



Adjectifs épithètes alternants en français parlé : premiers résultats

Alternating Attributive Adjectives in Spoken French: first results

Christophe Benzitoun



Édition électronique

URL : <https://journals.openedition.org/tipa/585>
DOI : 10.4000/tipa.585
ISSN : 2264-7082

Éditeur

Laboratoire Parole et Langage

Référence électronique

Christophe Benzitoun, « Adjectifs épithètes alternants en français parlé : premiers résultats », *TIPA. Travaux interdisciplinaires sur la parole et le langage* [En ligne], 29 | 2013, mis en ligne le 19 décembre 2013, consulté le 15 avril 2024. URL : <http://journals.openedition.org/tipa/585> ; DOI : <https://doi.org/10.4000/tipa.585>

Ce document a été généré automatiquement le 15 avril 2024.



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

Adjectifs épithètes alternants en français parlé : premiers résultats

Alternating Attributive Adjectives in Spoken French: first results

Christophe Benzitoun

1. Introduction

- 1 La place de l'adjectif épithète par rapport au substantif auquel il se rapporte est un sujet qui a donné lieu à une littérature conséquente en linguistique française. Pour s'en rendre compte, il suffit de consulter, entre autres, les deux articles de synthèse sur la question (Delomier, 1980 ; Forsgren, 1997), le livre de Reiner (1968) comportant, déjà à l'époque, vingt-et-une pages de bibliographie, ainsi que les ouvrages de Forsgren (1978) et Larsson (1994). Une conclusion largement partagée est celle de Reiner (1968 : 4) qui dit que « les deux constructions *adjectif + substantif* et *substantif + adjectif* [seraient], le plus souvent, également possibles ». Wagner & Pinchon (1962) en tirent la même conclusion et Goes (1999) fait même de la possibilité d'alternance par rapport au nom un critère d'adjectivité. Mais il existe tout de même des règles et des contraintes difficiles à acquérir pour les non francophones :

S'il fallait énumérer les principaux domaines de la syntaxe faisant problème à l'étudiant de français langue étrangère, notamment de souche nordique, il est certain [que] [...] serait aussi mentionnée la place de l'adjectif épithète. Redoutable pour sa complexité, le problème de la position variable – globalement parlant – de l'adjectif ne cesse d'attirer [...] l'intérêt des grammairiens et des linguistes. (Forsgren, 1997 : 115)

- 2 Pour étudier un tel sujet, l'utilisation du corpus s'est rapidement imposée, accompagnée (ou non) d'exemples inventés par introspection. Le nombre d'occurrences pris en compte s'échelonne entre 2.000 (Glatigny, 1967) et 29.016 (Wilmet, 1981) exclusivement issues de corpus. Les données prises en compte sont très majoritairement des textes émanant de la littérature et de la presse (de l'écrit donc), exception faite de Larsson (1994), qui a élargi la description en y incluant des

ressources provenant de catalogues de voyage, de guides touristiques et d'autres types de prose non littéraire. Les relevés ont été faits manuellement, ce qui a nécessité un travail long et fastidieux.

- 3 Très tôt, la question des contraintes, catégoriques et préférentielles, a été au centre des préoccupations des linguistes pour rendre compte de la place des adjectifs. Certains linguistes ont même proposé des théories explicatives générales permettant selon eux de rendre compte de l'ensemble des exemples possibles à partir d'une ou plusieurs contraintes catégoriques. Cependant, comme le note Larsson (1994 : 229-230), aucune théorie n'est véritablement parvenue à remplir l'objectif que son auteur lui a assigné. D'où une tendance récente à s'orienter presque exclusivement vers les contraintes préférentielles plutôt que les règles rigides (Thuilier, 2012 ; Fox, 2012). Dans ce paradigme, les contraintes préférentielles peuvent être vues comme des paramètres ayant une influence non systématique sur la place. L'influence de ces paramètres est visible à travers les fréquences observées. Forsgren (1978) a par exemple montré que la forme du déterminant avait un impact sur l'ordre relatif entre adjectif et substantif. Voici une liste non exhaustive de ces contraintes préférentielles mises en évidence dans les études antérieures : sens de l'adjectif, ellipse de *être*, nombre (singulier/pluriel), présence d'autres adjectifs, présence d'un modifieur/adverbe pré- ou post-adjectival, forme du substantif constructeur, emplois spécifiques/figements, coordination avec une autre unité, nature du déterminant, etc.
- 4 Dans une étude récente, Thuilier (2012) s'est attelée à quantifier précisément l'influence d'un ensemble de contraintes préférentielles en prenant appui sur une approche basée principalement sur l'extraction automatique de données à partir du *French Treebank* (Abeillé et al., 2003), corpus annoté en syntaxe et composé d'articles du quotidien *Le Monde*. Cette étude montre clairement l'apport décisif de ressources annotées syntaxiquement. Cela permet d'extraire rapidement une grande quantité d'exemples et rend possible des analyses difficiles à effectuer manuellement eu égard au volume d'exemples et de paramètres à traiter. Les résultats obtenus à partir du *French Treebank* ont été complétés par des questionnaires psycholinguistiques ainsi que d'autres ressources telles que *l'Est Républicain* (quotidien régional lorrain), *ESTER* (corpus oral radiophonique) et *C-ORAL-ROM* (corpus oral comprenant des situations de parole variées).
- 5 Concernant spécifiquement le français parlé, il existe les études de Blasco-Dulbecco & Cappeau (2004 ; 2005) ainsi qu'un bref passage dans Blanche-Benveniste (2010). Mais ces travaux ne sont pas centrés sur la place des adjectifs épithètes. Ils abordent la question du fonctionnement des adjectifs en général et contiennent seulement quelques remarques sur la place. Fox (2012), quant à elle, s'est intéressée notamment à l'acquisition de la place des adjectifs chez de jeunes enfants francophones à partir de corpus longitudinaux.
- 6 Comme on peut le voir, la majorité des études précédemment citées se concentrent exclusivement sur le français écrit ou ne sont pas centrées sur la problématique de la place de l'adjectif dans une perspective grammaticale, à l'exception de Thuilier (2012). Cependant, même si l'auteur délivre les résultats chiffrés obtenus à partir des données orales, ceux-ci sont exploités de manière embryonnaire pour l'instant. Mais dans le présent volume, J. Thuilier se propose justement d'approfondir ses analyses sur l'oral en modélisant et quantifiant l'impact des contraintes préférentielles. Dans ce qui suit, nous menons une étude complémentaire à celle de Thuilier (ce volume) en privilégiant

