec un écart moindre entre les deux auxiliaires.

Nous examinerons successivement les principales constructions rencontrées dans les deux genres puis les verbes auxiliaires autres que "être" et "avoir". Nous verrons que ces deux auxiliaires remplissent des fonctions différentes.

5.2 L'auxiliaire "avoir"

L'auxiliaire "avoir" suivi d'un participe passé indique un « fait passé et achevé au moment où l'on parle » (Chevalier & Al.). Cependant, le passé composé peut aussi laisser entendre que le fait, ou certaines de ses conséquences, peuvent se prolonger jusqu'au présent de l'énonciation : "il a bien parlé" peut signifier que le discours s'achève à l'instant et qu'il aura peut-être des effets dans le futur. C'est la raison pour laquelle des grammairiens, comme Wagner et Pichon (1962), parlent de « passé indéfini », voire de « forme composée du présent ».

En revanche, le plus-que-parfait – auxiliaire à l'imparfait - indique que le fait s'inscrit dans un passé révolu ("il avait bien parlé"). Enfin, il faut rappeler que l'auxiliaire devant un participe passé peut aussi être mis au conditionnel ou au subjonctif et prendre tous les temps, selon la nuance que le locuteur veut donner à son propos, notamment le passé simple (formant le "passé antérieur") et le futur ("antérieur"). Dès lors la question se pose : quel est le poids de ces divers temps dans les deux genres ? Le tableau 8 répond à cette question pour l'auxiliaire "avoir".

Tableau 8. Les temps et modes de l'auxiliaire avoir devant un participe passé

Temps de l'auxiliaire	% Parlé	% Ecrit	Ecart net
Indicatif, subjonctif Présent	83,6	54,4	+54
Imparfait	10,7	32,5	-67
Infinitif	0,9	3,1	-72
Futur	0,4	0,5	-23
Participe présent	0,1	2,3	-94
Participe passé	0,1	0,1	-
Passé simple	0,0	1,7	-100
Conditionnel (présent et passé)	4,3	5,5	-23

Lecture : dans le français parlé, l'auxiliaire avoir devant un participe passé est mis au présent dans 83,6% des occurrences contre 54,4 % dans le français écrit, soit un écart net de + 54 % dans le premier par rapport au second. Pour le participe passé, la faiblesse des effectifs interdit le calcul de l'écart.

Les tendances, déjà notées pour l'ensemble des verbes du français parlé, se confirment au niveau du principal auxiliaire (avoir) : suprématie du présent et poids négligeable des autres temps excepté une présence de l'imparfait (qui recule cependant des deux tiers par rapport à l'écrit). Le passé simple et le participe présent disparaissent. En fait, dans le français parlé, le passé composé se distingue du présent "indéfini" grâce à l'adjonction d'adverbes (comme "hier", "autrefois"...) ou de petits groupes nominaux (comme *l'an dernier*) qui ancrent le propos dans un point, plus ou moins précis, du passé, ce que ne fait pas le passé composé seul. Quels sont les verbes compléments de l'auxiliaire *avoir* ? Le tableau 7 ci-dessus a déjà suggéré qu'ils sont en grand nombre. La liste du tableau 9 ci-dessous s'arrête lorsqu'on atteint la moitié de cette surface car elle suffit pour constater qu'elle redouble largement celle présentée dans le tableau 5 ci-dessus (les verbes usuels du français oral).

Tableau 9. Les verbes les plus utilisés au participe passé avec l'auxiliaire "avoir" dans le français parlé

Complément	Parlé		Ecrit		écart net
	rang	%o participe passé	rang	%o participe passé	
être	1	112,9	1	72,2	+56
avoir	2	107,7	4	36,5	+195
faire	3	104,7	2	55,6	+88

dire	4	79,6	3	41,0	+94
voir	5	27,4	5	33,7	-19
mettre	6	22,7	13	10,4	+119
pouvoir	7	18,2	6	27,0	-33
prendre	8	17,9	7	19,6	-9
changer	9	12,7	32	4,9	+159
demander	10	12,3	11	11,3	+9
devoir	11	12,1	8	19,3	-37
parler	12	11,3	17	9,2	+23
donner	13	10,9	14	10,2	+7

Lecture : dans le français parlé le verbe "être" (été) est le plus utilisé au participe passé avec l'auxiliaire "avoir", il représente 129,5 pour mille participes passés précédés de "avoir". A l'écrit, c'est aussi le participe passé "été" qui est le plus employé (72,2 % des participes passés avec avoir). La densité nette d'utilisation de cette construction est de 79 % plus élevée dans le français parlé par rapport à l'écrit.

Les quatre premiers figurent presque dans le même ordre dans les deux corpus mais avec une densité d'emploi nettement plus élevée dans le français parlé. C'est particulièrement vrai pour la composition "avoir eu", nettement plus fréquente dans le français parlé, sur le modèle : "il y a eu un (petit) problème", "j'ai eu un (gros) souci", "un accident", etc.

Beaucoup des verbes les plus usuels figurent dans la tête de liste des verbes les plus fréquents, ce qui prouve que les formes composées avec l'auxiliaire "avoir" sont utilisées pour marquer le temps aussi bien à l'écrit qu'à l'oral et que c'est même la manière usuelle de mettre un verbe au "passé".

Enfin, la combinaison "avoir été" est la plus fréquente. Comme on le verra plus bas, elle introduit un attribut du sujet (adjectif, nom, voire participe passé) mais le français parlé ajoute un autre sens qu'on ne trouve pas à l'écrit.

A part "donner", tous les écarts sont significatifs. La majorité d'entre eux sont positifs. Cela confirme que le français parlé utilise un nombre plus limité de verbes et qu'il se concentre sur les plus usuels.

Cependant, on note que, outre "prendre", les verbes "voir", "pouvoir" et "devoir" sont significativement sous-employés dans le français parlé. Un peu plus bas dans la liste on trouve les participes passés de "vouloir" (-18 %) et "savoir" (-32 %). Nous avons déjà signalé que ces verbes (*pouvoir*, *devoir*, *vouloir* et *savoir* comme bien d'autres, notamment *voir*) sont des auxiliaires "modaux". Manifestement beaucoup de locuteurs répugnent à utiliser des "surcompositions" consistant à mettre au passé composé le premier élément de la combinaison d'un verbe "modal" suivi d'un infinitif (*il a voulu faire*).

5.3 L'auxiliaire "être"

Les fonctions de l'auxiliaire "être" sont plus diverses que celles de "avoir". Certes, la majorité relative de ses emplois ont le même rapport avec le passé que ceux de l'auxiliaire "avoir", soit parce que le verbe n'admet que l'auxiliaire "être" (comme *aller, venir, partir* ou *mourir* : *il est mort, il était mort*), soit que le verbe admette les deux auxiliaires, avec des sens différents (*appeler, descendre, monter, paraître, passer...*). Pour ce dernier cas, les grammaires signalent environ une cinquantaine de verbes, mais l'examen du vocabulaire usuel du français parlé montre que leur liste est nettement plus longue. Il faut ajouter à cela les formes passives : "il est appelé" n'est pas le symétrique de "il a appelé", car le premier est situé dans le présent de l'énonciation – le participe passé fonctionne ici comme un adjectif attribut du sujet - et il faudrait dire "il a été appelé", ou mieux "il avait été appelé", pour situer explicitement cette action dans le passé (le symétrique à la forme active serait le plus-que-parfait : "il avait appelé"). L'auxiliaire "être" est également utilisé dans les formes réfléchies (*il s'est souvenu, il s'était souvenu, il s'en est allé...*) mais ici les formes composées sont impossibles (le pronom réfléchi ne peut être employé devant "avoir").

Enfin, le passé composé avec "être" sert aussi à exprimer une "nuance d'atténuation" (Wagner et Pichon) par rapport au même fait exprimé au présent : "je suis venu vous dire" est perçu comme moins direct ou plus respectueux que "je viens vous dire".

Comme on le voit, l'auxiliaire "être" remplit une multitude de fonctions – au-delà de l'expression du passé. Le poids de ces différents emplois n'est pas le même dans les deux registres comme le suggère l'examen des temps et modes de l'auxiliaire (tableau 10).

Tableau 10. Les temps et modes de l'auxiliaire *être* devant un participe passé

Temps et modes	% Parlé	% Ecrit	Ecart
Indicatif et subjonctif			
Présent	70,4	45,2	+56
Imparfait	11,3	35,1	-68
Participe passé (été)	10,0	6,8	+47
Infinitif	4,9	5,7	-13
Futur	1,1	1,4	-19
Passé simple	0,0	3,1	-100
Participe présent	0,6	0,5	-
Conditionnel	1,4	2,7	-48

Lecture : voir tableau... ci-dessus. Les effectifs du participe présent sont trop faibles pour calculer un écart

Dans le français parlé, la prédominance du présent est presque aussi forte qu'avec l'auxiliaire *avoir*. On observe également le recul considérable de l'imparfait et la disparition du passé

simple. La seule surprise réside dans le participe passé "été" dont la densité augmente de près de moitié par rapport à l'écrit. Nous verrons plus bas les raisons de cette étrangeté.

Comme indiqué ci-dessus, le verbe "être" fonctionne surtout comme une copule entre un sujet et un attribut (adjectif, nom ou participe passé). Les verbes au participe passé les plus utilisés avec "être" sont récapitulés dans le tableau 11.

Tableau 11. Les principaux verbes au participe passé avec l'auxiliaire "être" dans le français parlé

compléments	rang	Parlé % participes passés	rang	Ecrit % participes passés	Ecart
faire	1	53,5	4	20,7	+159
venir	2	40,8	1	32,7	+25
arriver	3	34,6	5	20,5	+69
obliger	4	33,9	18	8,1	+316
passer	5	29,9	3	20,7	+44
partir	6	27,9	9	15,6	+79
dire	7	21,2	20	7,3	+189
aller	8	20,3	7	16,7	+22
mettre	9	14,2	8	15,6	-9
rester	10	11,8	10	15,1	-21
rendre	11	11,5	27	4,9	+134
prendre	12	11,2	16	9,2	+21
revenir	13	11,0	14	9,9	+11
donner	14	10,6	23	5,8	+83
rentrer	15	9,8	22	5,9	+66
concerner	16	9,7	707	0,2	+4667
compliquer	17	8,9	74	2,3	+288
ouvrir	18	8,5	25	5,6	+51
finir	19	7,9	19	8,0	0
régler	20	7,5	126	1,4	+447
tomber	21	7,0	12	11,0	-37
satisfaire	22	6,9	119	1,4	+377
élire	23	6,8	380	0,5	+1361
payer	24	6,6	64	2,5	+168
devenir	25	6,5	6	19,0	-66
lier	26	6,5	81	2,0	+231
chauffer	27	6,3	659	0,2	+2590
apercevoir	28	6,1	53	2,9	+107
installer	29	6,1	29	4,8	+27
marier	30	6,1	56	2,7	+126
décevoir	31	6,0	29	4,8	+26
isoler	32	6,0	424	0,4	+1366
poser	33	5,9	54	2,9	+100
intéresser	34	5,5	144	1,2	+350
élever	35	5,5	106	1,6	+237
fermer	36	5,4	33	4,2	+31
habituer	37	5,2	69	2,4	+117
convaincre	38	5,0	92	1,8	+176
sortir	39	5,0	13	10,2	-51
naître	40	5,0	26	5,0	-1
terminer	41	4,5	34	4,1	+10
marquer	42	4,3	122	1,4	+206

prévoir	43	4,3	154	1,2	+268
présenter	44	4,1	90	1,9	+120
retrouver	45	4,0	45	3,3	+23
placer	46	3,9	70	2,4	+62
informer	47	3,8	211	0,9	+320
entendre	48	3,8	52	3,0	+27
mourir	49	3,8	2	24,1	-84
occuper	50	3,7	36	4,0	-6

Lecture : dans le français parlé, "faire" est le verbe le plus utilisé avec l'auxiliaire "être". Il représente 53,5 pour mille participes passés avec cet auxiliaire. A l'écrit, "faire" n'est qu'au quatrième rang, avec 20,7 pour mille participes passés. L'écart entre les deux genres est de +159 % dans le français parlé par rapport à l'écrit.

Ces cinquante premiers verbes ne couvrent que 55 % de tous les emplois d'un participe passé avec "être" dans le français parlé, ce qui montre bien la dispersion considérable du phénomène. Cette brève liste comporte au moins deux singularités.

Premièrement, les verbes usuels ne figurent pas tous dans ce tableau qui comprend ceux qui n'admettent que l'auxiliaire "être" (*venir, arriver, partir, aller...*) ou les verbes transitifs qui permettent les formes passives ou pronominales (*faire, obliger, passer, dire...*).

Deuxièmement, les différences entre les deux registres sont considérables (dernière colonne du tableau). Les suremplois les plus importants signalent les expressions courantes de la conversation. Le modèle est donné par : *c'est fait, c'est obligé, dit, compliqué, réglé, payé, marqué, prévu...* On dit de quelqu'un qu'il *est venu, est allé, concerné, intéressé, impliqué, fatigué, satisfait, déçu*, etc. Dans tous ces cas, le "passé composé" décrit une action ou une situation présente et seul le plus-que-parfait, rare dans la conversation, peut situer le propos dans le passé. De plus, ces participes passés se comportent comme des adjectifs attributs, même s'ils appartiennent bien à la catégorie du verbe.

La présence des thèmes abordés dans les entretiens – "être chauffé", "être isolé", "être élu" permet de comprendre qu'on se situe ici à la limite entre le système de la langue (qui donne les règles d'association d'un participe passé et d'un auxiliaire mais aussi les verbes susceptibles d'entrer dans ces combinaisons) et le sujet de la conversation qui amène à privilégier tel ou tel verbe comme attribut du sujet.

Enfin, les verbes sous-employés sont d'abord certains de ceux qui n'admettent que l'auxiliaire être : *rester, tomber, devenir, sortir, mourir...* sans doute parce que, dans le français écrit, ces verbes sont surtout utilisés au passé et que la conversation répugne à utiliser le passé, surtout le passé simple et, dans une moindre mesure l'imparfait.

5.4 La singularité de "avoir été" dans le français parlé

Dans le tableau 9, on aura remarqué que "été" est le participe passé le plus utilisé avec l'auxiliaire "avoir". Dans le français écrit, la plupart de ces emplois correspondent à une forme passive au passé que nous venons d'examiner dans le paragraphe précédent (*avoir été fait, avoir été dit*, etc.) ou à un adjectif attribut (*avoir été facile*), cas qui sont bien décrits dans les grammaires. L'examen des emplois de "avoir été" dans le français parlé révèle qu'il peut aussi être utilisé comme un synonyme de "allé" (au passé composé) : "j'ai été le voir" au lieu de "je suis allé le voir". Les 6 390 occurrences de "avoir été" se répartissent comme suit :

- 44,1 % sont des formes passives au passé composé ou, plus rarement des plus-que-parfait (il a été cassé, il avait été cassé ; il a été satisfait, déçu, blessé...) ;
- 41,8 % introduisent un attribut qui peut être un adjectif (il a été rapide, efficace, discret...) ou un nom (il a été ouvrier, chef, peintre, étudiant) ;
- 14,1 % sont employés pour le participe passé de "aller" dont 1,8 introduisent un verbe à l'infinitif ("il a été voir" pour "il est allé voir").

Les deux premiers cas de figure montrent que la fonction du passé composé est bien intégrée par la plupart des locuteurs (marquer une situation ou un événement passé ponctuel, contrairement à l'imparfait qui suggère une durée) même s'il est beaucoup plus rare qu'à l'écrit. Dans l'ensemble du corpus du français parlé, un emploi sur sept de "avoir été" survient donc en lieu et place de "être allé", soit 778 occurrences alors qu'il y a 654 participes passés "allé(e,s)". Si l'on ajoute qu'une proportion importante de ces "allé" sont suivis d'un verbe à l'infinitif (examiné plus bas), on se rend compte que la majorité des locuteurs considèrent "avoir été" comme le passé composé (ou le plus-que-parfait) du verbe "aller". Ces locuteurs semblent en être conscients puisqu'ils utilisent aussi beaucoup "aller" au présent, voire au futur (*je vais faire, j'irai faire*). Par exemple, un enquêté, à propos de son chauffage : « ça a été mais maintenant ça va plus du tout ». Parmi les expressions qui reviennent régulièrement : "il a été à l'école, au travail, au lycée, au collège", "été voir quelque chose ou quelqu'un", "été dire quelque chose". On trouve également un grand nombre d'emplois avec le pronom démonstratif : "ça a été trop loin, trop vite, bien, mal..." ; on dit qu'une personne "a été" dans un pays, une région ou une ville ; autre forme courante : "il a été au plus simple", etc.

Répons-le : la majorité des locuteurs emploient avec aisance et naturel ces tournures. Les contextes de ces emplois éclairent cette situation.

Quand "il est allé" est employé, l'action décrite par le passé composé prend sa source avant l'énonciation mais elle peut s'appliquer à la situation présente. "Il était allé" signifie

normalement qu'il ne l'est plus aujourd'hui mais il peut impliquer la répétition de l'action et ne convient donc pas quand on veut parler d'un événement unique situé dans le passé. Le passé simple est abandonné à l'oral et "avoir été allé" est impossible puisque "aller" est intransitif. Dans certains cas, le locuteur conserve "être allé" mais ajoute un adverbe ou une locution pour situer l'action dans le temps (*hier, je suis allé...*). Sans cela, "aller" ne semble pas apte à rendre correctement l'idée d'une action unique, située dans le passé et révolue par rapport à la conversation, contrairement à "avoir été". Enfin et surtout, "j'ai été là" (ou dans tel ou tel lieu) - comme forme passée de "je suis là" - est tout à fait correct. Ce sont des explications possibles de la substitution de "avoir été" à "être allé" dans le français parlé.

En tous cas, même passé sous silence ou condamné par les grammairiens, cet usage est si répandu qu'il semble difficile de l'ignorer. C'est aussi le cas des combinaisons où un verbe - autre que "être" ou "avoir" - joue le rôle d'auxiliaire devant un autre verbe placé à l'infinitif (et non plus au participe passé).

5.5 Les principales combinaisons verbe + verbe infinitif

Les plus employés sont présentés dans le tableau 12. Pour consulter ce tableau, il faut se souvenir que le recensement porte sur les lemmes. Dans ces constructions, le second terme est toujours à l'infinitif. En revanche le premier (le lemme de l'auxiliaire modal) regroupe sous une seule entrée les conjugaisons à toutes les personnes, tous les temps et modes. Enfin, il peut s'intercaler un adverbe ou une préposition (voire plusieurs) entre les deux verbes. Par exemple, "vouloir *bien* faire", "ne vouloir *pas* faire" sont comptés comme "vouloir faire" ; de même "venir *de* dire", est compté comme "venir dire", etc.