une analyse plus qualitative de chaque adjectif et en observant un corpus plus volumineux.

- 7 Pour résumer, on peut donc dire que, malgré une activité de publication foisonnante, il existe peu d'études sur la place des adjectifs épithètes en français parlé non planifié et aucune description systématique. Or, nous pensons qu'il est fondamental de prendre en compte des données orales car elles permettent de dégager les « faits majeurs de la distribution, qui [sont] ainsi moins « parasités » par des phénomènes d'ordre stylistique (et donc plus atypiques) » (Cappeau, 2002 : 11). Et, comme l'ont montré les études antérieures, la variable stylistique est un paramètre important pour rendre compte de la place des adjectifs.
- 8 Cependant, la faible quantité de données orales disponibles jusqu'à une époque récente explique sans doute à elle seule l'inexistence d'étude sur la question. En effet, la place de l'adjectif épithète est un sujet qui nécessite, au minimum, des données de l'ordre du million de mots pour pouvoir initier une étude quantitative. A la lecture des fréquences mentionnées par Thuilier (2012 : 334-335) portant sur C-ORAL-ROM (comprenant environ 300.000 mots pour la partie française), on voit par exemple qu'il y a un nombre d'occurrences extrêmement réduit pour des adjectifs aussi courants que *grand* (100), *bon* (45), *haut* (13), *fort* (9), *court* (2), et aucune pour *chaud*, *utile*, *sensible*, *sage*, *prudent*, etc. De plus, les adjectifs sont moins fréquents à l'oral spontané que dans la presse écrite ou la littérature (cf. ci-dessous pour les comparaisons entre l'oral et l'écrit). Il faut donc des corpus de plus grande taille qu'à l'écrit pour pouvoir mener une étude sur l'oral, alors même que les données orales demandent un travail initial beaucoup plus conséquent que les données écrites (transcription précise, multiples phases de correction, etc.).
- 9 A l'heure actuelle, la situation évolue très vite et on dispose désormais de données orales en quantité et en qualité significatives. Grâce à ce contexte favorable, nous avons donc entrepris de lancer une description systématique de la place des adjectifs en français parlé (majoritairement non planifié), en commençant par nous limiter aux adjectifs apparaissant à la fois à la droite et à la gauche du substantif auquel ils se rapportent. Ces adjectifs seront qualifiés d'« adjectifs alternants » dans la suite de ce travail. C'est le cas dans des exemples tels *une sublime jeune femme* / *une jeune femme sublime* où il semble n'y avoir aucune différence entre les deux exemples (hormis la place de l'adjectif justement), mais aussi *le dernier métro* / *la semaine dernière*¹ où la différence de sens est patente. Il s'agira notamment d'observer dans quelle proportion les adjectifs peuvent se postposer ou s'antéposer dans des contextes similaires ou identiques, sans paramètre distinctif apparent (comme dans l'exemple avec *sublime* ci-dessus). Je désignerai ce phénomène par le terme d'alternances « neutres ». Les alternances « neutres » s'opposent aux alternances « motivées ». Voici un exemple d'alternance « motivée » dans lequel la différence de sens permet de rendre compte de la place de l'adjectif : *mes propres difficultés* / *une serviette propre*. Nous émettons l'hypothèse que le choix d'une place motivée serait plus systématique en français parlé non planifié par rapport aux écrits planifiés. La difficulté à appréhender le phénomène de l'alternance viendrait donc des données prises en compte dans les études antérieures, données tendant à masquer les régularités.
- 10 Pour mener à bien cette étude, nous avons adopté une démarche « conduite par les corpus » (« corpus driven », selon la terminologie de Tonigni-Bonelli (2001)), en nous limitant exclusivement aux attestations rencontrées. Ainsi, les adjectifs sont considérés

comme alternants si et seulement s'ils possèdent au moins une occurrence en antéposition et en postposition.

- 11 Notre article est organisé de la manière suivante. Dans un premier temps, nous présentons les corpus et la méthodologie suivie pour extraire les exemples. Dans un second temps, nous exposons les résultats globaux en les comparant à ceux obtenus dans des travaux antérieurs. Puis, nous proposons des regroupements par fonctionnement similaire, basés sur l'étude détaillée de neuf adjectifs. Pour finir, nous concluons sur l'intérêt de travailler à partir de transcriptions de français parlé non planifié, données permettant de mettre en évidence les constructions préférentielles nécessitant un temps d'élaboration court. Il est important de souligner qu'il s'agit d'une étude exploratoire pour laquelle nous ne mentionnons, dans ce qui suit, que les premiers résultats.
- 12 Afin de mettre en lumière l'influence, sur les résultats, de la nature des données prises en compte, une comparaison sera menée avec des données de taille identique provenant d'articles de presse. Il s'agira de se focaliser plus particulièrement sur les cas où les tendances observées ne sont pas congruentes entre l'oral et la presse afin de montrer que la mise en évidence de régularités peut être fortement affectée par les données prises en compte.

2. Corpus et concordances

- 13 La ressource orale utilisée fait environ 2.300.000 mots et provient d'un regroupement de corpus divers. Les formats et les conventions de transcription ont été uniformisés afin de pouvoir les interroger à l'aide d'un concordancier et les annoter automatiquement en parties du discours (cf. Benzitoun & Bérard, 2010). Sans un tel traitement, les requêtes automatiques auraient été plus aléatoires, étant donné le nombre difficilement prévisible de graphies différentes pour un même lexème (en fonction des conventions de transcription adoptées). De même, la qualité de l'annotation automatique en parties du discours aurait été grandement affectée, le nombre de formes différentes démultipliant le nombre de formes inconnues du logiciel d'annotation automatique et donc le taux d'erreurs.
- 14 La ressource est composée des corpus suivants : *Corpus de Français Parlé Parisien*² - CFPP (Branca-Rosoff et al., 2012), *Corpaix* (Blanche-Benveniste, 1999), *Corpus de Référence du Français Parlé - CRFP* (Equipe DELIC, 2004), *C-ORAL-ROM* (Cresti & Moneglia, 2005), *Phonologie du Français Contemporain*³ - PFC (Durand et al., 2002 ; Durand et al., 2005), *Choix de Textes en Français Parlé - CTFP* (Blanche-Benveniste et al., 2002). Les situations de parole représentées sont diversifiées, allant de discussions à bâtons rompus à des émissions de télévision en passant par des entretiens et des interviews. Nous disposons des métadonnées pour les deux tiers de la ressource environ. Cependant, il est possible de se faire une idée du contexte et du contenu de la transcription en recourant au texte intégral, lorsque nous n'avons pas accès aux métadonnées. Etant donné l'influence de la situation de parole sur les phénomènes observés (comme nous le verrons plus loin), il est important de pouvoir disposer de cette indication ou, à défaut, d'essayer de la déduire de la lecture des données. Notamment, il est fondamental de savoir si les personnes, dans un enregistrement particulier, parlent de manière spontanée ou, au contraire, ont plus ou moins planifié leur discours.

- 15 L'intégralité de la ressource a été annotée automatiquement en parties du discours et lemmatisée grâce à un outil que nous avons spécialement entraîné sur du français parlé. La correction manuelle d'environ 100.000 mots du corpus oral TCOF (André & Canut, 2010) nous a servi de référence pour effectuer l'apprentissage⁴ avec le logiciel *TreeTagger*. Nous avons évalué le taux de précision de l'annotation automatique à 93,6 % (toutes étiquettes confondues et pas seulement les adjectifs).
- 16 Pour ce qui est de la partie écrite destinée à effectuer les comparaisons, nous avons regroupé des corpus de presse, à savoir le corpus Chambers-Rostand (composé d'articles extraits des quotidiens *L'Humanité*, *Le Monde* et *La Dépêche du Midi*), une journée du corpus de l'*Est Républicain* et la partie presse du *Corpus Evolutif de Référence du Français (CERF)*. Cette tranche représentant la presse écrite possède une taille comparable à celle de la partie orale.
- 17 Les requêtes ayant permis d'extraire automatiquement les adjectifs en position d'épithète ainsi que les substantifs auxquels ils se rapportent ont été formulées à l'aide de *WinLoX*⁵, logiciel permettant de faire des recherches sur des fichiers étiquetés de type tabulaire. Toutefois, notre ressource n'étant pas annotée en syntaxe, nous avons été contraint de formuler un ensemble de requêtes comportant un substantif et un adjectif dans une fenêtre relativement étroite en essayant de décrire un maximum de contextes possibles. L'objectif était de limiter au maximum à la fois le bruit (i.e. les fausses détections) et le silence (i.e. les exemples pertinents ratés). Nous avons donc cherché tous les adjectifs suivis ou précédés d'un substantif dont il pouvait être séparé par un (ou deux) adverbe(s) et/ou une (ou deux) particules discursives. Nous avons donc bien évidemment ratés des exemples tels que le suivant, dans lequel l'adjectif *embêtants* est trop éloigné pour être détecté et où, de toute façon, c'est l'adjectif *petit* qui aurait été repéré car plus proche du substantif *endroits* :