Tableau 12. Les syntagmes verbaux les plus fréquents dans les français parlé et écrit

Syntagmes	Parlé		Ecrit		Ecart
	Rang	% Syntagmes	Rang	% Syntagmes	
vouloir dire	1	40,2	3	14,2	+ 182
pouvoir faire	2	30,1	4	14,0	+ 115
aller faire	3	24,7	5	12,5	+ 98
pouvoir dire	4	22,8	6	8,8	+ 158
pouvoir être	5	22,1	2	15,7	+ 40
aller être	6	18,2	13	5,2	+ 249
aller dire	7	17,1	24	3,2	+ 433
devoir être	8	17,1	1	17,6	-3
falloir faire	9	16,2	16	4,0	+ 306
pouvoir avoir	10	15,9	10	5,9	+ 168
aller voter	11	15,5	-	-	-
vouloir faire	12	12,5	12	5,3	+ 135

aller voir	13	11,4	7	6,9	+ 66
aller avoir	14	10,2	43	2,2	+ 370
devoir avoir	15	8,2	8	6,6	+ 26
devoir faire	16	7,8	9	6,0	+ 30
falloir être	17	7,5	39	2,3	+ 222
pouvoir mettre	18	6,5	60	1,7	+ 281
aller aller	19	6,5	31	2,7	+ 142
entendre parler	20	6,3	23	3,4	+ 88
aller mettre	21	5,7	45	2,0	+ 186
pouvoir aller	22	5,6	21	3,4	+ 67
aller chercher	23	5,5	11	5,6	-2
aller passer	24	4,9	28	3,0	+ 66
falloir aller	25	4,9	85	1,3	+ 287
falloir avoir	26	4,4	148	0,9	+ 396
falloir dire	27	4,8	38	2,4	+ 102
falloir mettre	28	4,5	166	0,8	+ 464
falloir voir	29	4,3	81	1,3	+ 225
aller prendre	30	3,9	22	3,4	+ 16
savoir dire	31	3,7	68	1,6	+ 136
pouvoir donner	32	3,6	42	2,2	+ 64
pouvoir parler	33	3,6	36	2,4	+ 47
pouvoir passer	34	3,6	29	2,9	+ 24
aller donner	35	3,5	70	1,5	+ 141
venir voir	36	3,5	19	3,8	-10
savoir faire	37	3,4	26	3,0	+ 15
vouloir être	38	3,3	27	3,0	+ 11
pouvoir arriver	39	3,2	62	1,7	+ 95
vouloir avoir	40	3,2	94	1,2	+ 162
falloir savoir	41	3,2	120	1,0	+ 209
aller travailler	42	3,1	242	0,6	+ 441
faire venir	43	3,0	73	1,4	+ 105
aller changer	44	2,9	237	0,6	+ 389
pouvoir payer	45	2,9	307	0,5	+ 507
aller arriver	46	2,8	101	1,2	+ 137
vouloir aller	47	2,8	69	1,5	+ 84
aller essayer	48	2,8	113	1,1	+ 159
faire faire	49	2,7	141	0,9	+ 196
aller pouvoir	50	2,6	59	1,7	+ 54

Lecture : le syntagme "vouloir dire" est le plus employé dans le français parlé et le troisième à l'écrit. Il fournit 40,2 pour 1000 de la surface du texte couvert par ces modalités contre 14,2 à l'écrit, soit un écart net de +182 %

La quasi-totalité des écarts sont positifs et de grande ampleur. Il faut se souvenir qu'il s'agit des écarts nets (une fois neutralisée la propension plus forte à utiliser des verbes à l'oral). Autrement dit, non seulement la conversation utilise un peu plus ce genre de constructions (tableau 6 ci-dessus), mais il y a une concentration plus forte sur le haut du tableau (les verbes les plus usuels).

La présence de "aller voter" dans cette brève liste rappelle à nouveau que, malgré la grande dimension des deux corpus, le poids de certains thèmes – comme la politique pour le corpus du français parlé – se fait sentir dans ces calculs.

Cependant, les 52 premiers ne couvrent que 42 % de toutes les combinaisons. Les 50 % sont atteints à la 86^e ; les 90 % au 1 973^e (d'effectifs 3). Au-delà des quelques syntagmes largement partagés (qui figurent en tête du tableau), on assiste donc à une très grande diversité. Toutefois, on remarque que les mêmes modalisateurs reviennent régulièrement dans le tableau : *pouvoir*, *devoir*, *falloir*, *aller*. Le recensement des principaux verbes en position d'auxiliaires modaux permet de vérifier cette intuition.

Le tableau 13 compare ces auxiliaires dans les deux corpus. Dans ce tableau, la densité de chaque auxiliaire est calculée en rapportant le nombre de fois où le verbe apparaît dans cette position à l'effectif total des constructions modales rencontrées dans le corpus sous revue (colonnes % Modalisations). Par exemple, dans le français parlé, sur 1 000 occurrences d'une construction modale, 250,8 – soit le quart – sont construites avec l'auxiliaire "aller" alors qu'à l'écrit, cette proportion n'est que de 162,5 ‰. Dans la colonne suivante, on rapporte le nombre de fois où ce verbe joue un rôle d'auxiliaire à son effectif total (% Effectif). Par exemple, sur 100 occurrences de "aller", dans le corpus du français parlé, les deux tiers (65,9 %) sont en position d'auxiliaire modal alors que, dans le corpus écrit, cette proportion est de six sur dix. L'écart de 54,3 % indique qu'il y a une chance infinitésimale de se tromper en affirmant que l'expression orale spontanée privilégie l'auxiliaire modal "aller" (ce qui avait déjà été montré précédemment) et que cet auxiliaire permet de se situer dans le futur proche tout en utilisant le présent (*je vais faire*).

Tableau 13. Les auxiliaires modaux de la conversation comparés à ceux de l'écrit

Modalisateur	Parlé			Ecrit			Ecart
	Rang	% Mod.	% Effectifs	Rang	% Mod.	% Effectifs	
aller	1	250,8	65,9	2	162,5	60,1	+ 54,3
pouvoir	2	245,1	85,0	1	226,8	90,1	+ 8,1
falloir	3	134,6	58,5	6	57,5	56,3	+ 134,1
vouloir	4	114,2	55,5	5	93,2	60,5	+ 22,6
devoir	5	75,9	88,4	3	110,9	91,4	-31,5
faire	6	62,2	9,3	4	97,0	21,0	-35,9
venir	7	18,9	23,1	8	32,1	23,5	-41,2
savoir	8	18,4	6,5	9	22,9	12,1	-19,6
être	9	13,5	0,4	16	7,3	0,3	+ 85,8
aimer	10	10,7	30,7	14	10,0	18,6	+ 6,9
entendre	11	8,0	22,1	12	12,6	19,0	-36,1
laisser	12	7,1	25,0	7	35,1	45,1	-79,8
préférer	13	5,3	52,3	18	6,4	50,7	-17,0
voir	14	4,8	2,2	11	19,1	9,2	-74,9
penser	15	2,8	1,6	25	2,4	42,0	+19,5

Lecture : l'auxiliaire "aller" est le plus fréquemment utilisé dans le français parlé. Il représente 250,8 pour mille utilisations d'un auxiliaire modal et 62,9 % des occurrences de "aller". A l'écrit, cet auxiliaire vient au deuxième rang avec 162,5 pour mille utilisations d'un auxiliaire modal et 60 % des occurrences de "aller". La densité de "aller" comme auxiliaire modal est de 54,3 % plus élevée dans le français parlé comparé à l'écrit.

A eux seuls, les deux premiers – *aller* et *pouvoir* - fournissent la moitié des cas de modalisation. La proportion des trois quarts est atteinte avec le quatrième ("vouloir"), et celle des neuf-dixièmes avec "venir". Le tableau couvre 98 % du total des combinaisons modales. Dans la pratique, il suffit donc de connaître une quinzaine de verbes pour maîtriser la modalisation verbale.

Comme l'avait laissé pressentir l'analyse des verbes seuls au début de ce chapitre, la conversation privilégie surtout "falloir", "aller", "vouloir" et "ce+être" et se concentre sur un nombre plus limité de modalisateurs que l'écrit. Pour les combinaisons les plus fréquentes, l'écrit met davantage l'accent sur "devoir", "venir" ou "faire".

On aura certainement remarqué que le phénomène présente une forte parenté avec les temps composés à l'aide des auxiliaires "avoir" ou "être" ou avec la forme passive ou le réfléchi (auxiliaire "être"). Dans tous les cas, un verbe auxiliaire est associé à un verbe complément et lui donne une dimension (temporelle ou modale).

Tableau synthétique de l'auxiliation en français

En plus des auxiliaires "purs" ("avoir" et "être"), une quinzaine d'auxiliaires expriment ces dimensions que nous nommons "modales". En y ajoutant quelques quasi-synonymes, on obtient un tableau assez complet des modalisations possibles par ordre d'importance :

1. Avec le verbe modalisé au participe passé :

- le passé : avoir, être (pour quelques verbes n'admettant pas "avoir") ;
- le passé, le passif et le réfléchi : être ;

2. Avec le verbe modalisé à l'infinitif.

- le futur immédiat (par rapport à l'instance de l'énonciation) : aller, compter, prévoir ;
- la possibilité ou la capacité : pouvoir ;
- nécessité : falloir ;
- volonté : vouloir, prétendre, souhaiter, désirer ;
- obligation (ou forte probabilité) devoir ;
- action : faire, dire, laisser, entendre, voir, oser ;
- passé immédiat : venir (de), faillir ;
- déplacement : venir, revenir, retourner, passer, monter, descendre ;
- la connaissance : savoir ;
- pensée et le sentiment : penser, croire, aimer, préférer, détester, souhaiter, espérer ;

En y ajoutant la possibilité d'une liaison entre l'auxiliaire et le verbe assurée par une préposition (*proposer de faire, venir pour voir, finir par dire, parvenir à faire...*), on obtient un tableau complet du phénomène. Certains verbes admettent les deux constructions avec des sens plus ou moins différents. Par exemple "venir" immédiatement suivi du verbe modalisé à l'infinitif exprime un déplacement dans l'espace - *venir faire, revenir voir*, etc. – l'association avec les prépositions "à" et "pour" implique un processus (*en venir à faire, venir pour voir...*) alors que l'association avec la préposition "de" (*venir de faire*) donne un passé immédiat (symétrique au futur immédiat construit avec "aller faire") dont nous avons déjà dit les intérêts, notamment la possibilité de s'épargner des conjugaisons complexes et la concordance des temps. Enfin, à part certaines exceptions, il semble que pratiquement tous les verbes peuvent se trouver en seconde position sous la forme d'un participe passé ou d'un infinitif. Le système est donc assez simple et peut être facilement enseigné avec quelques exemples bien choisis.

Chapitre 6.

Les satellites du verbe

Dans la première partie de ce rapport, il a été montré que trois autres catégories grammaticales sont liées au verbe et suivent ses variations, souvent en les amplifiant : les pronoms, les adverbes et les conjonctions de subordination.

6.1 Les pronoms

Comme les verbes auxquels ils sont associés, les pronoms sont beaucoup plus utilisés dans le français parlé qu'à l'écrit. La hausse est de 13 % pour les pronoms personnels. Ceux-ci sont logiquement liés aux verbes et suivent leurs fluctuations alors que les autres pronoms – que l'on peut qualifier d'"impersonnels" - connaissent, en moyenne, un suremploi nettement plus considérable (+76 %). C'est particulièrement le cas des relatifs et spécialement de "que". Pour tenir compte de cette différence, l'analyse portera successivement sur les pronoms personnels puis sur les autres avec des coefficients de pondération calculés sur ces deux catégories (tableau 14). Les vocables présents dans ces tableaux couvrent la quasi-totalité des emplois d'un pronom dans la conversation (taux de couverture supérieur à 99 %).

Tableau 14. Les pronoms personnels et impersonnels dans les français écrit et parlé (fréquences pour mille mots)

Vocable	Parlé (A)			Ecrit (B)			(A-B)/B %
	Rang	Fréquence (‰)	‰ Pronoms	Rang	Fréquence (‰)	‰ Pronoms	
Pronoms personnels							
je	1	33,15	317,8	2	23,9	258,6	+23
on	3	21,24	203,6	9	4,9	52,7	+286
il	4	18,95	181,7	1	27,1	293,2	-38
ils	8	8,22	78,8	12	3,8	40,9	+93
se	10	6,12	58,6	3	11,8	127,9	-54
vous	12	5,06	48,5	8	5,8	62,8	-23
moi	13	4,72	45,3	17	2,1	22,9	+98
nous	15	3,4	32,5	14	3,2	34,4	-5
tu	18	1,56	15,0	10	4,5	48,4	-69
lui	21	1,26	12,1	11	4,2	45,0	-73
nous autre	32	0,15	1,5	51	0,0	0,2	+726
toi	36	0,1	0,9	26	0,5	5,9	-84
lui-même	37	0,07	0,7	30	0,3	3,6	-81
soi	39	0,06	0,6	41	0,1	0,7	-12
moi-même	40	0,056	0,5	36	0,1	1,2	-53
eux-mêmes	41	0,055	0,5	43	0,1	0,7	-19
nous-même	47	0,024	0,2	52	0,0	0,2	+33
soi-même	48	0,021	0,2	47	0,0	0,3	-33
vous-même	50	0,003	0,0	48	0,0	0,3	-89

Autres pronoms							
ce	2	27,14	294,4	4	9,7	185,0	+59
ça	5	15,81	171,5	15	3,0	57,4	+199
qui	6	10,93	118,6	6	7,6	145,9	-19
y	7	9,6	104,1	13	3,4	64,1	62
que	9	6,28	68,1	7	6,2	117,7	-42
le	11	6,11	66,3	5	8,0	152,1	-56
en	14	3,92	42,5	16	2,9	54,9	-23
quoi	16	2,68	29,1	25	0,7	14,1	+107
tout	17	2,5	27,1	18	1,7	33,2	-18
rien	19	1,39	15,1	19	1,7	32,6	-54
autre	20	1,27	13,8	20	1,1	21,3	-35
un	22	0,86	9,4	22	0,9	17,9	-48
leur	23	0,74	8,1	29	0,4	7,6	+6
celui	24	0,63	6,8	21	1,0	19,7	-66
quelqu'un	25	0,45	4,9	31	0,3	5,8	-16
chacun	26	0,25	2,7	33	0,2	3,5	-23
lequel	27	0,21	2,2	27	0,5	8,9	-75
personne	28	0,19	2,1	28	0,4	8,5	-75
cela	29	0,17	1,9	23	0,9	16,2	-88
celui-là	30	0,17	1,9	34	0,1	2,2	-15
dont	31	0,17	1,8	24	0,8	14,8	-88
même	33	0,15	1,6	40	0,1	1,3	+30
certain	34	0,14	1,5	39	0,1	1,3	+15
quel	35	0,12	1,3	38	0,1	1,6	-19

Lecture : voir sous le tableau 5.

Le pronom "on" est le plus caractéristique de la conversation. L'univers lexical de ce pronom montre un large éventail d'emplois. Le plus souvent, il signifie "une personne ou des gens que je ne nomme pas" et il se substitue à "il" et "ils". Parfois, le locuteur ne nomme pas ces autres parce qu'il ne connaît pas leur identité (ou que c'est un groupe flou) mais souvent c'est une personne (ou un groupe) qu'on ne veut pas nommer car son comportement est réprouvé par celui qui parle. Dans ce cas, les constructions les plus fréquentes sont : "on me..." ou "on nous..." suivis d'un verbe ("on me fait", "on me dit", "on veut me faire"). Enfin, "on" peut être un "nous" familier, dans lequel le locuteur s'inclue. Dans ce cas, il n'y a généralement pas de pronom relatif entre "on" et le verbe. Ceci explique pourquoi le pronom "nous" semble stable alors que "nous-même" et "nous autres"²² ont des densités d'emploi en nette augmentation (puisque'il n'existe pas de "on-même").

Par contre-coup, la première personne du singulier semble moins nettement privilégiée dans la conversation. Certes, la nature particulière de l'enquête sociologique peut expliquer que l'entretien ne se centre pas sur l'enquêté lui-même mais sur les sujets spécifiques de ces enquêtes. Cependant, l'une des caractéristiques de l'oral spontané apparaît clairement :

²²Pour "nous autres", la présence d'un corpus québécois explique en partie la densité exceptionnelle de ce pronom à l'oral.

l'importance du "moi" dont le sens principal est : "pour ma part", bien plus souvent qu'un simple renforcement de la première personne ("moi je").

L'enquêteur n'étant pas un familier, il est impossible de le tutoyer, d'où une densité du "tu" des deux tiers inférieure à celle de l'écrit ; "toi" et "toi-même" connaissent des mouvements similaires. Le caractère particulier de l'entretien sociologique peut également expliquer la faible densité du "vous", l'enquêteur s'effaçant autant que possible en face de l'enquêté. Il semble donc difficile d'en conclure que les secondes personnes sont peu employées dans la conversation.

En revanche, les reculs considérables de la troisième personne du singulier (*il, se, lui, lui-même, soi et soi-même*) ne peuvent être rattachés aux conditions particulières de l'enquête sociologique. L'examen des contextes d'emploi de ces pronoms montre que l'impersonnel n'est pas en cause : il est au moins autant utilisé dans le français parlé que dans le français écrit (en particulier "il faut"). En revanche, dans l'usage anaphorique - pour désigner une personne extérieure à l'interlocution ou une chose déjà nommée - la conversation semble préférer la répétition du nom ou le démonstratif "ce" ou encore le "on". En revanche, on parle plus aisément de plusieurs personnes ou des "gens" avec le pronom personnel de la troisième personne du pluriel.

Le français parlé privilégie le démonstratif – et en premier lieu "ce" et "ça", placés la plupart du temps en position de sujet et non de complément ("c'est", "ça fait", "ça peut être", "ça ne fait rien", etc.)

Le "quoi" est avant tout une interjection placée au début d'une intervention – en question rhétorique ou en interjection pour souligner le désaccord ou l'incompréhension – ou en clôture d'une période en manière de point d'exclamation.

Parmi les pronoms délaissés par le français parlé : "lequel" (et ses flexions : lequel(les), duquel, desquel(les) – recul de 75 % - mais aussi "dont" – recul net de 88 % - ou encore les pronoms composés avec "même" (moi-même, etc.) sont rares à l'écrit et disparaissent quasiment dans le français parlé. Le principe d'économie qui régit la communication orale conduit à sacrifier ces constructions trop complexes. D'autres comme "ceci" et "cela" sont remplacés par "ça", etc.

Autre constat surprenant : les pronoms possessifs – mien, tien, sien, nôtre, vôtre – sont absents du tableau car ils sont rares à l'écrit et plus encore dans le français parlé. Seul surnage "leur" mais en fait, il s'agit du pronom relatif : "je leur dis" et non pas "le leur" ou "les leurs" qui ne sont pratiquement jamais employés dans le français parlé (on les rencontre en tout une trentaine de fois contre 2 667 emplois du relatif).

Enfin, les pronoms vont souvent par couple, ainsi le "je" et le "moi" ; "il" et "lui" ; "ils" et "leur" ; "ce" et "ça", etc. Comme l'indique le tableau 5, les membres de ces paradigmes évoluent dans le même sens.

6.2 Les adverbes

Adverbe signifie "adjoint au mot" et ce mot peut appartenir au groupe du verbe comme à celui du nom. Adjoint au verbe, il se place après celui-ci : adjoint à un adjectif il se place avant. De fait, la majorité des usages se trouvent dans le groupe verbal. Dans tous les corpus, la densité d'emploi des adverbes suit celle des verbes en l'amplifiant. Dans le français parlé, ils sont donc beaucoup plus utilisés qu'à l'écrit. De plus, les emplois de l'oral sont assez différents de ceux de l'écrit (tableau 15).

Tableau 15. Les adverbes usuels en français parlé et écrit (fréquences pour mille mots).

Vocable	Parlé (A)			Ecrit (B)			(A-B)/B %
	Rang	Fréquence (%)	% Adverbes	Rang	Fréquence (%)	% Adverbes	
pas	1	20,6	170,6	2	9,8	127,6	+34
ne	2	12,2	101,3	1	13,8	180,7	-44
plus	3	6,2	51,1	3	5,0	65,3	-22
oui	4	6,0	49,4	18	0,8	10,3	379
là	5	5,5	45,3	8	1,5	19,6	131
non	6	5,3	44,0	12	1,3	17,0	159
bon	7	4,6	38,0	56	0,2	2,6	1336
bien	8	4,2	34,6	4	2,6	34,4	0
même	9	3,6	29,9	10	1,4	18,2	64
peu	10	3,1	25,7	7	1,6	20,4	26
alors	11	3,1	25,5	15	1,1	14,3	78
enfin	12	2,9	23,7	28	0,5	6,9	245
puis	13	2,8	23,3	14	1,2	15,3	52
très	14	2,6	21,5	13	1,2	15,9	35
aussi	15	2,2	18,1	11	1,3	17,4	4
peut-être	16	1,8	15,2	21	0,8	10,1	51
beaucoup	17	1,7	14,3	29	0,5	6,5	119
où	18	1,5	12,8	5	1,8	24,0	-47
toujours	19	1,4	11,9	17	1,1	13,8	-14
tout	20	1,4	11,7	9	1,5	19,1	-39
vraiment	21	1,4	11,7	41	0,4	4,7	148
trop	22	1,3	10,5	20	0,8	10,1	4
moins	23	1,3	10,5	22	0,8	10,0	5
maintenant	24	1,2	9,6	26	0,6	8,0	21
comment	25	1,1	8,8	27	0,5	7,0	26
déjà	26	1,1	8,8	19	0,8	10,2	-14
encore	27	1,0	8,3	6	1,6	20,5	-60
jamais	28	0,8	6,9	16	1,1	13,9	-50
ici	29	0,8	6,7	24	0,6	8,3	-19
assez	30	0,8	6,4	31	0,5	5,9	8

mal	31	0,6	5,3	35	0,4	5,5	-5
surtout	32	0,5	4,3	45	0,3	4,0	8
pourquoi	33	0,5	4,3	25	0,6	8,1	-47
mieux	34	0,5	4,3	33	0,4	5,8	-26
ailleurs	35	0,5	3,9	37	0,4	5,1	-22
souvent	36	0,4	3,7	49	0,2	3,1	20
justement	37	0,4	3,5	85	0,1	1,2	191
plutôt	38	0,4	3,2	48	0,3	3,3	-3
tellement	39	0,4	3,2	55	0,2	2,7	19
tant	40	0,3	2,9	40	0,4	4,8	-40
effectivement	41	0,3	2,9	180	0,0	0,2	1053
d'accord	42	0,3	2,8	77	0,1	1,4	99
aujourd'hui	43	0,3	2,6	51	0,2	2,9	-11
finalement	44	0,3	2,5	95	0,1	1,0	143
si	45	0,3	2,5	23	0,8	9,9	-75
loin	46	0,3	2,2	32	0,4	5,8	-62
forcément	47	0,3	2,1	125	0,0	0,6	258
actuellement	48	0,2	2,0	235	0,0	0,1	1228
complètement	48	0,2	2,0	80	0,1	1,3	52
là-dessus	50	0,2	1,8	128	0,0	0,6	212

Lecture : voir sous tableau 5.

A part "bien", "aussi", "trop", "moins", "mal" et "plutôt", tous les écarts sont significatifs.

Rappelons que, dans le français parlé, la densité moyenne des adverbes augmente de 58 % par rapport à l'écrit (tableau 2). Pour interpréter les fréquences, c'est cette moyenne qu'il faut garder en tête. On la neutralise pour calculer les écarts nets présentés en dernière colonne.

Ainsi, en comparant le pourcentage des adverbes fourni par "ne" dans les deux registres, on peut considérer qu'il en manque près de la moitié dans le français parlé par rapport à l'écrit. (Encore faut-il se souvenir que les secrétaires ont parfois corrigé les transcriptions). Si l'on se souvient que "ne" est nécessairement suivi d'un autre adverbe (essentiellement "pas" ou "plus"), c'est alors plus de six "ne" sur dix qui seraient manquants.

"Je veux **pas**" est presque la règle et "je **ne** veux **pas** (ou **plus**)" l'exception. Quoi de plus logique au fond puisque la place normale de l'adverbe est derrière le verbe auquel il est adjoind ? De plus, "ne" ne vient jamais seul et doit être complété par un autre adverbe de négation après le verbe. Certes, ce redoublement a de nombreuses raisons d'être mais le fait demeure que la majorité des locuteurs du français ne sont pas spontanément sensibles à ces raisons.

Autre remarque : certains des écarts les plus élevés (en dernière colonne) sont liés à la situation d'énonciation particulière à la conversation. Les exclamations "bon !" et "là" (ou "oh là !") ainsi que les "oui" et "non" introduisent les réponses des enquêtés ou relient des éléments de récits ou de raisonnement. "Alors", "enfin", "puis" fonctionnent de la même façon pour enchaîner les épisodes du récit. De même, "là", "là-dessus" (dans le français parlé, comme synonyme de "et alors", "ensuite") "justement", "finalement", "effectivement" et "forcément" permettent également d'articuler le raisonnement et d'amener les conclusions.

Enfin, ces mots usuels forment système et leur densité varie de conserve : oui et non ; bon, bien et mal ; toujours et jamais ; aujourd’hui, hier et demain ; dessus et dessous (avec les composés : par-dessus, par-dessous, là-dessus, en-dessus, par-dessus...). Chacun de ces termes est plus facilement intelligible quand il est présenté avec les autres vocables constituant son paradigme. Deux remarques en conclusion sur ce point.

Premièrement, la liste des adverbes français est très longue mais les cinquante listés dans le tableau 15 fournissent 92 % du total des occurrences des adverbes dans le français parlé. A l’écrit, les cinquante premiers en représentent 84 % et ils figurent presque tous dans la liste du tableau 15. Il y a donc une forte concentration sur ces quelques outils. Pour la plupart des formes rares, il s’agit d’adverbes en "-ment". Ils sont facilement compris une fois connue la règle de composition des adverbes à l’aide de ce suffixe, ajouté au radical d’un adjectif.

Deuxièmement, l’emploi relativement rare du "ne" semble être l’une des caractéristiques les plus remarquables du français parlé. La majorité des locuteurs utilisent rarement "ne" et seulement pour mettre une emphase dans leur refus. Le débutant en FLE devrait en être averti – en même temps qu’on lui présente les principaux verbes - ainsi que de la double prononciation de "plus" : la négation se prononce "plu" (dans "plus du tout") ; et "plusse" si l’adverbe signifie "davantage" ("en vouloir plus").

6.3 Les conjonctions de subordination

Enfin, dernier élément du groupe verbal privilégié dans la conversation : les conjonctions de subordination. Il faut rappeler que, dans le français parlé, leur densité est considérable par rapport à l’écrit (tableau 2 : + 79 %). Bien que les grammaires en donnent des listes plus longues, dans les faits, il n’y en a que dix (récapitulées dans le tableau 16).

Tableau 16. Les conjonctions de subordination dans les français parlé et écrit

Vocables	Parlé		Ecrit		Ecart %
	Rang	% Subordination	Rang	% Subordination	
que	1	484,0	1	576,5	-16
parce que	2	156,3	5	31,2	401
quand	3	131,3	4	67,9	93
Si	4	116,4	3	116,3	0
comme	5	85,7	2	167,6	-49
puisque	6	11,8	8	7,3	62
sinon	7	8,2	9	3,5	136
lorsque	8	2,8	6	20,5	-86
tandis que	9	2,6	7	8,1	-68
quoique	10	1,0	10	1,2	-14

Les deux premières fournissent plus de six emplois sur dix dans les deux registres. Dans le français parlé, les 90 % sont atteints au 4^e (au cinquième dans le français écrit). Autant dire que, à partir de *comme*, l'emploi des six dernières est très marginal.

Outre l'incontournable "que" qui introduit près de la moitié des subordonnées dans la conversation, les subordinations les plus caractéristiques de ce genre sont "parce que" (cinq fois plus employé dans la conversation qu'à l'écrit), "quand" (multiplication par 1,9) et, plus marginalement : "puisque". Ces trois conjonctions assurent l'enchaînement logique de l'argumentation. En quelque sorte, ce sont des échafaudages qui sont en partie gommés à l'écrit. De plus, dans la conversation, ces trois conjonctions en remplacent plusieurs encore assez utilisées à l'écrit, notamment "comme", "tandis que" et "lorsque". Enfin, l'étude des contextes montre que les usages de la subordination en français parlé sont souvent assez éloignés de l'image qu'en donnent les grammaires (voir Labbé & Labbé 2013).

Partie III.

Le groupe du nom

Alors que le groupe du verbe décrit des actions et les pose dans le temps, l'espace et, éventuellement, dans différentes dimensions modales, le groupe du nom porte sur les objets, les idées ou les personnes (les thèmes). Il apporte aussi des précisions sur la qualité des individus et des choses dont on parle. Les noms sont au cœur de ce deuxième groupe.

Il faut distinguer les noms propres (à majuscule initiale) et les noms communs. Contrairement aux seconds, les premiers peuvent être utilisés seuls sans déterminant. Il s'agit essentiellement des patronymes, toponymes, noms de peuples et sigles à l'exception des faux noms propres auxquels l'usage met parfois une majuscule initiale - comme Etat, Constitution, Parlement - qui sont traités ici comme des noms communs. La densité des noms propres dans un texte n'a généralement pas de liaison statistique avec les autres catégories grammaticales. En effet, les noms propres n'appartiennent pas vraiment à la langue mais constituent plutôt des interfaces entre les propos et la réalité extérieure au discours. Ils assurent la même fonction que les dates et les chiffres : ils ancrent la conversation dans le temps, l'espace (géographique ou social), dans la situation d'énonciation et ils donnent une matérialité au discours. C'est pour cela que le français parlé en fait un usage plus important que le français écrit (tableau 2).

La présente analyse ne porte donc que sur les noms communs et leurs satellites : les adjectifs, déterminants et prépositions.

Chapitre 7.

Les noms communs

Ce chapitre porte sur les éléments que la syntaxe et le lexique français étiquettent comme "noms" (parfois "substantifs"). Les noms désignent des choses, des êtres, des entités, des concepts, c'est-à-dire les thèmes du propos. Comme les transcriptions que nous utilisons ici portent sur un nombre de thèmes limités – spécialement, la vie quotidienne, le confort de la maison, la politique, le couple, la précarité et les relations professionnelles –, on s'attend donc à ce que ces thèmes ressortent particulièrement dans cette partie du vocabulaire. Cependant, le recensement des noms fait également apparaître certaines caractéristiques singulières du français parlé.

La grammaire contemporaine semble s'être accordée sur deux critères formels qui, en cas de doute, permettent de reconnaître un nom. Premièrement, c'est le noyau du groupe nominal, le seul élément indispensable pour que ce groupe existe, ou encore ce qu'il reste quand ont été retirés de ce groupe tous les éléments facultatifs : déterminants, prépositions, qualificatifs (adverbes et adjectifs) et conjonctions de coordination. Deuxièmement, « le nom est nécessairement déterminé » (Arrivé & Al, 1986, p. 404). En pratique, cela signifie qu'il est précédé d'un article, d'un numéral ou d'un "adjectif" (qualificatif, démonstratif, possessif, indéfini). Comme on le verra dans ce chapitre, ce deuxième critère n'est pas toujours vérifié surtout dans le français parlé.

Comme indiqué en introduction, il s'agit ici des noms communs à l'exclusion des mots à majuscule initiale.

7.1 Les noms communs du français parlé

Comme pour le groupe du verbe, le français parlé est le point de repère. Le tableau 17 ne liste donc pas les substantifs les plus fréquents de l'écrit qui sont assez différents de ceux de l'oral. Ainsi "homme" – qui est le substantif le plus fréquent à l'écrit – n'apparaît qu'au 36^e rang dans le français parlé avec une densité d'utilisation divisée par deux. De même, "œil" et "main", qui arrivent en seconde et troisième positions à l'écrit, n'apparaissent pas dans ce tableau (mais ils figurent bien dans la liste des noms usuels donnée en annexe 3).

Tableau 17. Les principaux noms du français parlé comparés à leur utilisation à l'écrit (fréquence pour mille mots)

Vocabulaire	Parlé			Écrit			Ecart %
	Rang	Fréquence (%)	% Substantifs	Rang	Fréquence (%)	% Substantifs	
chose	1	2,75	21,13	5	1,17	6,63	+ 219
gens	2	2,44	18,75	42	0,40	2,28	+ 721
problème	3	1,74	13,37	227	0,13	0,72	+ 1753
heure	4	1,49	11,46	6	1,17	6,62	+ 73
fait	5	1,48	11,38	55	0,36	2,02	+ 464
temps	6	1,41	10,87	4	1,17	6,67	+ 63
an	7	1,32	10,15	21	0,63	3,58	+ 183
fois	8	1,24	9,55	7	1,14	6,45	+ 48
niveau	9	1,17	8,99	798	0,04	0,25	+ 3482
jour	10	1,09	8,41	8	1,13	6,42	+ 31
chauffage	11	1,04	7,96	3114	0,01	0,04	+ 18352
façon	12	0,99	7,63	96	0,26	1,48	+ 416
travail	13	0,93	7,17	112	0,24	1,33	+ 438
moment	14	0,93	7,15	14	0,75	4,24	+ 69
côté	15	0,83	6,36	20	0,64	3,62	+ 76
femme	16	0,82	6,31	9	1,02	5,76	+ 9
année	17	0,82	6,29	50	0,36	2,06	+ 205
personne	18	0,81	6,23	154	0,19	1,05	+ 494
mois	19	0,80	6,11	81	0,30	1,68	+ 264
enfant	20	0,76	5,85	24	0,58	3,28	+ 79
maison	21	0,75	5,77	27	0,52	2,92	+ 97
monde	22	0,73	5,61	17	0,70	3,98	+ 41
truc	23	0,72	5,50	215	0,13	0,75	+ 633
exemple	24	0,70	5,35	208	0,14	0,77	+ 599
négociation	25	0,69	5,31	2675	0,01	0,05	+ 9682
rapport	26	0,68	5,20	196	0,14	0,81	+ 538
tout	27	0,67	5,19	148	0,19	1,08	+ 382
franc	28	0,63	4,81	324	0,10	0,55	+ 774
question	29	0,60	4,59	34	0,43	2,46	+ 87
vie	30	0,60	4,58	13	0,76	4,33	+ 6
partie	31	0,58	4,46	138	0,20	1,15	+ 286
gauche	32	0,57	4,41	232	0,13	0,71	+ 520
politique	33	0,54	4,15	729	0,05	0,27	+ 1416
idée	34	0,53	4,06	52	0,36	2,04+	+ 99
droite	35	0,52	3,99	245	0,12	0,69	+ 481
homme	36	0,51	3,92	1	1,46	8,27	- 53
compte	37	0,51	3,90	84	0,28	1,59	+ 145
besoin	38	0,50	3,87	67	0,33	1,87	+ 107
part	39	0,50	3,85	51	0,36	2,05	+ 87
point	40	0,50	3,81	53	0,36	2,03	+ 88
système	41	0,49	3,76	538	0,06	0,36	+ 934
client	42	0,48	3,69	302	0,10	0,59	+ 527
électricité	43	0,46	3,55	1818	0,02	0,09	+ 3685
cas	44	0,41	3,14	80	0,30	1,71	+ 83
argent	45	0,41	3,11	142	0,20	1,11	+ 182
entreprise	46	0,40	3,08	640	0,05	0,31	+ 902
monsieur	47	0,39	3,02	12	0,80	4,53	-33
solution	48	0,39	2,96	710	0,05	0,28	+ 955
pays	49	0,38	2,95	98	0,26	1,47	+ 101
coup	50	0,38	2,94	11	0,90	5,11	- 42
raison	51	0,37	2,86	57	0,35	1,97	+ 45

mari	52	0,37	2,85	216	0,13	0,75	+ 280
envie	53	0,37	2,80	114	0,23	1,31	+ 115
parti	54	0,36	2,79	541	0,06	0,36	+ 674
place	55	0,36	2,79	33	0,45	2,55	+ 9
facture	56	0,35	2,65	2553	0,01	0,06	+ 4457
suite	57	0,34	2,63	64	0,33	1,87	+ 40
syndicat	58	0,34	2,59	2187	0,01	0,07	+ 3410
semaine	59	0,34	2,58	140	0,20	1,14	+ 127
fin	60	0,33	2,54	93	0,26	1,50	+ 70
gaz	60	0,33	2,54	1019	0,03	0,19	+ 1242
radiateur	62	0,33	2,53	3145	0,01	0,04	+ 5883
ligne	63	0,33	2,50	181	0,16	0,90	+ 178
soir	64	0,32	2,48	29	0,50	2,83	-12
avis	65	0,32	2,44	355	0,09	0,51	+ 375
impression	66	0,31	2,38	136	0,20	1,16	+ 106
matin	67	0,31	2,37	47	0,39	2,19	+ 8
famille	68	0,31	2,35	109	0,24	1,35	+ 74
chambre	69	0,31	2,34	41	0,40	2,29	+ 3
journée	70	0,30	2,31	168	0,17	0,95	+ 142
température	71	0,29	2,23	1451	0,02	0,13	+ 1671
début	72	0,29	2,21	182	0,16	0,90	+ 146
chaleur	73	0,29	2,20	262	0,11	0,65	+ 241
eau	74	0,28	2,15	39	0,41	2,34	-8
droit	75	0,28	2,15	189	0,15	0,88	+ 145
moyen	77	0,28	2,12	200	0,14	0,80	+ 165
hiver	78	0,27	2,11	481	0,07	0,39	+ 435
installation	79	0,27	2,10	2289	0,01	0,07	+ 2928
formation	80	0,27	2,10	1730	0,02	0,10	+ 2016
intérêt	81	0,27	2,09	342	0,09	0,53	+ 298
sens	82	0,26	2,03	129	0,21	1,21	+ 67
départ	83	0,26	1,99	272	0,11	0,63	+ 217
pièce	84	0,26	1,99	87	0,28	1,56	+ 27
comité	85	0,25	1,94	814	0,04	0,24	+ 691
jeune (n m)	85	0,25	1,94	1158	0,03	0,17	+ 1070
service	87	0,25	1,93	161	0,17	0,99	+ 95
emploi	88	0,25	1,92	976	0,04	0,20	+ 860
changement	89	0,25	1,90	1092	0,03	0,18	+ 976
appareil	90	0,24	1,87	420	0,08	0,44	+ 326
nuît	91	0,24	1,86	16	0,71	4,01	- 54
vue	91	0,24	1,86	190	0,15	0,87	+ 115
candidat	93	0,24	1,85	2690	0,01	0,05	+ 3332
école	94	0,24	1,84	304	0,10	0,58	+ 215
pouvoir	95	0,24	1,81	399	0,08	0,46	+ 296
parent	96	0,24	1,80	233	0,13	0,71	+ 154
courant	97	0,23	1,80	317	0,10	0,56	+ 220
situation	98	0,23	1,78	257	0,12	0,66	+ 171
gouvernement	99	0,23	1,77	237	0,13	0,71	+ 149
économie	100	0,23	1,76	1184	0,03	0,16	+ 993
bout	101	0,23	1,75	45	0,39	2,20	-20
compteur	102	0,23	1,75	3540	0,01	0,04	+ 4885
degré	103	0,23	1,75	847	0,04	0,23	+ 652
choix	104	0,22	1,71	531	0,07	0,37	+ 368
logement	105	0,22	1,71	2058	0,01	0,08	+ 2026
prix	106	0,22	1,70	315	0,10	0,57	+ 199
campagne	107	0,22	1,69	445	0,07	0,41	+ 307
air	108	0,22	1,68	18	0,68	3,89	- 57
gars	108	0,22	1,68	263	0,11	0,65	+ 160
époque	110	0,22	1,67	210	0,13	0,76	+ 122

affaire	111	0,22	1,67	43	0,39	2,22	-25
différence	112	0,22	1,65	1009	0,03	0,19	+ 758
information	112	0,22	1,65	456	0,07	0,41	+ 305
limite	114	0,21	1,62	662	0,05	0,30	+ 443
société	115	0,21	1,62	536	0,06	0,36	+ 344
peur	116	0,21	1,60	65	0,33	1,87	-14
appartement	117	0,21	1,60	284	0,11	0,62	+ 158
convention	118	0,21	1,59	2832	0,01	0,05	+ 3089
porte	119	0,21	1,59	15	0,73	4,12	- 61
salle	120	0,21	1,57	118	0,22	1,27	+ 24
machine	121	0,20	1,55	455	0,07	0,41	+ 278
attention	122	0,20	1,55	174	0,16	0,92	+ 68
produit	123	0,20	1,54	1007	0,03	0,19	+ 697
histoire	124	0,20	1,51	49	0,36	2,07	-27
fonction	125	0,20	1,50	847	0,04	0,23	+ 548
consommation	126	0,19	1,49	2235	0,01	0,07	+ 1964
train	126	0,19	1,49	70	0,32	1,80	- 17
confort	128	0,19	1,47	2289	0,01	0,07	+ 2013
fond	129	0,19	1,46	44	0,39	2,21	- 34
étude	130	0,19	1,45	626	0,06	0,31	+ 362
filles	131	0,19	1,44	23	0,60	3,42	- 58
nombre	131	0,19	1,44	365	0,09	0,50	+ 188
salaire	133	0,19	1,42	1785	0,02	0,10	+ 1387
ville	134	0,18	1,42	61	0,34	1,91	- 26
instant	135	0,18	1,41	46	0,39	2,20	- 36
état	136	0,18	1,41	88	0,27	1,55	- 9
sujet	137	0,18	1,41	202	0,14	0,78	+ 80
confiance	138	0,18	1,40	290	0,11	0,61	+ 129
programme	139	0,18	1,40	1105	0,03	0,17	+ 702
voiture	140	0,18	1,39	58	0,35	1,96	- 29
plan	141	0,18	1,38	266	0,11	0,64	+ 115
chômage	142	0,18	1,38	3438	0,01	0,04	+ 3625
mesure	143	0,18	1,37	230	0,13	0,72	+ 91
cadre	144	0,18	1,35	857	0,04	0,23	+ 487
route	145	0,18	1,35	79	0,30	1,72	- 22
mal	146	0,17	1,32	213	0,13	0,75	+ 76
type	147	0,17	1,31	59	0,34	1,93	- 32
ouvrier	148	0,17	1,30	888	0,04	0,22	+ 485
région	148	0,17	1,30	489	0,07	0,39	+ 236
conseil	150	0,17	1,29	389	0,08	0,47	+ 174
été	151	0,17	1,29	367	0,09	0,50	+ 158