1/ ça passe à des **endroits** euh quand même je veux dire un petit peu **embêtants**

- 18 Les concordances obtenues ont la forme suivante dans l'ordre des informations contenues dans les colonnes : les 6 mots précédant l'adjectif ; l'adjectif ; les 6 mots suivant l'adjectif ; le lemme de l'adjectif ; le lemme du substantif. Par cette méthode, nous avons obtenu 39.083 occurrences (19.532 en antéposition et 19.551 en postposition) avant correction. La répartition entre adjectifs antéposés et postposés de même que les calculs de fréquence ont été effectués automatiquement à partir des requêtes formulées.
- 19 Afin de nous assurer que nous disposions uniquement d'exemples correspondant à des adjectifs épithètes et à leurs substantifs têtes, nous avons vérifié chaque ligne de concordance en supprimant les détections erronées et en corrigeant les substantifs lorsqu'il ne s'agissait pas de la tête. Pour ce qui est du silence, nous avons essayé de quantifier le nombre de cas non détectés pour une trentaine d'adjectifs. Cela représente un taux de silence théorique d'environ 6 %. Nous avons profité de cette phase de vérification pour injecter dans les concordances les exemples non détectés initialement, ce qui a quelque peu diminué le silence. Dans le cadre de cette vérification, nous avons privilégié les adjectifs alternants.
- 20 Tous les repérages d'adjectifs sont basés uniquement sur la forme. Ainsi, par défaut, nous considérons que *propre* dans *mon propre domaine de spécialité* et *une serviette propre* est un adjectif alternant, même si le sens est très différent. Ce choix, provisoire, est lié à

l'hypothèse selon laquelle la forme a une incidence sur les possibilités de placement, la place étant un paramètre important de désambiguïsation. Dit autrement, il y aurait une tendance à ce qu'un adjectif donné soit en distribution complémentaire de part et d'autre du substantif auquel il se rapporte, sauf paramètre supplémentaire venant neutraliser cette répartition. Si on prend l'exemple de l'adjectif *ancien*, on s'aperçoit qu'il n'a pas le même sens dans *une ancienne église reconvertie en habitation* et *une église ancienne*. Or, il suffit d'ajouter le modifieur *très* pour que cela neutralise la différence de sens liée à la place : *une très ancienne église* vs *une église très ancienne*. Afin de tester cette hypothèse, nous avons mis ensemble les formes identiques, que celles-ci soient des homonymes, des unités polysémiques ou qu'elles aient exactement le même sens.

- 21 Toutefois, l'extraction automatique des exemples grâce à *WinLoX* a eu au moins deux conséquences problématiques : la perte des références permettant d'identifier chaque transcription et l'impossibilité de visualiser une occurrence particulière en accédant facilement au texte intégral. De plus, même si le silence ne paraissait pas extrêmement élevé, il était non nul et variable d'un adjectif à un autre. Pour toutes ces raisons, nous avons, dans un second temps, utilisé le logiciel *AntConc* directement sur les données brutes afin d'étudier les adjectifs alternants en détail. L'extraction automatique à partir de la version annotée en parties du discours nous a finalement servi uniquement à déterminer la liste des adjectifs alternants et à observer les grandes tendances. Sur ce point, nous pensons qu'un corpus de français parlé arboré représenterait une ressource précieuse permettant d'effectuer une extraction automatique plus fiable⁶. Mais il faut signaler que l'annotation automatique en parties du discours représente déjà un gain de temps important par rapport à un travail intégralement manuel. Sans un corpus annoté, nous aurions eu la plus grande difficulté ne serait-ce qu'à repérer les adjectifs alternants.
- 22 Dans la partie suivante, nous présentons les résultats généraux.

3. Résultats généraux

- 23 Après vérification, correction et insertion d'exemples supplémentaires non repérés automatiquement, nous parvenons au total de 33.397 occurrences se répartissant de la façon suivante : 16.037 en antéposition (soit 48 %) et 17.360 en postposition (soit 52 %). Comparés aux 27 % d'antéposition et 73 % de postposition mentionnés par Thuilier (2012), chiffres calculés sur le *French Treebank*, cela représente une différence notable. Cependant, ces chiffres masquent des disparités importantes. Pour avoir une idée des répartitions générales, il est important d'explicitier ce qui se cache derrière ces fréquences en présentant des données plus détaillées et en distinguant les adjectifs alternants des adjectifs bloqués en anté- et postposition :

Tableau 1. Répartition générale des adjectifs à l'oral

	Antéposé	Postposé	Alternant	Total
Lemmes	41 1,9 %	1943 90,6 %	162 7,5 %	2146 100 %
Occurrences	624	13643	19130	33397

	1,9 %	40,8 %	57,3 %	100 %
Ratio occ./lemmes	15,2	7	118,1	15,6

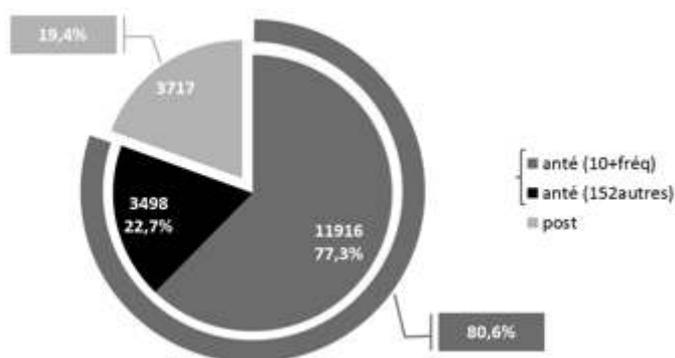
- 24 Comme on peut le voir dans le tableau ci-dessus, les adjectifs alternants sont peu nombreux du point de vue des lemmes différents, mais ils sont majoritaires du point de vue du nombre d'occurrences. Ce décalage est dû à la présence dans cette catégorie des adjectifs les plus fréquents tels *grand*, *petit*, *autre*, etc.
- 25 Par comparaison, Wilmet (1980) obtient la répartition suivante à partir d'un corpus composé d'œuvres littéraires contemporaines dont la taille théorique est de 1.200.000 mots⁷ : 29.016 occurrences en tout se répartissant en 3.835 lemmes différents dont 645 adjectifs alternants différents. Bien que le corpus de Wilmet (1980) soit près de deux fois moins volumineux que les données orales que nous avons rassemblées, le nombre d'adjectifs alternants est près de quatre fois supérieur. La mobilité des adjectifs est donc beaucoup plus importante dans les œuvres littéraires qu'en français parlé.
- 26 Au niveau du seul nombre total d'occurrences, il y a également une différence importante entre la fréquence mentionnée par Thuilier (2012), celle de Wilmet (1980) et celle que nous avons constatée :

Tableau 2. Comparaison de nos résultats avec ceux obtenus dans d'autres études

	<i>French Treebank</i> Thuilier (2012) ⁸	<i>Littérature</i> Wilmet (1980)	<i>Corpus oraux</i>
Taille totale	385.458	1.200.000	2.300.000
Occurrences	13.933	29.016	33.397
Lemmes	1.750	3.835	2.146
Adjectifs alternants	171	645	162

- 27 Comme on peut le voir, pour une taille près de six fois inférieure (dans le cas du *French Treebank*), la fréquence des adjectifs épithètes est seulement inférieure de l'ordre de 2,5 fois. L'usage des adjectifs épithètes est donc beaucoup plus fréquent dans la presse écrite et la littérature qu'à l'oral. Si l'on part du principe théorique selon lequel la fréquence des adjectifs est globalement proportionnelle à la taille des données prises en compte, à taille égale il y aurait environ deux fois plus d'occurrences dans la presse écrite qu'en français parlé. Et on peut faire le même constat en comparant l'oral à la littérature.
- 28 En ce qui concerne spécifiquement l'oral, si on se focalise sur les adjectifs alternants, on s'aperçoit que les cas d'antéposition sont largement majoritaires et que cette majorité concerne pour les trois quarts les 10 adjectifs les plus fréquents dans cette place⁹.

Figure 1. Répartition des adjectifs alternants en fonction de la place



- 29 Comme le montre le graphique précédent, les dix adjectifs les plus fréquents rendent compte à eux seuls d'une partie importante de la fréquence constatée pour les antépositions. Pour aborder la question de la place de l'adjectif, on voit là toute l'importance de ne pas se limiter à la présentation de chiffres généraux. Il est indispensable d'observer s'il y a des adjectifs qui sont surreprésentés et qui ainsi ont une influence cruciale sur les fréquences relevées.
- 30 En ce qui concerne l'étude des paramètres permettant d'expliquer l'alternance, nous avons choisi de partir de l'étude des adjectifs pris individuellement. Plutôt que de faire, à l'image de Thuilier (2012), une étude générale sur l'influence de chaque paramètre (forme du déterminant, présence d'un modifieur, coordination avec un autre adjectif, etc.), nous avons préféré nous focaliser sur l'observation des concordances pour chaque adjectif. En effet, nous nous sommes rapidement aperçu que d'un adjectif à un autre l'influence d'un paramètre particulier n'avait pas du tout le même poids.
- 31 Nous avons pour l'instant observé avec précision neuf adjectifs. Nous faisons état des principales tendances dans la suite du présent article.

4. Principales tendances

4.1. Alternances non spontanées

- 32 Sur les neuf adjectifs que nous avons étudiés, cinq n'alternent pas de manière « spontanée », bien qu'ils soient considérés comme alternants dans la littérature sur la question. Il s'agit d'*actuel*, *difficile*, *dur*, *important*, *puissant*, adjectifs que l'on peut considérer comme uniquement postposés à l'oral non planifié. Dans le cas de *puissant* et *difficile*, nous n'avons rigoureusement aucune occurrence d'antéposition dans les données orales observées. Or, pour *puissant* l'antéposition est majoritaire dans la presse écrite et la place réalisée dans les différents exemples n'est généralement pas explicable par l'une des contraintes, préférentielles ou catégoriques, identifiées dans les études antérieures. Pour preuve, l'exemple suivant, dans lequel les deux contextes sont strictement identiques :

2/ Dernière égalité à 16/16 avec une frappe **puissante** de Capet, à laquelle répond illico Sapinart, tandis que Ces enfonce le clou sur une balle sauvée in extremis. [Dépêche du midi]

3/ *Moukassa adressait une **puissante** frappe que Preau avait toutes les peines du monde à repousser (2')*. [Est Républicain]