Lecture : "chose" est le substantif le plus utilisé dans le français parlé avec une fréquence de 2,8 pour mille mots, soit 21,1 pour mille substantifs. A l'écrit, « chose » vient au cinquième rang avec une fréquence de 1,2 pour mille mots, soit 6,6 pour mille substantifs utilisés à l'écrit. Son utilisation dans le français parlé est de 135 % plus élevée qu'à l'écrit (ou encore multiplication par 2,35).

Les 151 premiers substantifs, donnés dans le tableau 17 couvrent à peine plus de la moitié de la surface occupée par les substantifs. Pour atteindre les 90 % - adoptés comme objectif pour une bonne couverture - il faut considérer les 1 423 premiers substantifs (voir annexe 3). Il y a donc une grande dispersion qui s'explique avant tout par la diversité des thèmes abordés dans ces entretiens. Dans le français écrit, ces proportions sont atteintes au 399^e (50 %) et 3 887^e pour les 90 %. Autrement dit, non seulement le français parlé utilise moins de noms, mais leurs emplois sont concentrés sur un nombre environ trois fois plus limité que dans le français écrit.

Cette concentration s'explique en partie par le nombre très grand de thèmes abordés dans le corpus écrit, mais elle correspond surtout à ce qui a déjà été constaté pour le groupe du verbe : dans la conversation, le vocabulaire est plus limité, ce qui pourrait faciliter la première initiation au FLE.

Dans la dernière colonne, à part *femme, vie, place, matin, chambre, eau et état*, tous les écarts sont significatifs au seuil choisi.

Le tableau 17 révèle la principale caractéristique de ce français ordinaire. Il privilégie quelques mots généraux ou vagues, en premier lieu : *chose* et *fait* pour désigner des objets mais aussi des actions ou des situations particulières ; *gens, personne* (une ou des...) pour désigner quelqu'un ou quelques-uns dont le nom est inconnu ou qu'il n'est pas nécessaire de nommer (voir ce qui a été dit plus haut à propos du pronom "on"). Ces mots sont des hyperonymes, c'est-à-dire des termes dont le sens inclut celui d'un ou plusieurs autres termes qui sont ses hyponymes car leurs signifiés sont plus restreints. Par exemple, "chose" est l'hyperonyme de n'importe quel objet mais aussi d'une situation, etc.

Plusieurs thèmes émergent de cette liste.

7.2 Les principaux thèmes

Le temps est le principal thème dans les deux registres. Il s'agit de la durée même si le temps qu'il fait (les conditions atmosphériques, spécialement les températures) est aussi important. Les principaux substantifs qui constituent ce thème sont les suivants (extraits de la liste en annexe 3, classement par ordre alphabétique) :

an, année, après-midi, avenir, date, déjeuner, délai, demi-heure, demi-journée, dîner, durée, échéance, époque, étape, heure, horaire, horloge, instant, jour (et les jours de la semaine du lundi au dimanche avec insistance sur le dimanche), journée, lendemain, matin, midi, minuit, minute, mi-temps, mois (et les mois de janvier à décembre), moment, nuit, passé, période, petit-déjeuner, quinzaine, quotidien, saison (automne, été, printemps, automne), seconde, semaine, soir, soirée, veille, week-end...

Généralement, chacun de ces vocables entre dans des groupes nominaux plus ou moins figés, des formules toutes faites. A titre d'exemple, voici les plus fréquentes : *l'année dernière* (ou *prochaine*), *à l'époque de*, *la veille (lendemain) de...* Puis viennent les rythmes de la *journée* (*matin, midi, soir*) les jours de la *semaine*, le *mois*, la *saison*, l'*année*...

Cette forte présence du thème du temps semble logique puisque l'enquête fait un récit. Mais, pour beaucoup de ces mots – l'écart entre les deux registres n'est pas grand. C'est que le temps est aussi un thème dominant du roman français. On ajoutera que, dans la conversation la majorité des verbes étant au présent, ces noms et ces formules courantes sont donc, avec certains adverbess, les principaux moyens utilisés pour situer le propos en dehors de l'instant de la

conversation. Là encore, les expressions sont parfois vagues : "il y a un moment", "un certain temps", "quelques heures" (ou "quelques jours, mois, années"), etc.

Puisque le substantif est le principal vecteur du contenu, il est logique que les thèmes des enquêtes soient signalés par les noms dont les écarts sont les plus forts. Par exemple, pour la politique, en fonction du rang dans le tableau (puis dans la liste en annexe), on trouve : *politique* (féminin) avec un écart de + 1416 , suivie de : gauche, droite, parti, candidat, gouvernement, débat, président, élection, ministre, vote, liberté, république, démocratie....

Signalons également les enquêtes sur le confort de la maison (électricité, chauffage, radiateur, facture, compteur, etc.) ou concernant les relations professionnelles (négociation, syndicat, emploi, convention, salaire, chômage...)

En dehors de ces questions spécifiques, il est possible d'isoler les principaux thèmes de la conversation. Après le temps vient la vie quotidienne (classement alphabétique des principaux noms rattachés à ce thème dans le corpus) :

argent, collège, cours, course, déjeuner, départ, dîner, école, éducation, foyer, linge, lycée, ménage, petit-déjeuner, quotidien, repas, retour, santé, sport, trajet, transport, vacance, vie

Puis la famille :

amant, ami(e), amoureux, beau-frère, beau-parent, bébé, belle-sœur, chien, copain, couple, divorce, enfant, entente, épouse, époux, ex-épouse, ex-mari, famille, fidélité, femme, fille, fils, fiston, frère, gamin, garçon, gars, gendre, gosse, grand-mère, grand-parent, grand-père, infidélité, jalousie, jeune, liaison, maîtresse, maman, mari, mariage, mère, mésentente, papa, parent, partenaire, père, petite-fille, petit-fils, relation, séparation, sœur

L'habitation, le quartier et la ville :

appareil, appartement, aspirateur, auto(mobile), baignoire, bâtiment, chaise, chaleur, chambre, chauffage, chauffe-eau, clé, chaise, chaudière, cité, climatisation, commune, compteur, confort, congélateur, consommation, construction, couloir, cour, cuisine, cuisinière, degré, domicile, douche, eau, électricité, énergie, entrée, escalier, essence, étage, évier, facture, fauteuil, fenêtre, four, froid, garage, gaz, grenier, habitation, immeuble, installation, internet, investissement, jardin, lampe, lavabo, lave-linge, lave-vaisselle, linge, lit, locataire, logement, lotissement, loyer, machine à laver, maison, ménage, meuble, pavillon, pièce, placard, plancher, poêle (nf et nm), porte, propriétaire, quartier, radiateur, réparation, résidence, salle de bain, salon, séjour, table, téléphone, télévision, température, toilette, toit, tuyau, vaisselle, village, ville, ventilation, voisin(e), volet, voisinage, voiture

Enfin, la vie professionnelle, en particulier les principaux noms de métiers et de professions :

activité, agent, agriculture, artisan, atelier, augmentation, avocat, bleu, boîte, boulot, bureau, cadre, caisse, caissière, capital, carrière, chantier, chef, chômage, chômeur, client, clientèle, collègue, commerçant, commercial, compétence, conflit, délégué, directeur, docteur, embauche, emploi, employé(e) , employeur, enseignant(e) , entreprise, établissement, expérience, exploitation, fabricant, fermeture, formation, gardien, grève, industrie, infirmière, ingénieur, licenciement, main-d'œuvre, maîtresse, médecin, métier, négociation, outil, ouvrier(ère), paie, patron, personnel, production, professeur, professionnel, profit, promotion, retraite, réunion, revenu, salaire, salarié(e), sanction, secrétaire, société, stage, succursale, syndicat, tâche, technicien, travail, usine, usine, vendeur

Peu d'entretiens ont abordé des thèmes comme les loisirs ou la culture. Ces listes ne sont donc pas complètes. Au total, – en l'état actuel des corpus, un peu plus de 1 400 substantifs semblent

former la base du français parlé. La plupart d'entre eux s'actualisent dans quelques formules usuelles. Certaines seront étudiées dans la suite de ce rapport. Au-delà de ce premier recensement, le concordancier en ligne permettra de compléter ces emplois.

7.3 L'intégration du nom dans le syntagme verbal

Au début du chapitre, il a été signalé qu'il y a un accord chez les grammairiens pour estimer que l'une des caractéristiques du nom est de ne pouvoir être utilisé sans un déterminant ou une préposition (outils qui seront présentés dans un prochain chapitre). Cependant, on observe, dans les deux corpus, un nombre important de compositions "verbe + nom", le nom venant directement derrière le verbe ("avoir besoin", "faire envie"...) avec parfois un adverbe intercalé ("avoir *très* envie"). En quelque sorte, le nom a été dépouillé de tout déterminant et intégré dans le groupe verbal, comme s'il avait été "satellisé" par le verbe. Le relevé systématique de ces compositions permet de mesurer l'ampleur du phénomène. Le tableau 18 donne les compositions les plus fréquentes.

Tableau 18. Les principaux syntagmes verbaux intégrant un nom sans déterminant dans les français parlé et écrit (classement par effectifs décroissants)

Syntagmes	Parlé		Ecrit		Ecart
	Rang	%Verbes	rang	%Verbes	
avoir besoin	1	2,26	1	1,50	+51
avoir envie	2	1,48	3	0,93	+60
rendre compte	3	1,14	4	0,85	+33
faire partie	4	1,02	8	0,29	+249
faire attention	5	0,80	9	0,25	+213
avoir peur	6	0,62	2	1,04	-40
donner envie	7	0,44	45	0,06	+585
être dommage	8	0,43	37	0,08	+442
avoir tendance	9	0,41	43	0,07	+502
faire confiance	10	0,31	20	0,13	+136
faire peur	11	0,30	10	0,25	+18
avoir droit	12	0,28	17	0,16	+81
tenir compte	13	0,27	33	0,09	+213
avoir raison	14	0,25	5	0,70	-64
avoir marre	15	0,22	21	0,13	+72
avoir mal	16	0,20	115	0,02	+813
avoir confiance	17	0,19	23	0,12	+62
faire plaisir	18	0,19	12	0,23	-18
être monsieur	19	0,16	57	0,05	+184
avoir grand-chose	20	0,14	52	0,06	+149
avoir lieu	21	0,14	6	0,31	-54
être question	22	0,14	13	0,21	-35
faire appel	23	0,13	32	0,09	+48
être président	23	0,13	288	0,01	+1588
avoir intérêt	25	0,11	50	0,06	+79

prendre conscience	26	0,11	24	0,11	-5
avoir affaire	27	0,10	19	0,14	-32
prendre rendez-vous	28	0,09	143	0,02	+421
avoir moyen	28	0,09	54	0,06	+61
faire face	30	0,08	22	0,13	-35
avoir tort	31	0,08	11	0,23	-66
être fonctionnaire	32	0,07	468	0,00	+1522
poser problème	33	0,06	333	0,01	+929
être chef	34	0,06	148	0,02	+265
être cadre	34	0,06	1273	0,00	+5915
faire grand-chose	34	0,06	152	0,02	+276
être ministre	37	0,06	299	0,01	+737
rendre service	38	0,06	64	0,05	+21
changer grand-chose	39	0,06	244	0,01	+498
être maire	39	0,06	585	0,00	+1694
faire grève	41	0,05	380	0,01	+945
avoir monsieur	41	0,05	355	0,01	+850
avoir accès	43	0,05	66	0,04	+18
être directeur	43	0,05	683	0,00	+1926
être conseiller	45	0,05	423	0,00	+990
être femme	45	0,05	93	0,03	+69
être locataire	45	0,05	-		
dire bonjour	48	0,05	59	0,05	-6
être artisan	49	0,05	-		
avoir conscience	49	0,05	35	0,09	-44

Lecture : le syntagme "avoir besoin" est le plus utilisé dans le français parlé avec une fréquence de 2,26 pour mille emplois de tous les verbes. A l'écrit, ce syntagme est également le premier avec une fréquence de 1,5 ‰ soit 51 % de plus dans le français parlé qu'à l'écrit.

Il est frappant de voir que la tête de liste est assez largement partagée entre les deux registres. Nous avons donc affaire à un phénomène langagier qui transcende les genres. En effet, la plupart de ces formules toutes faites sont familières au locuteur français. Le phénomène est donc assez fréquent, même si les effectifs se réduisent rapidement au fur et à mesure qu'on descend dans le tableau. Certains syntagmes montrent la difficulté de tracer une ligne étanche entre le nom, l'adjectif et l'adverbe. Par exemple, dans une expression comme "pas grand-chose" (précédé de *avoir, changer, faire, être, comprendre...*) ces deux mots peuvent être interprétés comme formant une "locution adverbiale". Autre exemple : "*C'est monsieur X* qui s'occupe de l'affaire". Le nom propre se passe habituellement de déterminant et "monsieur" est ici l'appendice d'un nom propre. Ces deux réserves admises (locutions et noms propres), il est impossible de considérer que ce phénomène ne met pas en cause le principe du nom nécessairement déterminé à moins d'admettre que les vocables ainsi satellisés par un verbe perdent leur qualité de noms dans une telle position.

Les 50 premières combinaisons couvrent 63 % des occurrences de ce type de construction survenant dans une conversation contre 51 % à l'écrit. Donc, dans le français parlé, il se produit une concentration sur les combinaisons plus fréquentes, ce qui indique les seules indispensables

à connaître pour soutenir une conversation. Au total, les expressions les plus fréquentes sont nettement suremployées dans le français parlé selon une tendance déjà constatée pour les autres catégories grammaticales. Cependant, ici l'étalement est nettement plus fort. Sur les 2 191 expressions différentes, 1 447 – soit les deux tiers - n'apparaissent qu'une fois et couvrent 11 % de la surface du corpus occupée par ce type de construction. Au fond, il suffit de connaître le principe pour engendrer une multitude de formulations adaptées au thème de la communication. Du point de vue des écarts (dernière colonne) un seul écart est non significatif ("prendre conscience" (26^e rang dans le français parlé et 24^e à l'écrit). Ce syntagme peut être suivi de la préposition "de" (*prendre conscience d'une chose*) ou par une conjonction de subordination (*prendre conscience qu'il est tard*). Il en est ainsi de la majorité des syntagmes listés dans le tableau 18 ce qui rappelle le caractère multifonction de ce type de construction.

Les principaux sous-emplois sont : "avoir peur", le couple "avoir raison / avoir tort" ; les expressions : *faire plaisir, avoir lieu, être question (de), avoir affaire (à), faire face (à), avoir conscience (de)*.

A l'inverse, la conversation fait un large usage des principales professions – être fonctionnaire, chef, cadre, ouvrier... - et, plus largement, des qualités puisque c'est la manière normale de se présenter et de désigner les autres. Nous revenons plus bas sur ce point fondamental.

On aura aussi remarqué que certains verbes apparaissent dans un grand nombre de combinaisons, en particulier "être" et "avoir" qui, à eux deux, fournissent plus de la moitié des 50 syntagmes les plus fréquents. Pour en être certains, nous avons effectué un recensement de tous les verbes suivis d'un nom non-déterminé. Le tableau 19 ci-dessous répond aux questions suivantes : Quels sont les principaux verbes concernés et combien de combinaisons différentes chacun d'entre eux génère-t-il ? Combien de fois apparaissent ces combinaisons et quelles proportions de ces verbes sont concernées ? Pour "être" et "avoir", les emplois comme auxiliaire d'un participe passé ont été retirés du dénominateur.

Ces 40 verbes représentent à eux seuls près de 90 % de la surface du texte occupée par ces constructions "verbe+nom" sans déterminant entre les deux. Sans surprise, on trouve en tête les trois verbes les plus usuels de la langue française (*être, avoir et faire*).

Les noms susceptibles de se trouver juste derrière ces verbes sont extrêmement nombreux et si divers qu'il semble difficile de réduire cette diversité à quelques règles simples. Cependant trois choses dominent.

Tableau 19. Les principaux verbes pivots dans les syntagmes "verbe+nom"

Verbe pivot	Combinaisons	Occurrences	% Verbes
avoir	302	4 938	6,80
être	819	3 120	2,37
faire	220	2 343	6,52
donner	23	328	8,22
prendre	43	234	0,32
dire	121	233	0,83
tenir	8	183	14,47
parler	51	102	2,14
appeler	45	66	4,15
devenir	44	59	7,66
voir	25	55	0,01
mettre	37	54	0,74
passer	19	45	0,84
porter	14	38	12,18
trouver	14	33	0,57
tester	23	29	1,38
connaître	7	25	0,93
voter	12	24	0,87
savoir	12	19	0,13
aller	16	17	0,08
demander	9	14	0,58
pouvoir	8	13	0,08
commencer	10	12	0,82
entendre	7	11	0,57
payer	8	11	0,38
rentrer	8	10	0,63
vouloir	8	8	0,07

Lecture : le verbe "avoir" est le pivot de 302 combinaisons différentes (avec un substantif sans déterminant), totalisant 4 938 occurrences soit 6,8 % du total des effectifs de ce verbe.

- La majorité de ces associations s'intègrent dans des syntagmes verbaux plus larges construits sur le patron : "verbe + nom + préposition + verbe infinitif" ("avoir envie de faire"), ce qui renvoie aux modalisations déjà étudiées à propos des verbes. Autre patron fréquent : "verbe+nom+subordination" – *avoir envie que* - qui souligne combien, dans ces constructions, le nom est devenu un constituant du groupe verbal.

- Précédé de "être", le nom sans déterminant se transforme en attribut du sujet. Ce sujet est d'abord le pronom démonstratif : "c'est dommage", "c'est décembre", "c'est lundi", "c'est nuit" (ou "jour"), etc. Derrière "être", on trouve ensuite des noms désignant des professions, des qualités ou des états, par exemple : "il est président, fonctionnaire, cadre, chef, ministre, député, directeur, artisan, commerçant, employé, ouvrier, professeur, plombier, cuisinier, chauffeur, étudiant, lycéen, etc. Naturellement, cet état peut ne pas être statique : on "devient fonctionnaire", on "passe cadre", on "reste ouvrier", etc. Tout verbe exprimant un état ou le changement de cet état peut donner lieu à une construction de ce genre, mais cette potentialité s'actualise habituellement avec le petit nombre de verbes listés dans le tableau ci-dessus.

- Pour "avoir" – et certains synonymes exprimant la possession - les noms les plus utilisés expriment des dispositions particulières de l'esprit ou du corps : le *besoin*, l'*envie*, la *peur* (et *grand-peur*), la *confiance*, la *tendance*, l'*horreur*, la *conscience*, la *faim*, la *soif*, la *peine*, le *plaisir*, l'*intention*, la *honte*, etc. Comme synonyme de *posséder* "avoir" sert également à l'expression d'une situation particulière : *avoir droit*, *raison*, *tort*, *affaire*, *accès*, *moyen*, *possibilité* (et *impossibilité*), *incompatibilité*, *rendez-vous...* et *pignon* (sur rue).

Certains noms non-déterminés utilisés avec "avoir" se retrouvent dans la même position avec "faire" : *envie*, *peur*, *confiance*, *attention*, *horreur*, *peine*, *honte...* D'autres sont plutôt spécifiques à "faire", comme (pour s'en tenir aux principaux) : *abstraction*, *allusion*, *appel*, *attention*, *barrage*, *face*, *fortune*, *grève*, *jour*, *marche (arrière)*, *nuit*, *partie*, *pression*, *preuve*, *référence*, *semblant*, *signe...* et naturellement "gaffe" !

D'autres verbes ne se rencontrent que dans quelques usages, souvent assez fréquents comme "tenir compte", "poser problème", "prendre position" ou "prendre garde".

En définitive, la règle demeure selon laquelle un groupe nominal, complet et normalement constitué doit comporter une détermination. Cependant, il n'en est pas de même pour un grand nombre de noms qui peuvent être précédés de verbes sans déterminant entre les deux. Ce phénomène d'intégration au groupe verbal est surtout idiomatique et il semble impossible de le réduire à quelques règles simples, sinon pour souligner le petit nombre de verbes usuels susceptibles d'entrer dans ces combinaisons.

Chapitre 8

Les adjectifs

Ce chapitre porte sur l'adjectif qualificatif à l'exclusion des vocables que la grammaire traditionnelle appelait "adjectifs déterminatifs" (démonstratifs, possessifs, numéraux, indéfinis) qui seront étudiés dans le chapitre suivant.

L'adjectif qualificatif a deux principales caractéristiques. En premier lieu, il est variable en genre et en nombre. Par exemple, les mots *grand*, *grands*, *grande*, *grandes* sont groupés sous le vocable (entrée de dictionnaire) : grand, adj) mais les mots "grand" et "grands" peuvent aussi être des noms (*les grands de ce monde*). Deuxièmement, l'adjectif qualificatif peut se trouver en position d'épithète (une chose *belle*) ou d'attribut (cette chose est *belle*). Cependant, en pratique, de nombreux adjectifs ne se trouvent jamais en position d'attribut même si, grammaticalement, rien ne s'y oppose. Enfin, du point de vue statistique, l'adjectif est lié au nom et sa densité d'emploi varie dans le même sens que celle du nom, comme on l'a vu en étudiant les parties du discours

L'étiquetage permet de distinguer l'adjectif du nom, du verbe et de l'adverbe homographes. Par exemple : *grand* cité ci-dessus, le *français* (la langue), le gouvernement *français* ; un *mort*, un homme *mort*, il est *mort* ; il fait *beau*, un *beau* garçon. Les étiquettes distinguent également l'adjectif pur de l'adjectif dérivé du participe passé (comme "homme *mort*"). Le tableau 2, plus haut, a indiqué que cette seconde catégorie est peu utilisée dans le français parlé. Sauf quelques adjectifs courants comme "mort", "blessé", "cassé", "satisfait", "déçu", "fatigué", etc. Ces adjectifs dérivés de verbes sont près de quatre fois moins présents dans le français parlé qu'à l'écrit. Comme le participe présent, l'adjectif dérivé du participe passé semble donc être un procédé plutôt littéraire.

Nous examinerons d'abord les adjectifs privilégiés par le français parlé avant de discuter une question spécifique : la place de l'adjectif devant ou derrière le nom.

8.1 Les adjectifs usuels du français parlé

Quelques adjectifs, en général très usuels, sont nettement plus employés dans le français parlé qu'à l'écrit (tableau 20).

Tableau 20. Les principaux adjectifs du français parlé comparés à leur utilisation à l'écrit (fréquence pour mille mots)

Vocable	Parlé			Ecrit			Ecart %
	Rang	Fréquence (%)	% Adjectifs	Rang	Fréquence (%)	% Adjectifs	
petit	1	1,79	53,01	1	1,63	30,97	+ 71
vrai	2	1,50	44,30	11	0,40	7,61	+ 482
bon	3	0,98	28,90	4	0,80	15,13	+ 91
sûr	4	0,87	25,74	12	0,40	7,53	+ 242
électrique	5	0,58	17,29	163	0,05	1,00	+ 1633
important	6	0,55	16,26	81	0,09	1,75	+ 827
grand	7	0,54	15,86	2	1,12	21,25	- 25
pareil	8	0,53	15,70	26	0,17	3,26	+ 381
seul	9	0,47	13,85	3	1,00	18,99	- 27
social	10	0,39	11,55	190	0,05	0,87	+ 1223
politique	11	0,38	11,20	108	0,07	1,39	+ 708
gros	12	0,37	10,95	15	0,35	6,75	+ 62
chaud	13	0,37	10,91	47	0,13	2,51	+ 335
différent	14	0,36	10,66	71	0,11	2,05	+ 421
difficile	15	0,34	10,06	37	0,14	2,75	+ 266
dernier	16	0,34	9,91	8	0,50	9,45	+ 5
nouveau	17	0,31	9,31	5	0,64	12,10	- 23
froid	18	0,30	8,90	28	0,17	3,19	+ 179
général	19	0,29	8,53	80	0,09	1,76	+ 384
raisonné	20	0,29	8,43	3311	0,00	0,03	+ 26802
normal	21	0,28	8,15	111	0,07	1,36	+ 497
possible	22	0,27	7,88	17	0,24	4,54	+ 74
cher	23	0,26	7,74	53	0,13	2,42	+ 220
intéressant	24	0,23	6,76	213	0,04	0,77	+ 783
facile	25	0,23	6,70	73	0,10	1,93	+ 246
clair	26	0,21	6,31	49	0,13	2,49	+ 154
capable	27	0,19	5,67	66	0,11	2,13	+ 167
syndical	28	0,19	5,58	2138	0,00	0,06	+ 8595
économique	29	0,19	5,54	318	0,03	0,52	+ 973
mauvais	30	0,19	5,50	21	0,20	3,83	+ 44
évident	31	0,18	5,30	210	0,04	0,78	+ 578
plein	31	0,18	5,30	16	0,34	6,46	- 18
actuel	33	0,18	5,24	568	0,02	0,30	+ 1675
haut	34	0,18	5,21	20	0,21	3,92	+ 33
simple	35	0,17	5,07	25	0,18	3,34	+ 52
dur	36	0,17	4,93	38	0,14	2,74	+ 80
donné	37	0,17	4,87	470	0,02	0,36	+ 1254
rouge	38	0,16	4,83	18	0,22	4,19	+ 15
jeune	39	0,16	4,71	6	0,63	11,98	- 61
beau	40	0,16	4,70	9	0,49	9,31	- 50
blanc	41	0,16	4,68	14	0,37	6,97	- 33
content	42	0,16	4,61	94	0,08	1,52	+ 204
fort	43	0,16	4,58	24	0,20	3,78	+ 21

collectif	44	0,15	4,40	707	0,01	0,23	+ 1777
français	45	0,14	4,27	23	0,20	3,80	+ 12
juste	46	0,13	3,83	97	0,08	1,50	+ 156
personnel	47	0,13	3,80	126	0,06	1,22	+ 211
creux	48	0,13	3,69	442	0,02	0,38	+ 862
traditionnel	49	0,12	3,67	749	0,01	0,22	+ 1541
ancien	50	0,12	3,59	27	0,17	3,21	+ 12

Lecture : voir sous tableau 7. +

Si l'on considère la surface du corpus occupée par les adjectifs (quatrième colonne) et que l'on cumule ces chiffres, on constate que les 50 premiers – listés dans ce tableau – occupent la moitié de cette surface. Les 75 % sont atteints avec le 227^e et les 90 % sont dépassés au 753^e. La liste de ces 753 adjectifs du français parlé courant est donnée en annexe 4. À l'écrit, ces bornes sont 201^e (50 %) et 2 394^e (90 %). Dans le français parlé par rapport à l'écrit, la concentration du propos sur les adjectifs les plus fréquents est donc plus nette que pour les noms. Cependant, contrairement aux verbes et à leurs satellites, les adjectifs usuels sont malgré tout plus nombreux (ou plus divers). Naturellement, l'influence des thèmes des entretiens est ici aussi bien visible avec, dans la liste en annexe 4, la présence de vocables comme "gaulliste", "centriste", "socialiste" ou "communiste" dont on peut évidemment contester qu'ils fassent partie du "français courant" !

Il faut rappeler que la densité des adjectifs, dans le français parlé, est de 35,6 % inférieure à celle de l'écrit, c'est-à-dire plus d'un tiers en moins. Pour interpréter la dernière colonne du tableau 20, c'est cette norme qu'il faut garder en tête. Ainsi "grand" (écart *net* -25 %) a une densité d'emploi *brute* de 60 % inférieure dans le français parlé par rapport à l'écrit. En sens inverse, les six premiers de la liste sont nettement suremployés. Dans son désamour pour les adjectifs, le français parlé manifeste une préférence marquée pour "petit" et quelques autres comme "sûr", "bon", "vrai", "important", "pareil" – surtout combinés avec le démonstratif "ce" – dont nous avons dit plus haut, à propos des satellites du verbe, qu'il s'agit d'une caractéristique du français parlé : "c'est petit", "c'est sûr", "c'est vrai", "c'est bon", "c'est pareil", etc. Ces tournures sont très employées dans la conversation courante. Ce sont des patrons tout prêts indispensables qui peuvent être enseignés en même temps que le verbe qui en est le pivot.

L'adjectif "dernier" est le seul à avoir une densité d'emploi à peu près semblable dans les deux corpus. Les 49 autres listés dans le tableau... présentent des écarts significatifs. Sept sont sous-employés : "grand", "seul", "nouveau", "plein", "jeune", "beau" et "blanc". Les 43 autres premiers adjectifs listés dans le tableau connaissent une utilisation nettement plus forte dans le français parlé par rapport à l'écrit.

Les adjectifs "électrique", "social" et "politique" doivent certainement leurs places dans la tête de liste aux thèmes d'une proportion importante des entretiens. Il en est de même pour "collectif" ou "raisonné" qui s'appliquent à la négociation des conventions collectives. En revanche, les autres sont incontestablement propres à la conversation en général.

Plus fondamentalement, les adjectifs vont par couples : *petit/gros* (*grand* à l'écrit), *chaud* et *froid*, *nouveau* et *ancien*, *pareil* et *différent* (ou *identique*), *bon/mauvais* ; *difficile/facile* ; *possible/impossible* ; *vieux/jeune*, etc. Quand l'un des deux termes est significativement suremployé par rapport à l'autre, l'écart fait sens. Par exemple, le français parlé privilégie "petit", parce que les problèmes quotidiens portent rarement sur de grandes choses (du moins les locuteurs n'osent pas le prétendre). Dans la littérature, le couple "petit/grand" est à peu près équilibré. A l'inverse, le discours politique français n'utilise pratiquement que "grand", même à l'oral, car les dirigeants ont de grands desseins et ne s'occupent guère des contingences qui sont au cœur de la conversation.

8.2 Devant ou derrière le nom ?

C'est un problème qui déconcerte les débutants en FLE. Certes, en écoutant les Français, on apprend vite qu'un "cher ami" ne désigne pas forcément un "ami cher" ou qu'un "petit problème" n'est pas forcément négligeable. Dans les grammaires, il est indiqué que la plupart des adjectifs peuvent se placer devant le nom et qu'il s'agit d'une mise en relief de l'adjectif. En quelque sorte, l'élément lourd de la combinaison serait placé devant l'autre. Le français parlé fait apparaître une autre dimension : la concentration du phénomène sur un (relativement) petit nombre de formules toutes faites.

A titre d'exemple, le tableau 21 donne les cent combinaisons (adj+nom) les plus fréquentes. Pour le consulter il faut se souvenir que le recensement est effectué sur les lemmes. Par exemple, "bon" est le lemme des adjectifs "bon, bons, bonne, bonnes". Ainsi, "bon chose" regroupe le singulier et le pluriel de cette formule ("bonne(s) chose(s)").

Au centième, les effectifs sont déjà assez faibles. D'ailleurs, si l'on se souvient qu'il y a 433 entretiens (sans compter les sondages), le syntagme le plus fréquent (*gros problème*) n'apparaît au mieux que chez un tiers des locuteurs.

De plus, dans le français parlé, assez peu d'adjectifs semblent susceptibles de se trouver dans la position d'élément premier dans le couple formé d'un nom et d'un adjectif. En s'en tenant aux combinaisons apparaissant au moins cinq fois – seuil au-dessus duquel l'association a peu

de chance d'être due au hasard – il n'y a que 34 adjectifs susceptibles d'être antéposés à un nom (tableau 22).

Tableau 21 Les cent syntagmes (avec adjectif antéposé au nom) les plus fréquents dans le français parlé avec leurs effectifs

Syntagmes	Effectifs				
gros problème	164	prochain fois	32	dernier moment	21
bon chose	144	vrai gens	32	gros chaleur	21
petit truc	126	gros différence	31	meilleur solution	21
seul chose	113	gros facture	31	nouveau président	21
dernier fois	110	nouveau système	31	vrai femme	21
bon sens	94	petit entreprise	31	bon chauffage	20
petit problème	93	bon produit	30	dernier temps	20
dernier année	91	petit commune	30	grand pièce	20
petit chose	88	petit fille	30	libre entreprise	20
petit village	88	petit moment	30	petit coin	20
bon idée	78	grand idée	29	petit convecteur	20
seul coup	76	petit ville	28	petit radiateur	20
grand surface	70	plein pot	28	gros truc	19
long terme	70	vieux maison	28	petit appareil	19
vrai problème	54	ancien système	27	plein heure	19
bon raison	44	bon partie	27	plein sud	19
vrai niveau	43	grand ligne	27	bas salaire	18
bon solution	41	petit peur	27	bon qualité	18
dernier élection	41	petit salaire	27	grand entreprise	18
petit bout	41	dernier facture	26	grand monde	18
grand ville	39	dernier négociation	26	grand principe	18
mis part	39	bas âge	24	petit moyen	18
petit coup	39	petit boîtier	24	plein milieu	18
petit maison	39	vieux pays	24	seul raison	18
plein hiver	39	haut direction	23	bon côté	17
bon question	38	juste milieu	23	bon femme	17
bon moment	36	petit commerçant	23	bon santé	17
bon heure	35	petit détail	23	grand public	17
grand différence	34	petit dommage	23	nombreux année	17
seul truc	34	petit niveau	23	nouveau façon	17
court terme	33	bon volonté	22	petit boîte	17
grand maison	33	gros travail	22	petit chauffage	17
grand partie	33	petit côté	22	petit papier	17
seul fois	33	petit gens	22	prochain élection	17
petit bouton	32	seul problème	22	vrai vie	17
		bon température	21		

Parmi ces 34 adjectifs - placés devant le nom un nombre significatif de fois - 26 apparaissent dans plus d'une combinaison. Ce sont donc ceux-là qui doivent être signalés comme pouvant, dans la "vraie vie", entrer dans la construction "adjectif + nom". Pour les autres, leur place normale est, en épithète, derrière le nom ou en attribut derrière un verbe.

Tableau 22. Les adjectifs antéposés apparaissant au moins dans une combinaison répétée au moins cinq fois (classement hiérarchique sur les effectifs).

Adjectif antéposé	Nombre compositions	Effectifs
petit	102	1 869
bon	43	924
grand	34	544
gros	26	488
vrai	24	367
nouveau	23	233
seul	13	354
mauvais	13	120
dernier	10	364
plein	10	155
ancien	9	107
meilleur	7	72
vieux	6	83
beau	6	38
prochain	5	79
juste	4	58

froid	4	31
double	4	26
bas	3	55
haut	3	44
chaud	3	22
important	3	21
moyen	3	16
libre	2	26
principal	2	25
cher	2	20
long	1	70
court	1	33
nombreux	1	17
moindre	1	9
joli	1	6
propre	1	6
blanc	1	5
pareil	1	5

Naturellement, cette tournure dépend des circonstances et de la sociologie des locuteurs. Par exemple, si c'est un *haut* fonctionnaire qui parle de son travail, les syntagmes avec adjectifs antéposés seront probablement un peu différents de ceux du corpus. L'on rencontrerait sans doute des syntagmes comme : "haute autorité", "haute juridiction", "haut conseil", "haut représentant"... Même s'il n'a jamais entendu ces formules auparavant, l'auditeur les comprendra puisqu'il sait que l'adjectif "haut" est connu comme pouvant se placer devant le nom pour générer ce type de syntagmes.

Chapitre 9

Les mots outils dans le groupe du nom

Rappelons que l'appartenance à l'un des deux groupes – par exemple celui du verbe - n'est pas exclusive d'apparitions dans l'autre (celui du nom). Ainsi, beaucoup d'adverbes peuvent s'insérer dans un groupe nominal, spécialement devant un adjectif (une chose *très* petite, un homme *vraiment* fort), de même des prépositions peuvent entrer dans un syntagme verbal (voir plus haut la discussion sur la modalisation). On étudiera ici leur usage dans les groupes nominaux du français parlé.

Le tableau 2 montre la faible densité d'utilisation des déterminants dans le français parlé (plus d'un cinquième en moins par rapport à l'écrit). Sauf les nombres, toutes les catégories de déterminants sont concernées y compris les articles pourtant considérés comme des outils incontournables.