- 33 Le cas de *puissant* illustre parfaitement l'existence de ce que l'on pourrait appeler la grammaire du français non planifié (ce que Blanche-Benveniste (1990) appelle « grammaire première ») opposée à la grammaire du français planifié (« grammaire seconde » chez Blanche-Benveniste (1990)). Ce second type de grammaire peut prendre le contre-pied de la grammaire première, ce qui semble être le cas ici.
- 34 L'absence d'occurrence de *puissant* en antéposition pourrait toutefois être un effet de bord des données prises en compte. En effet, un corpus étant clos par nature, il est possible qu'en augmentant son volume ou les situations de parole, on puisse trouver des occurrences de *puissant* antéposées. Mais cela n'irait pas à l'encontre de notre hypothèse sur la grammaire du français non planifié. Il ne s'agit ni d'une règle stricte, ni d'une question de grammaticalité, mais plutôt d'une place privilégiée se reflétant dans les fréquences, comme nous allons le voir dans les exemples ci-dessous.
- 35 En effet, il peut arriver que l'on trouve quelques exemples résiduels d'antéposition. Ceux-ci s'expliquent notamment par la situation de parole :

4/ *enfin l'**importante** réduction des participations communales pour les zones d'aménagement confiées à notre SEM liées à l'achèvement de plusieurs opérations et à leurs bons résultats vient elle aussi renforcer nos moyens* [CRFP, PUB-PNO-1, conseil municipal]

5/ *Frodo jouit pleinement de son **actuelle** suprématie en l'absence de Freud* [C-ORAL-ROM, fmedrp02, documentaire télévisé sur les chimpanzés]

- 36 Il apparaît clairement, dans les exemples (4-5), qu'il ne s'agit pas de discours spontané mais de paroles lues et/ou largement préparées (un conseil municipal dans le premier cas et un documentaire télévisé dans le second). En guise d'indice linguistique congruent avec la situation de parole, l'adjectif *important* se trouve, en (4), dans un syntagme nominal à fonction de sujet que la taille oriente clairement vers de la parole lue. Un sujet d'une telle longueur et d'une telle complexité se rencontre très rarement, voire jamais, dans un discours non planifié.
- 37 D'autres antépositions peuvent s'expliquer par un effet du cotexte. Dans l'exemple ci-dessous, nous faisons l'hypothèse que l'occurrence de *long travail* dans la proximité immédiate de *dur* a une influence sur l'antéposition de ce dernier :

6/ *c'est un **long** travail et un **dur** travail* [CRFP, PRI-TOU-1]

- 38 Par un effet de mimétisme, la répétition lexicale de *travail* favoriserait la reproduction à l'identique de l'ordre des mots réalisé dans le premier syntagme nominal.
- 39 Par ailleurs, dans la presse écrite, nous rencontrons à nouveau des exemples où il n'y a aucun indice linguistique permettant d'expliquer l'alternance, les contextes d'emploi étant presque totalement identiques :

7/ *des élus et représentants de partis politiques dont ceux de l'**actuel** gouvernement* [L'Humanité]

8/ *Chaque jour un peu mieux, les Français comprennent la brutalité du gouvernement actuel.* [L'Humanité]

- 40 Ces exemples clairs d'alternance neutre sont sans doute de ceux qui brouillent l'observation des régularités. Pour ces adjectifs, la postposition est sans aucun doute la place utilisée spontanément, le recours à l'antéposition pouvant être considéré comme un indice de planification du discours.
- 41 Ci-dessous, les fréquences des cinq adjectifs étudiés, à l'oral et dans la presse écrite :

Tableau 3. Comparaison des fréquences pour les cinq adjectifs étudiés

ADJECTIFS	ORAL		PRESSE ECRITE	
	Antéposé	Postposé	Antéposé	Postposé
<i>Puissant</i>	0	17	61	44
<i>Difficile</i>	0	98	15	154
<i>Actuel</i>	2	132	109	353
<i>Dur</i>	2	72	12	90
<i>Important</i>	9	276	149	338

4.2. Alternances en distribution complémentaire

- 42 Parallèlement à ces adjectifs sans alternance spontanée, il en existe qui sont en distribution complémentaire. Comme nous l'avons montré dans Benzitoun et al. (2010), l'adjectif *prochain* a une distribution régulière à l'oral non planifié. Hormis le substantif temporel *fois*, avec lequel *prochain* s'antépose très majoritairement, on retrouve seulement des substantifs à sémantisme temporel au singulier lorsque l'adjectif est postposé¹⁰. Dans ce contexte, il doit être accompagné d'un déterminant défini ou ne comporter aucun déterminant. Ainsi, on trouve des exemples tels *l'année prochaine*, *la semaine prochaine*, *lundi prochain* mais pas *?l'événement prochain* ou *?une année prochaine*. De même, nous n'avons trouvé aucun exemple où *prochain* est postposé et au pluriel (du type *?les années prochaines*), quelle que soit la classe sémantique du substantif.
- 43 Or, ce n'est pas ce que nous avons constaté dans des écrits planifiés. Dans un corpus écrit diversifié de taille identique, nous avons trouvé les exemples suivants :

9/ *le voyageur se hâte vers la station **prochaine*** [CERF, Roman]

10/ *qui doivent s'être arrêtés dans **les mois prochains*** [CERF, Chirac]

11/ *Il annonce **une** normalisation **prochaine** de la situation* [CERF, Le Monde]

- 44 De même en ce qui concerne les données obtenues par introspection, Berthonneau (2002) invente notamment l'exemple ci-dessous, qu'elle considère comme étant acceptable :

12/ Le maire **prochain** devra mettre un terme à ces pratiques.