Dans le langage, les chiffres et les nombres diffèrent sensiblement des autres catégories. Comme les noms propres et certains déictiques ils servent à ancrer le discours dans le temps et l'espace (social, économique, géographique). De plus, les chiffres donnent une apparence de précision ou d'objectivité. Leur densité est moitié moindre à l'écrit que dans la conversation. Ce sont avant tout des dates – comme indiqué ci-dessus la temporalité est un thème important dans la conversation - mais aussi des quantités (en durées, monnaies, degrés, mètres, kilomètres, kilowatts...) Autrement dit, les numéraux, les cardinaux et les unités sont des outils indispensables pour l'apprentissage du FLE.

9.1 Les déterminants dans les français parlé et écrit

Le tableau 23 donne les principaux déterminants classés par ordre décroissant de fréquence. Il est bon de rappeler une fois de plus que le recensement porte sur les lemmes. Par exemple, la première ligne ("le") comporte les "formes graphiques" : *l', le, la, les* – employées comme articles et non comme pronoms - auxquels s'ajoute le second membre des contractions "du", "des", "au", "aux" (de+le et à+le).

Tableau 23. Les déterminants dans le français parlé comparé à l'écrit (classement hiérarchique, en pourcentage du total des occurrences des déterminants)

Déterminant	Parlé		Ecrit		Ecart %
	Rang	%Déterm.	Rang	%Déterm	
le	1	568,4	1	562,2	+1
un (art)	2	148,1	2	147,2	+1
tout	3	32,6	6	19,1	+71
ce	4	29,6	4	44,1	-33
mon	5	24,7	5	27,4	-10
deux	6	17,3	7	12,4	+40
son	7	12,8	3	74,7	-83
autre	8	10,8	11	6,4	+69
vingt	9	9,7	18	3,9	+152
cent	10	9,6	16	4,0	+137
quelque	11	9,6	9	9,9	-3
trois	12	9,2	14	5,0	+85
notre	13	9,2	10	6,9	43
quatre	14	8,7	19	3,7	137
leur	15	8,4	8	10,8	-22
même	16	7,6	17	4,0	89
cinq	17	7,1	11	2,7	164
mille	18	7,1	9	3,3	112
dix	19	7,0	10	3,1	126
certain	20	5,7	25	2,7	112
premier	21	5,0	12	5,8	-13
huit	22	4,4	27	1,9	135
six	23	3,8	30	1,7	130
trente	24	3,4	31	1,5	123
neuf	25	3,2	26	2,4	35
sept	26	3,2	33	1,4	122
tel	27	2,6	29	1,8	46
soixante	28	2,5	18	1,0	155
chaque	29	2,4	23	3,1	-23
quinze	30	2,3	18	0,9	153
quel	31	2,3	21	3,2	-29
un (num)	32	2,3	22	0,8	183
aucun	33	2,1	22	3,2	-35
cinquante	34	2,1	37	0,9	133
votre	35	2,0	13	5,5	-64
plusieurs	36	1,8	32	1,4	25
quarante	37	1,7	34	1,0	67
deuxième	38	1,3	42	0,5	149
ton	39	1,2	15	4,4	-73
zéro	40	1,1	48	0,3	322
douze	41	1,0	24	0,6	76
seize	42	0,9	28	0,4	152
quatorze	43	0,9	27	0,4	102
troisième	44	0,7	23	0,6	16
treize	45	0,7	29	0,3	117
onze	46	0,6	26	0,5	22
cinquième	47	0,3	32	0,2	121
quatrième	48	0,2	31	0,2	10
sixième	49	0,2	33	0,1	130
second	50	0,2	21	0,9	-75

Lecture : l'article "le" est le déterminant le plus utilisé en français parlé avec 568,4 occurrences pour mille déterminants. A l'écrit, cette densité est de 562,2 pour mille déterminants. L'écart entre les deux genres n'est pas significatif (+1 % dans le français parlé par rapport à l'écrit)

Ces cinquante vocables couvrent plus de 99 % des emplois d'un déterminant. Les articles (le, un) fournissent à eux seuls sept occurrences sur dix et "le", à lui seul, plus de la moitié. Dans tout texte en français, cet article est toujours le vocable le plus employé. De façon remarquable, les deux articles ne connaissent aucun écart net significatif entre les deux corpus (c'est-à-dire une fois défalquée la propension beaucoup plus faible à utiliser la détermination dans le français parlé par rapport à l'écrit). Etant donné les différences considérables entre les deux genres, cette stabilité nette était peu attendue. Elle montre que ces deux outils sont au cœur de la langue.

A part l'indéfini "quelque", tous les autres écarts sont significatifs. Ils s'organisent ainsi.

- Dans les suremplois : les nombres et les indéfinis. Comme indiqué plus haut, les nombres (numéraux et cardinaux) sont des interfaces entre le discours et les thèmes dont on parle. Ils servent à ancrer le propos dans la réalité. Une exception : "second". Sauf pour quelques expressions toutes faites - comme "seconde main", "seconde chance"...- la conversation n'utilise que "deuxième". Quant aux indéfinis, ils se limitent à : *tout, autre, même, certain, plusieurs*. Ce sont les seuls indéfinis indispensables à la communication orale.

- Les sous-emplois portent sur le démonstratif et les possessifs. Le démonstratif ("ce") quoique encore fréquent dans la conversation, est d'un tiers moins employé qu'à l'écrit (une fois défalquée la propension moindre à utiliser les déterminants). En fait, il a deux fonctions différentes. Premièrement, il accompagne ou se substitue au geste consistant à montrer quelque chose ou quelqu'un. Cette fonction est très fréquente dans la conversation. Deuxièmement, il a une fonction "anaphorique" : il renvoie à une chose ou une personne déjà mentionnée auparavant dans le propos : très fréquent à l'écrit – car il permet d'éviter les répétitions et les lourdeurs -, le procédé est moins utilisé dans la conversation, sans doute à cause de sa complexité et parce que les répétitions passent mieux qu'à l'écrit.

Enfin, le possessif est nettement sous-employé, sauf le "notre", à mettre en relation avec le suremploi du pronom "nous" : la plupart des locuteurs ont le souci de ne pas parler en leur seul nom, en utilisant trop la première personne, mais de s'inclure dans un groupe qui varie, selon le thème traité, du couple (ou de la famille) – pour la maison ou la vie quotidienne - au peuple français quand on parle de politique. En tous cas, cela permet de ne pas sembler assumer seul le propos.

9.2 Les prépositions

La plupart des prépositions peuvent aussi régir un verbe à l'infinitif mais elles se trouvent devant lui ("demander de faire", "se préparer à dire", "finir par penser", etc.) ce qui leur conserve le statut de *prépositions*.

Quelques prépositions peuvent avoir des emplois adverbiaux : *voilà, devant, derrière*, etc. Elles sont associées à un verbe non suivi d'un groupe nominal derrière (*il marche devant, derrière, il est dessus, dessous*). Au lieu de diminuer quand on passe de l'écrit à la conversation, la densité de ces quelques prépositions augmente, car ces emplois adverbiaux de la préposition sont associés à des verbes qui sont plus fréquents dans le français de la conversation (tableau 24).

Tableau 24. Les prépositions dans le français parlé comparé à l'écrit (fréquence en pour mille prépositions)

Vocables	Parlé		Ecrit		Ecart %
	Rang	% Prépositions	Rang	% Prépositions	
de	1	443,9	1	466,6	-5
à	2	179,2	2	181,9	-2
en	3	74,5	3	61,4	21
pour	4	72,8	5	43,5	67
dans	5	52,3	4	60,1	-13
par	6	30,2	8	24,0	26
sur	7	28,6	6	38,5	-26
avec	8	28,3	7	27,1	4
voilà	9	16,4	23	2,5	554
après	10	14,7	11	7,1	108
avant	11	7,9	13	5,7	38
chez	12	6,9	18	5,1	34
entre	13	5,5	17	5,2	6
sans	14	5,1	8	15,6	-68
contre	15	5,0	19	4,3	19
depuis	16	4,4	16	5,5	-20
près	17	3,6	20	3,5	3
pendant	18	3,4	21	3,2	5
jusque	19	3,1	15	5,6	-45
vers	20	1,7	10	7,7	-78
derrière	21	1,6	22	2,8	-44
dès	22	1,6	24	2,2	-27
sous	23	1,5	12	6,6	-78
sauf	24	1,4	27	0,8	83
devant	25	1,1	14	5,6	-81

Lecture : la préposition "de" est la plus utilisée dans les français parlé et écrit avec, dans le français parlé 443,9 emplois et sur mille occurrences d'une préposition et 466,6 dans le français écrit, soit un écart net de -5 % dans le français parlé par rapport à l'écrit.

Une fois de plus, on observe une forte concentration sur quelques vocables très usuels. Dans le français parlé les 25 prépositions du tableau 24 couvrent 99,6 % des emplois de la conversation

et 99,1 % à l'écrit. Certes, la liste des prépositions françaises est beaucoup plus longue mais celles qui ne figurent pas dans le tableau ci-dessus ne sont guère utiles pour la pratique quotidienne du français.

Il a été signalé plus haut que certaines prépositions sont également attachées à des verbes (elles interviennent notamment dans la modalisation). C'est surtout le cas de "de" et "à". La stabilité de ces deux prépositions – pas d'écart net significatif entre les deux registres - résulte donc d'un double mouvement : une augmentation des emplois adverbiaux due à la plus forte propension de la conversation à utiliser des verbes et une diminution à peu près équivalente résultant de la plus faible densité des noms, spécialement des syntagmes nominaux complexes (examinés ci-dessous). Au moins deux autres sont concernés par ce phénomène : "par" et "pour" mais ils sont nettement préférés dans la conversation, sans doute à cause de leur simplicité d'emploi.

Outre ces deux prépositions (par et pour), quelques autres sont caractéristiques du français parlé. En premier lieu : "voilà" qui se substitue entièrement à "voici" – et qui a deux emplois : désigner une chose ou une personne (peu importe que ce qui est montré soit éloigné ou proche) mais aussi pour marquer la fin du propos : "voilà !" (ce que je voulais vous dire) et, parfois, marquer de l'impatience en face d'une situation ou d'un propos que l'on n'approuve pas. Puis viennent "avant" et "après" qui servent surtout à situer le propos par rapport à des événements (*c'était avant, après X ou Y ; avant ou après une date, un événement*).

En revanche, toutes les prépositions qui servent à indiquer une position dans l'espace sont nettement sous-employées : *dans, sur, sous, devant, derrière, vers, jusque*. Mais, contrairement à "voici" elles n'ont pas disparu et leur connaissance est nécessaire pour entrer en conversation.

9.3 Les prépositions dans les groupes nominaux complexes

Ces vocables servent souvent à former des syntagmes nominaux construits sur le patron { nom + préposition + nom } qui sont moins utilisés dans la conversation à cause sans doute de leur relative complexité, du moins quand elles ne sont pas suffisamment lexicalisées pour être immédiatement disponibles.

Le tableau 25 ci-dessous donne leurs effectifs absolus dans le corpus du français parlé car il semble difficile de leur affecter un dénominateur pour calculer une fréquence relative. Pour interpréter ces effectifs, il faut se souvenir qu'il y a 433 entretiens et quatre sondages. La colonne des effectifs permet donc de se rendre compte que – passé les premières – chacune de ces formules est assez rare et que le phénomène est fortement dispersé.

Tableau 25. Les principaux syntagmes "nom+préposition+nom" dans le français parlé (classement par effectifs décroissant)

Syntagme	Effectifs		
point de vue	764	tas de trucs	45
salle de bain	450	économie d'énergie	44
en fin de compte	447	gouvernement de droite	44
de temps en temps	370	moyen de pression	44
d'heure en heure	256	rapport de force	44
heure du matin	214	employé de bureau	43
relation de travail	180	type de chauffage	43
de gauche à droite	138	de mois en mois	42
chauffage au gaz	133	pays de l'est	41
tas de choses	129	bout du mois	40
en fin d'année	115	dizaine d'année	40
heure du soir	115	mode de vie	40
président de la république	111	d' année en année	39
système de chauffage	109	bout du compte	39
genre de choses	103	chauffage d' appoint	38
table des négociation	98	femme de ménage	38
pouvoir d'achat	95	nombre de choses	38
incompatibilité d' humeur	92	emploi du temps	36
quart d'heure	90	fois par mois	36
du jour au lendemain	86	horaire de travail	36
plupart du temps	80	besoin d'électricité	35
fin de mois	71	président du syndicat	35
gens de droite	71	processus de négociatio	35
heure de travail	71	création d'emploi	34
état d'esprit	69	fin de semaine	34
niveau de chauffage	69	peine de mort	34
sécurité d'emploi	68	solution du problème	34
heure par jour	67	nombre de personne	33
droit de vote	62	équipe de négociation	32
jour par semaine	61	prise de conscience	31
résolution du problème	61	problème d'argent	31
chef d'entreprise	59	agent d'accueil	29
fois par an	59	genre de problème	29
courant d'air	58	mise à pied	29
gens de gauche	58	moyen de chauffage	29
problème de niveau	58	pièce par pièce	29
cadre de vie	57	problème de fond	29
genre de truc	56	règle du jeu	29
facture d'électricité	55	heure de l'après-midi	28
qualité de vie	55	mise en place	28
au bout d'un moment	54	numéro de téléphone	28
condition de travail	53	train de vie	28
conseil d'administration	52	prise de position	27
mise en main	52	droit de l'homme	26
vie de famille	52	monde du travail	26
début d'année	51	niveau de consommation	26
temps de travail	51	nombre d'heure	26
chef d'établissement	50	plupart des gens	26
au jour le jour	49	bout de ligne	25
niveau de vie	48	chambre d'enfant	25
fois par semaine	47	coup de main	25
mode de chauffage	46	au niveau du travail	25
mode d'emploi	45	prise en charge	25
		problème de santé	25

chauffage au fuel	25
contrat d'entretien	24
jour de l'an	24
porte d'entrée	24
avenir des enfants	24
coût de l'installation	23
fois par jour	23
mode de fonctionnement	23
partie des gens	23
porte de la chambre	23
besoin de changement	22
carte à puce	22
chef d'équipe	22
climat de confiance	22
consommation d'électri	22
en cours de route	22
parti de droite	22
recherche d'emploi	22
recherche de solution	22
coup sur coup	21
majorité des gens	21
mère de famille	21
bâton dans les roues	20
chauffage au bois	20
homme de droite	20
idée de gauche	20
liste d'appel	20
niveau de l'emploi	20
niveau de l'entreprise	20
niveau d'étude	20
parti de gauche	20
changement de gouvernem	19
climat de travail	19
coup de téléphone	19
coût de la vie	19
mal de tête	19
niveau de confiance	19
part des choses	19
prise de courant	19
tout pour le tout	19
besoin de travailler	18
bout du chemin	18
chef d'état	18
coût du chauffage	18
débat d'idée	18
fin de journée	18
jeu de rôle	18
niveau du confort	18
niveau de la facture	18

parent d'élève	18
période de crise	18
aménagement du temps	18
appel d'offre	17
chambre d'ami	17
chose simple	17
condition de vie	17
directeur du service	17
homme de gauche	17
idée de droite	17
mise en route	17
niveau des prix	17
niveau de la santé	17
plan de fidélisation	17
problème de chômage	17
problème d'emploi	17
question d'argent	17
vie de couple	17
argent du ménage	16
bout de l'année	16
chef d'agence	16
coup de fil	16
débat de fond	16
feu de cheminée	16
fonction de l'heure	16
groupe de travail	16
intérêt des gens	16
manque d'argent	16
milieu de travail	16
mois d'hiver	16
nombre d'année	16
période de froid	16
problème de logement	16
tas de gens	16
vie des gens	16
bouche à oreille	15
chambre du petit	15
chauffage de la maison	15
chef de service	15
coup d'œil	15
gens d'à côté	15
gens du pays	15
lieu de travail	15
manque de confiance	15
niveau de l'environnement	15
niveau de la famille	15
secteur d'activité	15
syndicat d'employeurs	15
vision des choses	15

Là encore, il est bon de rappeler que le recensement se fait sur les lemmes. Par exemple, "point" peut être au singulier ou au pluriel.

Le plus utilisé ("du point de vue" ou "au point de vue") est d'abord un quasi-synonyme de "à propos de", "sur le sujet de" (*au point de vue prix, confort, revenu...*). Secondairement, "un point de vue" signifie "opinion" mais en la relativisant (c'est plutôt une "vision des choses"). Par exemple, en disant "c'est mon point de vue", le locuteur affirme qu'il s'en tient à ce qu'il vient de dire, mais il admet implicitement qu'il peut y avoir d'autres opinions et qu'elles peuvent valoir la sienne. Beaucoup d'expressions courantes – comme "fin de mois" et "fin du mois" – ont ainsi plusieurs sens qu'il est possible de retrouver en utilisant le concordancier.

Pour le reste, on trouve dans ce tableau, les principaux thèmes des entretiens déjà évoqués. Le plus important est le temps – *de temps en temps, d'heure en heure, de jour en jour, heure du matin, du soir, mois de janvier*, etc. – suivi de la maison et de la vie quotidienne. Puis viennent la politique et les relations professionnelles. Mais il y a aussi des "tas de choses" ou ce "genre de truc" et beaucoup d'autres expressions toutes faites aisément mobilisables dans la conversation.

A titre comparatif, dans la littérature des cent dernières années, les dix syntagmes nominaux de ce type les plus fréquents sont (dans l'ordre) : *coup d'œil, temps en temps, quart d'heure, femme de chambre, tour à tour, point de vue, valet de chambre, nom de dieu, chemin de fer, coup de pied*. Quatre seulement figurent dans les usuels de la conversation mais seuls "de temps en temps" et "point de vue" figurent parmi les dix premiers dans les deux genres. C'est assez dire combien la littérature donne une image décalée du français usuel.

9.4 La coordination

D'un point de vue théorique, les éléments d'un groupe verbal ont autant besoin de coordination que ceux d'un groupe nominal. En pratique, la majorité des conjonctions de coordination est employée avec le nom et ses satellites (ce qui explique que nous la traitions dans cette partie consacrée au nom). Dès lors, on pourrait s'attendre à ce que, dans le français parlé, leur densité diminue comme celle des autres éléments du groupe du nom. Pourtant, la fréquence des conjonctions de coordination augmente de 17 % quand on passe de l'écrit au parlé. En effet, la coordination sert aussi à enchaîner les éléments constituant un propos et nous avons déjà souligné combien cette question est cruciale dans la conversation. C'est pourquoi, certaines conjonctions sont des outils privilégiés du français parlé (tableau 26).

Tableau 26. Les conjonctions de coordination dans le français parlé comparé à l'écrit (classement hiérarchique, fréquences en pour mille du total des occurrences des conjonctions de coordination)

Conjonctions	Parlé		Ecrit		Ecart %
	Rang	‰ Conjonc	Rang	‰ Conjonc	
et	1	472,9	1	691,2	-32
mais	2	279,1	2	175,7	59
donc	3	138,7	5	18,3	659
ou	4	93,0	3	66,1	41
ni	5	7,2	4	20,9	-66
soit	6	6,4	9	2,6	149
or	7	1,3	8	3,1	-56
car	8	1,2	6	16,5	-93
cependant	9	0,3	7	5,6	-95

Les conjonctions "et" et "ni" qui servent à construire des groupes (verbaux ou nominaux) complexes sont pour cette raison nettement sous-employées dans le français parlé qui privilégie "donc", "mais" et "ou". "Donc" est un cas d'école : à l'oral, il sert beaucoup à enchaîner les éléments d'un raisonnement - "A donc B" – au même titre que plusieurs autres vocables dont "parce que". En quelque sorte, le locuteur, pris dans le flux de la conversation, cherche à donner de la cohérence à son propos à l'aide de ces petits mots de liaison. Quant à "mais", et secondairement "soit", ils servent souvent à réagir aux propos tenus par d'autres, d'une manière contradictoire pour "mais" et concessive pour "soit !"

Les trois dernières (or, car, cependant) – assez rares à l'écrit – disparaissent presque complètement dans le français parlé, par ce phénomène de concentration sur les vocables les plus usuels que nous avons rencontré pour toutes les catégories grammaticales.

Conclusions

L'apprentissage d'une langue comme le FLE commence logiquement par l'oral et il est aujourd'hui acquis qu'il doit être fondé sur l'observation des locuteurs natifs dans les situations de la vie quotidienne et de la conversation. On rappellera à ce propos qu'il y a près de 70 ans, G. Gougenheim et ses collaborateurs avaient utilisé les premiers magnétophones pour enregistrer les Français dans les situations de la vie quotidienne et élaborer un dictionnaire et une grammaire de base (Gougenheim & Al. 1956). Malheureusement, pour de multiples raisons, cette piste n'a guère été poursuivie malgré l'importance apparente que l'on accorde aux enquêtes d'usage (Hug 2002).

Notre présentation donne quelques éclairages concernant cette importante question même si nos corpus – constitués pour de tout autres objectifs – ne répondent pas à certaines exigences des enquêtes d'usage. Il s'agissait de présenter les principales données accompagnées d'un minimum de commentaires et de conclusions. Nous reviendrons sur certains aspects remarquables dans des recherches ultérieures.

Naturellement, on peut faire certaines objections à ce rapport.

En premier lieu, dans les corpus utilisés pour cette présentation, les individus qui parlent et qui écrivent ne sont pas les mêmes et ne s'expriment pas dans le même cadre. Pour neutraliser l'influence respective du genre, du style personnel, de la situation et du thème, il faudrait par exemple, prendre le même homme, lui demander de rédiger des textes sur un sujet donné tout en enregistrant ses propos spontanés sur le même thème. Ce cas existe : il s'agit du général de Gaulle qui rédigeait toujours soigneusement ses discours et conférences de presse... sauf en décembre 1965, où il a improvisé une série d'émissions télévisées sur les mêmes questions politiques (Labbé 2005). La comparaison de ces émissions avec ses autres interventions révèle exactement les caractéristiques qui viennent d'être mises à jour : augmentation considérable du groupe du verbe — spécialement les verbes "faire", "dire", "aller", "falloir", de la première personne, du singulier et surtout du pluriel, et du pronom démonstratif "ce" —, augmentation tout aussi nette de "chose" et de "fait"... ; recul de certains pronoms relatifs et conjonctions propres à l'écrit, renforcement notable des occurrences de "mais" et "donc", etc. Il s'agit là d'une véritable expérience de laboratoire qui valide l'essentiel des conclusions présentées dans ce rapport.

En second lieu, on peut estimer que certaines des conclusions de cet article confirment des choses déjà connues. Mais elles changent de statut : elles étaient fondées sur l'intuition des chercheurs ou sur des recensements limités ; maintenant, elles sont chiffrées assez précisément

et, surtout, elles deviennent vérifiables grâce à la bibliothèque en ligne et aux corpus étiquetés que nous tenons à la disposition des chercheurs intéressés.

On sait donc que le français parlé privilégie le verbe au présent de l'indicatif et au passé composé, quelques pronoms et adverbes, la subordination plutôt que la coordination (sauf pour l'enchaînement des phrases). Il utilise un petit nombre de noms communs, qui couvrent des champs sémantiques larges, dont les archétypes sont "chose" et "gens". Au total à peine plus de 2 500 vocables – dont 163 verbes et 1 423 noms communs – suffisent pour accéder au français parlé de la conversation ordinaire.

Dès 1950, le statisticien P. Guiraud – à l'aide de quelques observations bien organisées - avait signalé que le nombre des substantifs et celui des verbes varient en sens inverse et que le nom domine dans la prose abstraite. A sa suite, on peut donc penser que l'effort d'élaboration qu'implique l'écriture s'accompagne d'un mouvement d'abstraction au-delà de la perception ou de la visée immédiate qui dominant la conversation. L'écriture s'accompagne aussi d'un effort de précision qui se traduit notamment par le remplacement des hyperonymes par des hyponymes (vocables dont le signifiant est plus précis, plus limité) et par la construction de groupes nominaux complexes et mieux coordonnés. Mais, pour écrire, il faut du temps, ressource dont on ne dispose pas dans la conversation !

Enfin, nous pouvons témoigner que les enseignants de français, spécialement de français langue étrangère, peuvent trouver dans les grands corpus, comme ceux utilisés pour cet article, un réservoir d'exemples et de tournures utiles pour leur enseignement.

Remerciements.

M. Becue, J.-G. Bergeron, M. Brugidou, J. et N. Leselbaum, S. Pionchon, E. Schweisguth nous ont confié les transcriptions d'entretiens ou de sondages qui forment l'essentiel du corpus "français parlé".

J. Savoy a relu une première version de ce rapport et ses remarques ont permis d'améliorer grandement le texte.

La bibliothèque électronique du français moderne, dont est extrait le corpus de la littérature contemporaine, est hébergée par l'équipe Sigma (Laboratoire d'Informatique de Grenoble).

Références

- Arrivé M., Gadet F. & Galmiche M. (1986), *La grammaire d'aujourd'hui. Guide alphabétique de linguistique française*, Paris : Flammarion.
- Arnold E. (2019). Le vocabulaire et le style du général de Gaulle. *Document Numérique*. (numéro spécial), 2019, p. 1–16.

- Arnold E., Labbé C., Labbé D. & Monière D. (2016). *Parler pour gouverner, Trois études sur le discours présidentiel français*. Grenoble, Laboratoire d'Informatique de Grenoble.
- Arnold E. & Labbé D. (2015). Vote for me. Don't vote for the other one. *Journal of World Languages*. January 2015, p. 1-18.
- Benveniste, E. (1956). La nature des pronoms. *Problèmes de linguistique générale I*. Paris : Gallimard, 1970, 251-257.
- Benveniste, E. (1959). Les relations de temps dans le verbe français. *Problèmes de linguistique générale I*. Paris : Gallimard, 1970, 237-250.
- Benveniste, E. (1965). Structure des relations d'auxiliarité. *Problèmes de linguistique générale II*. Paris : Gallimard, 1974, 177-193.
- Berger G. & Leselbaum N. dir. (2002), *La prévention des toxicomanies en milieu scolaire : éléments pour une évaluation*, Montpellier, CNDP.
- Bergeron, J.-G. & Bourque R. (2005). La formation et la pratique de la NBI au Québec. Esquisse d'un bilan. In Deschênes P., Bergeron J.-G., Bourque R., Briand A.(dir). *Négociations en relations de travail. Nouvelles approches*. Québec : Presses de l'Université du Québec, 66-99.
- Bertot, F. & Hassan R. (2015). *Didactique et enseignement de l'oral*. Paris : Publibook
- Blanche-Benveniste C. (2005). *Le français parlé. Études grammaticales*. Paris : CNRS.
- Blanche-Benveniste C. (2010). *Approches de la langue parlée en français*. Paris : Ophrys.
- Blanche-Benveniste C. & JeanJean C. (1987). *Le français parlé. Edition et transcription*. Paris : Didier-Erudition.
- Breckx M. (2017). *Grammaire française*. Louvain-la-Neuve : De Boeck.
- Burnard L. (1995). *Users Reference Guide for the British National Corpus*. Oxford : Oxford University Computing Service.
- Charaudeau P. (1992). "La modalisation et les modalités énonciatives", *Grammaire du sens et de l'expression*, Paris : Hachette, p. 569-629.
- Chevalier J.-C., Blanche-Benveniste C., Arrivé M. & Peytard J. (1964). *Grammaire Larousse du français contemporain*. Paris : Larousse.
- Cressot M. (1963). *Le style et ses techniques*. Paris: PUF.
- Cuq J.-P. & Gruca I. (2002). *Cours de didactique du français langue étrangère et seconde*. Grenoble : Presses Universitaires de Grenoble.
- Dubois J. (1965). *Grammaire structurale du français*. Paris : Larousse.
- Durand J., Laks B. & Lyche C. (2009). Le projet PFC (Phonologie du Français Contemporain) : une source de données primaires structurées. In *Phonologie, variation et accents du français*. Paris : Hermès.
- Frérot C. & Pecman M. (Dir.) (2021). *Des corpus numériques à l'analyse linguistique en langues de spécialité*. Grenoble : UGA Editions.
- Gadet F. (1989). *Le français populaire*. Paris : A. Colin.
- Gadet F. (1997). *Le français ordinaire*. Paris : Colin.

- Gougenheim G., Michea R., Rivenc P. & Sauvageot A. (1956). *L'élaboration du français élémentaire : étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Paris : Didier.
- Guiraud P. (1950). *Les caractères statistiques du vocabulaire*. Paris : PUF.
- Habert B., Nazarenko A. & Salem A. (1997). *Les linguistiques de corpus*. Paris : Armand Colin.
- Halliday M. A. K. (1985). *Spoken and written language*. Oxford : Oxford University Press.
- Hug M. (2002), *L'enquête d'usage linguistique*, Paris : Champion.
- Kerbrat-Orrecchioni C. (1981), *L'énonciation de la subjectivité dans le langage*. Paris : A. Colin.
- Labbé C. & Labbé D. (1994). *Que mesure la spécificité du vocabulaire ?* Grenoble : CERAT. Reproduit dans *Leximetrica*, 3, 2001.
- Labbé C. & Labbé D. (2013). La modalité verbale en français contemporain. Les hommes politiques et les autres. In Banks David. *La modalité, le mode et le texte spécialisé*. Paris : L'Harmattan, p. 33-61.
- Labbé D. (1990). *Normes de saisie et de dépouillement des textes politiques*. Cahier du CERAT n° 7. Grenoble : CERAT-IEP, avril 1990,
- Labbé D. (2001). Normalisation et lemmatisation d'une question ouverte. Les femmes face au changement familial, *Journal de la Société Française de Statistique*, 142-4, décembre 2001.
- Labbé D. (2002), *Analyse des représentations du confort électrique à partir d'un corpus d'entretiens*. GREST-EDF, Grenoble : CERAT, juin 2002.
- Labbé D. (2005). Le général de Gaulle en campagne. In Banks D. (Ed.). *Aspects linguistiques du texte de propagande*. Paris : L'Harmattan, p. 213-233.
- Labbé D. (2007). Coordination et subordination en français oral. In Banks D. *La coordination et la subordination dans le texte de spécialité*. Paris : L'Harmattan, p. 161-182.
- Lafon P. (1980). Sur la variabilité des formes dans un corpus. *Mots*, 1, octobre 1980, p 127-165.
- Lafon P. (1984). *Dépouillements et statistiques en lexicométrie*. Genève-Paris: Slatkine-Champion.
- Mandelbrot B. (1957). Étude de la loi d'Estoup et de Zipf. Fréquences des mots dans le discours. In Apostel L. et al. *Logique, langage et théorie de l'information*. Paris, PUF, p. 22-53.
- Monière D. & Labbé D. (2018). Le vocabulaire des campagnes électorales. In Iezzi D. F., Celardo L., Misuraca M. *Proceedings of the 14th International Conference on Statistical Analysis of Textual Data*. Roma: UniversItalia, 2018, p. 522-540.
- Muller C. (1963). Le mot, unité de texte et unité de lexique en statistique lexicologique. In Muller C. *Langue française et linguistique quantitative*. Genève-Paris : Slatkine-Champion, 1979, p 125-143.
- Muller C. (1977). *Principes et méthodes de statistique lexicale*. Paris, Hachette.

- Paternostro R. (2022). Français oral ou français parlé ? Quelles implications pour l'enseignement ? *Mélanges CRAPEL*, 43/1, p 65-100.
- Parisse C., Benzitoun C., Etienne C. & Liégeois L. (2021). Agrégation automatisée de corpus de français parlé. In Frérot C. et Pecman M. (Dir.) (2021). *Des corpus numériques à l'analyse linguistique en langues de spécialité*. Grenoble : UGA Editions, p. 61-86.
- Pionchon S. (2001), *Les Françaises et la politique*, Thèse pour le doctorat de science politique. Grenoble : Institut d'Etude Politique.
- Ravazzolo E., Traverso V., Jouin E. & Vigner G. (2015). *Interactions, dialogues, conversations : l'oral en français langue étrangère*. Paris : Hachette.
- Riegel M., Pellat J.-C. & Rioul R. (1994). *Grammaire méthodique du français*. Paris : PUF.
- Savoy, J. (2020). *Machine Learning Methods for Stylometry. Authorship Attribution and Author Profiling*. Cham : Springer Nature.
- Schweisguth E. (1986). Les avatars de la dimension gauche/droite. In Dupoirier Elisabeth & Grunberg Gérard. *La drôle de défaite de la gauche*. Paris : PUF, p. 51-70.
- Sinclair J. (1991). *Corpus, Concordances, Collocations*. Oxford : Oxford University Press.
- Surcouf C. (2021). Le français oral quotidien : un objectif spécifique en FLE ? Retour sur les défis de la création d'un corpus de français parlé annoté à visée pédagogique. In Frérot C. et Pecman M. (Dir.) (2021). *Des corpus numériques à l'analyse linguistique en langues de spécialité*. Grenoble : UGA Editions, p 107-133.
- Troncy C., Carette E. & Miras G. Eds (2022). Enseignement du français parlé aujourd'hui : Recherches et expériences de terrain. *Mélanges CRAPEL*, n° 43/1.
- Wagner R.-L. & Pichon J. (1962). *Grammaire du français classique et moderne*. Paris : Hachette.
- Williams G. (dir.) (2005). *La linguistique de corpus*. Rennes : Presses universitaires de Rennes.
- Weber C. (2013). *Pour une didactique de l'oralité. Enseigner le français tel qu'il est parlé*. Paris : Didier.
- Zipf G. K. (1935). *La psychobiologie du langage*. Paris : CEPL, 1974.

Annexe 1.
Le français oral

Les Français et la politique (Schweisguth 1986) : 64 entretiens, 722 888 mots et 10 475 vocables
Les Français(es) et la politique (Pionchon 2001) : 32 entretiens, 345 752 mots, 6 540 vocables
La négociation raisonnée au Québec (Bergeron & Bourque, 2003) : 61 entretiens, 409 225 mots, 6 591 vocables
La prévention des toxicomanies en milieu scolaire (Berger & Leselbaum, 2002) : 15 entretiens, 92 992 mots, 4 255 vocables
Le confort électrique réalisé par les sociologues d'EDF en six enquêtes (1994-1999) (Labbé 2002) : 201 entretiens, 1 270 307 mots, 10 904 vocables
La formation du jugement politique (élection présidentielle 2007) : 44 entretiens, 389 707 mots, 8 058 vocables
Divers : 6 entretiens réalisés au début des années 2000 : 115 494 mots, 4 922 vocables
Questions ouvertes dans un sondage auprès des femmes divorcées réalisé par l'INED (Labbé, 2001) : 2 329 enquêtes, 56 107 mots, 2 786 vocables
Questions ouvertes dans un sondage auprès des citoyens belges sur la droite et la gauche : 760 enquêtes, 22 294 mots, 1 706 vocables différents
Questions ouvertes dans un sondage à propos d'un grand chantier d'EDF (2004) : 980 enquêtes, 107 487 mots 2 775 vocables
Questions ouvertes dans un sondage sur la formation du jugement politique (2007) : 1 467 enquêtes, 142 284 mots, 3 695 vocables
Total transcriptions de l'oral :
433 entretiens et quatre sondages : 3 634 737 mots.

Annexe 2
Le français écrit dans la littérature contemporaine
(Auteur, nombre de livres, nombre de mots)

Arnaud Georges-Jean, 2, 88 170 ; Audiard Michel, 2, 117 549 ; Bazin Hervé, 3, 216 720 ; Benacquista Tonino, 3, 208 885 ; Bernanos Georges, 3, 280 892 ; Blondin Antoine, 3, 147 258 ; Boileau Pierre-Louis et Narcejac Thomas, 3, 133 110 ; Bussi Michel, 2, 275 947 ; Camus Albert, 3, 155 890 ; Céline Louis-Ferdinand, 2, 430 804 ; Charles Frédéric, 3, 92 098 ; Dabit Eugène, 2, 81 441 ; Dard Frédéric, 8, 454 678 ; Darrieusecq Marie, 2, 102 114 ; Gaulle Charles de, 4, 515 421 ; Desportes Virginie, 2, 115 039 ; Djian Philippe, 1, 113 536 ; Echenoz Jean, 3, 175 822 ; Gary Romain (et Ajar Emile), 13, 865 900 ; Gavalda Anna, 2, 165 797 ; Gide André, 6, 322 873 ; Giebel Karine, 3, 264 193 ; Giono Jean, 3, 276 561 ; Giraudoux Jean, 7, 205 903 ; Grangé Jean-Christophe, 3, 354 771 ; Hélène André, 3, 281 783 ; Houellebecq Michel, 4, 362 528 ; Le Clézio Jean-Marie, 12, 870 310 ; Lemaitre Pierre, 6, 618 532 ; Makine Andreï, 2, 164 691 ; Malet Léo, 5, 275 989 ; Malraux André, 4, 361 519 ; Manchette Jean-Patrick, 1, 42 304 ; Martin du Gard Roger, 3, 175 874 ; Mauriac François, 3, 146 302 ; Minier Bernard, 2, 341 283 ; Modiano Patrick, 2, 92 109 ; Musso Guillaume, 2, 196 123 ; Nothomb Amélie, 4, 145 656 ; Pennac Daniel, 7, 377 460 ; Queneau R, 1, 12 816 ; Radiguet Raymond, 2, 66 839 ; Sagan Françoise, 1, 30 083 ; Saint-Exupéry Antoine de, 3, 92 392 ; Simenon Georges, 7, 507 909 ; Simonin Albert, 3, 181 362 ; Thilliez Franck, 4, 523 308 ; Vargas Fred, 5, 403 182 ; Vian Boris, 5, 245 786 ; Yourcenar Marguerite, 2, 203 506.
Soit : 51 auteurs, 134 livres, 12 881 018 mots

Annexe 3

Les 1420 noms les plus utilisés dans le français parlé (classement alphabétique).