- 45 A la régularité observée pour les données orales non ou peu planifiées se substitue donc un placement plus libre pour les productions écrites (ayant majoritairement un degré de planification élevé). Insistons sur le fait que si l'on tenait compte de manière indifférenciée de l'ensemble des exemples – inventés, écrits, oraux – aucune régularité claire ne se dégagerait en ce qui concerne la grammaire de *prochain*. On voit bien là les limites d'une approche agrégeant l'ensemble des exemples, sans distinguer minimalement des sous-types de données.
- 46 Pour résumer, voici les résultats tirés de Benzitoun et al. (2010) :

Tableau 4. Comparaison du fonctionnement de *prochain* à l'oral et à l'écrit

ORAL (164 occ.)		ECRIT (290 occ.)	
Singulier (149 occ.)		Singulier (217 occ.)	
Antéposé (51 occ.)	Postposé (98 occ.)	Antéposé (95 occ.)	Postposé (122 occ.)
Noms temporels (51 %) (<i>la prochaine fois</i> (25 occ./26))	Noms temporels (100 %) <i>année/an</i> (48 occ.) <i>semaine</i> (23 occ.)	Noms temporels (20 %)	Noms temporels (86,1 %)
Autres noms (49 %)		Autres noms (80 %)	Sens de « proche » (9 %)
			Divers (4,9 %)
Pluriel (15 occ.)		Pluriel (73 occ.)	
Antéposé (15 occ.)	Postposé (0 occ.)	Antéposé (64 occ.)	Postposé (9 occ.)
		Noms temporels (62,5 %)	
		Autres noms (37,5 %)	

4.3. Alternances en quasi-distribution complémentaire

- 47 A l'image de *prochain*, il existe d'autres adjectifs dont les différents emplois sont en distribution complémentaire. Mais, à la différence de *prochain*, quelques rares exemples s'écartent de ce fonctionnement. C'est le cas de l'adjectif *propre*. En antéposition, le sens est systématiquement « appartenance spécifique » (ex : *mon propre problème* ; désormais

sens 1) et jamais « bien lavé » (ex : *une chemise propre* ; désormais sens 2). En antéposition toujours, *propre* est accompagné soit d'un déterminant possessif soit d'un syntagme prépositionnel exprimant la possession :

13/ à chaque fois qu'il parle d'un étage il utilise **sa propre** façon de de de compter les étages [C-ORAL-ROM, fnatte02]

14/ le **propre** frère **de mon ami** est un sportif euh qui fait à peu près deux cent cinquante à trois cents kilomètres de vélo le dimanche [CRFP, PRI-AMI-2]

- 48 Sur les 150 occurrences antéposées, il y a seulement deux exemples ne respectant pas cette règle :

15/ l'individu est vraiment quelqu'un d'unique qui a **une propre** histoire [C-ORAL-ROM, ffammn19]

16/ ils avaient trouvé chacun **une propre** règle pour eux [CORPAIX, 34SOPH]

- 49 Nous touchons là à une difficulté posée par l'approche « corpus driven ». L'occurrence d'un exemple n'oblige pas à le considérer comme appartenant au système linguistique d'une langue donnée. Il est important de discuter de son statut. On peut en effet avoir affaire à des exemples comportant une erreur ou faisant partie d'un idiolecte. C'est l'hypothèse que nous formulons ici. Dans (15) et (16), nous considérons que la présence du déterminant indéfini couplée à l'antéposition de *propre* est un indice de mauvaise formation et que cela ne poserait aucun problème si ces exemples comportaient un déterminant possessif.
- 50 En postposition, la majorité des exemples que l'on trouve avec le sens 1 sont ceux où il y a un modifieur post-adjectival (qui bloque l'antéposition et permet une liberté plus grande concernant le déterminant) ou un pronom possessif en guise de constructeur :

17/ il y a une culture **propre à l'endroit** [PFC, 38aep1]

18/ je pense que à partir du moment où l'homme a éliminé l'autre en tant que valeur euh je dirai(s) égale à égale à la à la **sienne propre** [CTFP, les_gris-gris]

- 51 Les autres exemples de *propre* postposé sont ceux où il s'agit du sens 2 :

19/ on leur avait appris dans le cours qu'il fallait enlever les bijoux qu'il fallait enlever le vernis à ongles et qu'il fallait donner euh une **chemise propre** [CRFP, PRO-LYO-1]

- 52 ou de figements ou quasi-figements : *nom propre, fonds propres, amour propre, en mains propres*.
- 53 Les exemples qui vont à l'encontre de ce fonctionnement général en distribution complémentaire entre le sens 1 et le sens 2 sont soit liés à la situation de parole et à un discours planifié (ce qui était déjà le cas pour les alternances non spontanées) :

20/ l'essentiel c'est que chacun euh se se réalise d'après ses qualités **propres**
[CORPAIX, APOSTROP, émission télévisée]

54 soit à un effet du cotexte :

21/ chacune des communautés faisait passer son **intérêt propre** avant l'**intérêt national** [C-ORAL-ROM, ffammn17]

55 Dans (21), nous faisons l'hypothèse que l'occurrence de l'ordre substantif-adjectif dans *intérêt national* a une incidence sur la postposition de *propre* par un phénomène de calque (voir aussi l'exemple (6) ci-dessus). En plus de ces explications, on peut également ajouter que les sujets grammaticaux des exemples (20) et (21) sont très proches (*chacun* et *chacune des communautés*) et que ces sujets comportant une dimension sémantique distributive pourrait également avoir une influence sur la postposition.