aberration, abonnement, absence, absentéisme, abus, accent, accès, accident, accord, accréditation, accueil, accumulateur, accumulation, achat, acquis, acte, acteur, action, actionnaire, activité, actualité, adjoint, administration, ado, adresse, adulte, aération, affaire, âge, agence, agent, agriculteur, agricultrice, agriculture, aide (nm et nf), air, aise, ajustement, alarme, alcool, alcoolisme, alentour, alimentation, allocation, alternance, ambiance, ambition, amélioration, aménagement, ami, amie, ampère, ampoule, an, analyse, ancien, ancienneté, anglais, animal, animateur, animation, année, août, appareil, appart, appartement, appel, application, appoint, apport, apprentissage, approche, après-midi, arbre, architecte, argent, argument, arme, armée, armoire, arrêt, arrivée, article, artisan, aspect, aspirateur, assemblée, assistante, association, assurance, atelier, atmosphère, atout, attente, attention, attitude, augmentation, automne, autorité, autoroute, avance, avant, avantage, avenir, avion, avis, avocat, avortement, avril, bac, bagarre, bain, baisse, balance, balcon, balle, ballon, banlieue, banque, barre, barrière, bas, base, bataille, bateau, bâtiment, beau-frère, beaux-parents, bébé, belle-mère, bénéfice, besoin, bête, bêtise, béton, biais, bien, bien-être, bilan, blanc, bleu, bloc, blocage, bois, boisson, boîte, boîtier, bon, bonheur, bonhomme, bonjour, bord, boss, bouche, boulot, bouquin, bourse, bout, bouteille, bouton, brainstorming, branchement, bras, brevet, bricole, brique, bruit, budget, bureau, bus, but, cabinet, câble, cadeau, cadre, café, cahier, caisse, calcul, camion, camp, campagne, cancer, candidat, candidate, cantine, canton, capacité, capital, capitalisme, capteur, caractère, carrelage, carrière, carte, cas, cassette, catalogue, catastrophe, catégorie, catholique, cause, cave, centaine, centime, centimètre, centrale, centre, céréale, certificat, chaîne, chaleur, chambre, champ, chance, changement, chantier, charge, chargé, chasse, chaud, chaudière, chauffage, chauffe-eau, chef, chemin, cheminée, chèque, cheval, chien, chiffre, choix, chômage, chômeur, chose, cigarette, cinéma, circonstance, circuit, circulation, cité, citoyen, classe, clause, clé, client, clientèle, clim, climat, climatisation, club, coeur, coin, colère, collectif, collectivité, collège, collègue, combat, combattant, comité, commande, commerçant, commerce, commercial, commission, commun, communauté, commune, communication, communisme, communiste, compagnie, comparaison, compétence, comportement, compresseur, compromis, comptabilité, compte, compteur, con, conception, concertation, conciliateur, conciliation, conclusion, concours, concurrence, condensation, condition, conduite, confiance, conflit, confort, confrontation, congé, congélateur, connaissance, connerie, conscience, conseil, conseiller, conseilère, consensus, conséquence, considération, consigne, consommateur, consommation, constitution, constructeur, construction, contact, contenu, contexte, contrainte, contraire, contrat, contremaître, contrôle, convecteur, convention, conversation, conviction, copain, copine, corps, côté, couche, couleur, couloir, coup, couple, coupure, cour, courage, courant, courrier, cours, course, coût, couverture, crainte, création, crédit, crise, critère, critique, cuisine, cuisinière, culture, cumulus, curé, cuve, dalle, dame, danger, date, débat, début, décalage, décembre, déception, décision, défaut, défense, déficit, définition, dégât, dégradation, degré, délai, délégué, délestage, délesteur, délinquance, demande, démarche, demie, demi-heure, demi-journée, démocratie, départ, département, dépense, déplacement, dépôt, député, dernier, dernière, dessous, dessus, détail, détriment, dette, dévaluation, développement, devis, devoir, diagnostic, dialogue, dictature, dieu, différence, difficulté, dimanche, dimension, diplôme, directeur, direction, directrice, dirigeant, discipline, discours, discussion, disjoncteur, disposition, dispute, distance, distribution, division, divorce, dizaine, docteur, document, documentation, doigt, dollar, domaine, domicile, dommage, domotique, donnée, dos, dossier, double, douche, doute, drogue, droit, droite, durée, eau, écart, échange, échéance, échec, échelle, échelon, éco, école, écologie, économie, écoute, écran, éducateur, éducation, effectif, effet, efficacité, effort, égalité, église, électeur, élection, électricien, électricité, électroménager, électronique, élément, élève, élu, embauche, émission, emplacement, emploi, employé, employeur, emprunt, encadrement, endroit, énergie, enfance, enfant, engagement, enjeu, ennui, enquête, enquêté, enseignant, enseignement, ensemble, entente, entourage, entraide, entrée, entreprise, entretien, enveloppe, envie, environnement, époque, épouse, équilibre, équipe, équipement, erreur, escalier, espace, espèce, espoir, esprit, essai, essence, essentiel, est, esthétique, estimation, établissement, étage, étape, état, été, étiquette, étoile, étranger, être, étude, étudiant, étudiante, euro, évaluation, événement, évolution, examen, excès, exécutif, exemple, exercice, exigence, expérience, explication, exploitation, expression, extérieur, extrême, extrême droite, extrême gauche, extrémiste, fabricant, fabrication, face, facilitateur, facilité, façon, facteur, facturation, facture, faculté, faillite, faim, fait, famille, faute, faveur, fédération, femme, fenêtre, fer, ferme, fête, feu, feuille, février, fidélisation, figure, fil, filière, fille, film, fils, filtre, fin, final, finance, financement, foi, foire, fois, folie, fonction, fonctionnaire, fonctionnement, fond, fonds, force, forfait, formation, forme, formule, fortune, four, fournisseur, foyer, fraîcheur, frais, franc, français, frein, frère, fric, frigidaire, frigo, froid, front, frontière, fruit, fuel, fuite, fusible, gain, gamin, gamme, garage, garantie, garçon, gardien, gars, gauche, gaz, gel, gendarme, génération, genre, gens, gestion, gestionnaire, gosse, goût, gouvernement, grâce, grand, grand-chose, grand-mère, grève, grief, grille, groupe, guerre, gueule, guillemet, habitant, habitation, habitude, halogène, handicap, hasard, haut, hauteur, heure, histoire, hiver, homme, hôpital, horaire, horloge, horreur, hors-gel, hôtel,

humeur, humidité, idéal, idée, identité, idéologie, illusion, image, immeuble, immigration, immigré, impact, importance, impôt, impression, incidence, incompatibilité, inconvenient, indemnité, indépendance, individu, industrie, inégalité, inertie, infidélité, infirmière, inflation, influence, information, informatique, infrastructure, ingénieur, initiative, inquiétude, insécurité, insertion, instabilité, installateur, installation, instance, instant, instituteur, institutrice, institution, intention, intérêt, intérieur, intérim, interlocuteur, intermédiaire, interne, internet, interprétation, intervenant, intervention, inversion, investissement, isolation, jalousie, jambe, janvier, jardin, jeu, jeudi, jeune, jeunesse, job, jour, journal, journaliste, journée, jugement, juillet, juin, justice, kilo, kilomètre, kilowatt, label, laine, lait, lampe, langage, langue, lave-linge, lave-vaisselle, lecture, législative, lendemain, lessive, lettre, liaison, libéral, libération, liberté, licence, lien, lieu, ligne, limite, linge, liste, lit, litre, livre (nf et nm), local, locataire, location, logement, logique, loi, loisir, long, longueur, lotissement, loyer, lumière, lundi, lutte, luxe, lycée, machin, machine, madame, magasin, magnétoscope, mai, main, main-d'oeuvre, maintien, maire, mairie, maison, maître, maîtresse, maîtrise, majorité, mal, malade, maladie, malheur, maman, mandat, manière, manifestation, manque, marche, marché, mardi, marge, mari, mariage, marketing, marque, marre, mars, masse, matériel, matière, matin, maximum, mazout, mec, mécanique, mécanisme, médecin, médecine, média, médiateur, médiation, mélange, membre, mémoire, ménage, mentalité, mer, merci, mercredi, merde, mère, mérie, mécontente, message, messe, mesure, méthode, métier, mètre, meuble, micro-onde, midi, milieu, militant, milliard, million, mine, minimum, ministère, ministre, minitel, minorité, minuit, minute, miracle, mise, misère, mission, mi-temps, mode (nf et nm), modèle, modification, mois, moitié, môme, moment, monde, mondialisation, monnaie, monsieur, montagne, montant, morceau, mort (nf et nm), mot, moteur, motif, motivation, mouvement, moyen, moyenne, municipale, municipalité, mur, musique, naissance, nation, nationalité, nature, nécessité, négociateur, négociation, neige, nez, niveau, noir, nom, nombre, nord, norme, note, notice, notion, nourriture, nouvelle, novembre, nucléaire, nuisance, nuit, numéro, objectif, objet, obligation, occasion, occupation, octobre, oeil, oeuvre, office, offre, opération, opinion, opposition, option, optique, orage, ordinateur, ordre, oreille, organisation, organisme, orientation, origine, ouest, outil, ouverture, ouvrage, ouvrier, ouvrière, page, paie, paiement, pain, paix, panne, panneau, papa, papier, paquet, paramètre, parc, parcours, pardon, parent, parenthèse, parité, parole, part, partage, partenaire, partenariat, parti, participation, particulier, partie, pas, passage, passé, patrimoine, patron, patronat, patte, pauvre, pavillon, pays, paysage, paysan, peau, peine, peinture, pensée, pension, perception, père, période, permanence, permis, personnage, personnalité, personne, personnel, perte, petit, petit-déjeuner, pétrole, peuple, peur, phase, phénomène, philosophie, photo, phrase, picot, pièce, pied, pierre, pile, piscine, piste, placard, place, plafond, plage, plaisir, plan, plancher, planète, plaque, plat, plateau, plombier, plupart, poche, poêle, pognon, poids, point, police, politicien, politique (nf et nm), pollution, pompe, population, porte, porte-parole, pose, position, possibilité, poste (nf et nm), pot, poteau, poubelle, pourcentage, poussière, pouvoir, pratique, prélèvement, préoccupation, préparation, présence, présent, présentation, président, présidente, présidentielle, presse, pression, prêt, preuve, prévention, prime, principal, principe, printemps, priorité, prise, prison, privé, privilège, prix, problématique, problème, procédure, processus, producteur, production, produit, prof, professeur, profession, professionnel, profil, profit, programmeur, programmation, programme, progrès, projet, promesse, promotion, proportion, propos, proposition, propriétaire, propriété, protection, province, proximité, pub, public, publicité, puissance, pull, pylône, qualification, qualité, quantité, quart, quartier, question, questionnaire, quinzaine, quota, quotidien, racisme, radian, radiateur, radio, rafraîchissement, raison, raisonnement, rappel, rapport, ras-le-bol, rayon, réaction, réalisation, réalité, recette, recherche, reconnaissance, recul, réduction, référence, référendum, réflexe, réflexion, réforme, refus, regard, régime, région, réglage, règle, règlement, regroupement, régulation, relais, relation, relevé, religion, remarque, remise, remplacement, rencontre, rendement, rendez-vous, renouvellement, rénovation, renseignement, rentrée, répartition, repas, repassage, repère, réponse, représentant, représentation, reprise, reproche, république, réseau, réserve, résidence, résistance, résolution, respect, responsabilité, responsable, ressource, restaurant, reste, résultat, retard, retour, retraite, retraité, réunion, réussite, rêve, revenu, révolution, riche, richesse, rigueur, risque, rôle, rond, roue, rouge, route, rue, rythme, sac, saison, salaire, salarié, salle, salon, samedi, sanction, santé, satisfaction, séance, seconde, secours, secrétaire, secteur, sécurité, sein, séjour, semaine, sens, sensation, sensibilité, sentiment, séparation, septembre, série, service, serviette, session, sexe, siècle, signature, site, situation, ski, social, socialisme, socialiste, société, soeur, soin, soir, soirée, sol, soleil, solidarité, solution, somme (nm et nf), sondage, sonde, sorte, sortie, sou, souci, soufflerie, source, sous-sol, soutien, souvenir, spectacle, sport, stade, stage, standard, statistique, statut, stop, stratégie, structure, style, subvention, succursale, sud, suite, suivi, sujet, supermarché, surface, surprise, syndicat, système, tabac, table, tableau, tâche, tarif, tarification, tas, taux, taxe, taxi, technicien, technicienne, technique, technologie, télé, télécommande, téléphone, télévision, témoin, température, temporaire, temps, tendance, tension, tentative, terme, terrain, terrasse, terre, test, tête, texte, thème, théorie, thermomètre, thermostat, tiers, titre, toilette, toit, tort, total, totalité, touche, tour (nm, nf), tourisme, tournée, tout, tradition, train, traitement, tranche, transport, travail, travailleur, travers, tribunal, trou, truc, tube, tuyau, type, union, unité, université, untel, urgence, usage, usine, utilisation, utilité, vacance, vache, vaisselle, valeur, véhicule, veille, vélo, vendeur, vendredi, vent, vente, ventilateur, ventilation, vérité, verre, vêtement, viande, vice-président, vie, vieux, villa, village, ville, vin,

vingtaine, violence, virgule, vision, visite, vitesse, vitrage, voie, voisin, voisinage, voisine, voiture, voix, vol, volet, volonté, volume, vote, voyage, vue, watt, week-end, zone

Annexe 4

Les 753 adjectifs les plus utilisés dans le français parlé (classement alphabétique)

aberrant, absent, absolu, acceptable, accessible, acquis, actif, actuel, adapté, adéquat, adjoint, administratif, aérien, affectif, affreux, africain, âgé, agréable, agressif, agricole, aimable, aîné, aisé, alcoolique, alimentaire, allemand, allumé, ambiant, ambitieux, américain, ami, amusant, ancien, anglais, annuel, anonyme, anormal, antérieur, antillais, apparent, appréciable, apte, arrêté, arrivé, artificiel, assis, associatif, athée, atomique, atroce, attentif, automatique, automobile, autonome, autoritaire, autre, avantageux, bas, basé, basique, battu, beau, belge, bénéfique, bête, bizarre, blanc, bleu, bon, bourgeois, bref, brillant, brut, brutal, bruyant, budgétaire, caché, calme, capable, capitaliste, caractériel, carré, catastrophique, catégorique, catholique, célibataire, censé, central, centriste, certain, changé, charmant, chaud, chauffant, chauffé, cher, chiant, choisi, choquant, chrétien, citoyen, civil, civique, clair, classique, cohérent, coïncé, collectif, commercial, commode, commun, communal, communiste, comparable, comparé, compatible, compétent, compétitif, complémentaire, complet, complexe, compliqué, compréhensible, compréhensif, compris, comptable, con, concerné, concret, confédéral, confiant, conflictuel, conforme, confortable, conjoint, conjugal, connu, conscient, conséquent, conservateur, considérable, constant, constructif, construit, content, continu, contradictoire, contraignant, convaincant, convaincu, conventionnel, cool, correct, coté, coupé, courageux, courant, coureur, court, coûteux, couvert, crédible, creux, croyant, culturel, curieux, dangereux, décent, décevant, déçu, défavorisé, défini, définitif, dégoûtant, dégueulasse, délicat, démagogique, demi, démocratique, démuni, départemental, dépendant, dépensier, dépressif, dérisoire, dernier, désagréable, déterminant, déterminé, développé, différent, difficile, digital, dingue, direct, directeur, discret, disponible, distributif, dit, divergent, divers, domestique, donné, double, doux, dramatique, droit, drôle, dû, dur, dynamique, écolo, écologique, écologiste, économe, économique, écrit, éducatif, efficace, effrayant, égal, égoïste, élargi, électoral, électrique, électroménager, électronique, élevé, éligible, élu, embêtant, enceinte, encombrant, énergétique, engagé, ennuyeux, énorme, entendu, enterré, entier, épais, épouvantable, équilibré, équitable, équivalent, espagnol, essentiel, esthétique, estimé, établi, étonnant, étranger, étroit, étudiant, européen, éventuel, évident, évolué, exact, excellent, exceptionnel, excessif, exigeant, existant, exploratoire, extérieur, externe, extraordinaire, extrême, extrémiste, fabuleux, facile, faible, faisable, fait, fameux, familial, fantastique, farfelu, fatigant, faux, favorable, féminin, féministe, ferme, fermé, fiable, fidèle, fier, fin, final, financier, fiscal, fixe, flagrant, flou, foncier, fonctionnel, fondamental, forcé, formel, formidable, fort, fou, fragile, frais, franc, français, fréquent, frileux, froid, futur, gagnant, gamin, gauche, gaulliste, gênant, général, généreux, génial, gentil, géographique, global, gouvernemental, grand, gratuit, grave, gris, gros, grossier, habitable, habituel, handicapé, haut, heureux, historique, honnête, honteux, horaire, horrible, hospitalier, humain, humide, idéal, identique, idiot, immédiat, immense, immigré, immobilier, impeccable, impliqué, important, impossible, inadmissible, inaudible, incapable, incroyable, indépendant, indifférent, indispensable, individualiste, individuel, industriel, inférieur, infernal, informatique, informé, initial, injuste, inquiet, inquiétant, insatisfait, instable, installé, institutionnel, insuffisant, insupportable, intégré, intellectuel, intelligent, intéressant, intéressé, intérieur, intermédiaire, international, interne, inutile, inverse, invivable, isolé, issu, italien, jaloux, japonais, jaune, jeune, joli, juridique, juste, laitier, lamentable, large, légal, léger, législatif, légitime, libéral, libre, lié, limité, local, locatif, logique, long, lourd, lucide, lyonnais, maghrébin, magnifique, majeur, majoritaire, mal, malade, malheureux, malhonnête, malin, malsain, mandant, manuel, marginal, marié, marrant, masculin, matériel, maternel, mauvais, maximum, mécanique, méchant, mécontent, médiatique, médical, méfiant, meilleur, même, ménager, mensuel, mental, merveilleux, mesquin, militaire, militant, mince, mineur, minime, minimum, mis, mixte, moche, moderne, modeste, moindre, mondial, monétaire, monstrueux, moral, motivant, motivé, moyen, multiple, municipal, mutuel, naïf, national, nationalisé, naturel, nécessaire, néfaste, négatif, négligeable, nerveux, net, neuf, neutre, noble, nocif, noir, nombreux, normal, normatif, nouveau, nucléaire, nul, objectif, obligatoire, obligé, occasionnel, occidental, officiel, onéreux, opposé, optimiste, ordinaire, organisé, originaire, ouvert, ouvrier, parallèle, pareil, parfait, parisien, paritaire, parlant, parti, participatif, particulier, partiel, partisan, passé, passionnant, patronal, pauvre, payant, payé, paysan, pédagogique, péjoratif, pénible, perdant, perdu, performant, permanent, personnel, pertinent, pessimiste, petit, phénoménal, philosophique, physique, pilote, pire, placé, plat, plein, pointu, politicien, politique, polonais, ponctuel, populaire, porteur, positif, possible, pourri, poussé, pratiquant, pratique, précaire, précédent, précis, préférable, présent, présidentiel, prêt, prévu, primaire, primordial, principal, prioritaire, pris, privé, privilégié, problématique, prochain, proche, professionnel, profond, programmable, progressif, propre, propriétaire, provincial, prudent, psychologique, public, puissant, pur, qualifié, quelconque, quitte, quotidien, raciste, radiant, radical, raisonnable, raisonné, rapide, rare, ras, rassurant, rayonnant, réalisable, réaliste, récent, réduit, réel, réfractaire, régional, réglé, regrettable, régulier, relatif, relationnel, religieux, rendu, rentable, réparti, représentatif, résidentiel, respectif, responsable, restreint, réversible, révolutionnaire, riche, ridicule, rigide, rigolo,

rigoureux, rond, rose, rouge, roulant, routier, rural, sacré, sain, salarial, sale, sanitaire, satisfaisant, satisfait, sauvage, scandaleux, sceptique, scientifique, scolaire, sec, secondaire, sectaire, semblable, sensationnel, sensé, sensible, séparé, sérieux, seul, sévère, sexuel, significatif, similaire, simple, sincère, social, socialiste, soi-disant, solaire, solidaire, solide, sophistiqué, souhaitable, soûl, souple, souterrain, soviétique, spécial, spécialisé, spécifique, spontané, sportif, stable, stratégique, strict, stupide, suédois, suffisant, suivant, super, superbe, superficiel, supérieur, supplémentaire, supportable, sûr, surprenant, susceptible, sympa, sympathique, syndical, syndiqué, systématique, tarifaire, technicien, technique, technologique, téléphonique, télévisé, temporaire, tenu, terminé, terrible, théorique, thermique, tiers, tolérant, total, touristique, traditionnel, tranquille, transparent, triste, typique, uniforme, unique, universel, universitaire, urbain, urgent, utile, utopique, vague, valable, variable, vaste, verbal, véritable, vert, vicieux, vide, vieux, vif, vigilant, violent, visuel, vital, vivable, vivant, voisin, volontaire, vrai, vu