56 Abstraction faite des cas ci-dessus, pour lesquels nous avons pu donner une explication concernant la place, les deux seuls exemples où *propre* est postposé (alors qu'il pourrait également être antéposé selon nous) sont les suivants :

22/ le patois était le euh leur ben leur langage **propre** [CORPAIX, UTOPIE]

23/ L2 comme ça de de but en blanc il ne me vient pas de euh d'autres d'autres plats à l'esprit si ce n'est ce que nous fabriquons nous-mêmes

L1 oui oui bien sûr

L2 à savoir nos spécialités **propres** [CRFP, PRO-SAI-1]

57 Cependant, il est possible de faire l'hypothèse d'une nuance sémantique entre *nos spécialités propres* et *nos propres spécialités*. Dans ce cas, l'alternance serait motivée sémantiquement et on pourrait donc inclure ces exemples dans le système d'ensemble décrit plus haut à condition d'ajouter un sens supplémentaire. Nous ne trancherons pas cette question en l'absence d'argument décisif.

58 En plus de ces exemples où *propre* pourrait s'antéposer (cf. 22-23), nous en avons repéré deux autres dans lesquels *propre* est théoriquement ambigu avec le sens 2 mais où il est vraisemblablement bloqué en postposition :

24/ ils ont gardé le terme ils /l', n'/ ont pas créé dans la langue **propre** un terme qui correspond à ce concept [CRFP, PUB-PSE-1]

25/ je je vais pas parler de patois parce que c'en est pas mais on a euh un langage **propre**
[CORPAIX, CYRIL]

59 En effet, *la propre langue* paraît peu acceptable et nous n'en avons trouvé aucun exemple. Nous avons pour l'instant du mal à intégrer ces exemples au système général de l'adjectif *propre* que nous avons décrit ci-dessus. Comme pour (22) et (23), nous les mettons provisoirement de côté dans l'attente d'une analyse plus approfondie. Mais même en intégrant ces exemples, on observe tout de même une tendance générale à

éviter les contextes ambigus entre les deux sens clairement identifiés de *propre* et à favoriser une place plutôt qu'une autre dans des contextes précis.

- 60 Dans la presse écrite, en revanche, il y a dix occurrences où *propre* est postposé et permettraient une alternance libre¹¹, comme l'attestent les exemples ci-dessous :

26/ *Les éditions ont donc aussi leur histoire propre.* [Le Monde]

27/ *N'est-ce pas non plus une façon pour les artistes qui inventent des façons neuves de se mouvoir, de court-circuiter les formes « classiques » du hip-hop, d'inventer, de détourner, de raconter leur propre histoire ?* [L'Humanité]

- 61 Et il y a dix autres exemples bloqués en postposition et théoriquement ambigu avec le sens 2, comme (24) et (25).
- 62 En conclusion, ce qui transparaît principalement dans l'étude de *propre*, c'est la tendance, à l'oral non planifié, à éviter les ambiguïtés entre les deux sens. En antéposition, seul le sens 1 est attesté. En postposition, les deux sens sont représentés mais les contextes d'emploi sont différents, ce qui fait qu'il n'y a presque aucun cas de recouvrement. Seuls quatre occurrences à l'oral non planifié s'écartent de ce principe, alors qu'elles sont beaucoup plus nombreuses dans la presse écrite, comme le montre le tableau récapitulatif ci-dessous :

Tableau 5. Comparaison du fonctionnement de *propre* à l'oral et dans la presse écrite

	ORAL		PRESSE ECRITE	
	Antéposé	Postposé	Antéposé	Postposé
SENS 1	150	25 dont 4 cas à part ¹²	334	37 dont 20 cas à part
SENS 2	0	12	0	22
FIGEMENT	0	11	0	27
TOTAL	150	48	332	86

- 63 Nous allons maintenant voir le dernier cas de figure, à savoir l'alternance libre.

4.4. Alternances libres

- 64 Bien que la majorité des adjectifs semblent posséder une place dépendant de paramètres précis, nous avons tout de même rencontré des adjectifs pour lesquels l'alternance semble beaucoup moins contrainte. Il s'agit d'*énorme* et *immense*. Cette liberté de placement se reflète clairement dans les fréquences observées à l'oral. Une fois de plus, la répartition est fort différente dans la presse écrite, comme on peut le constater dans le tableau suivant :

Tableau 6. Place de énorme et immense à l'oral et dans la presse écrite

ADJECTIFS	ORAL		PRESSE ECRITE	
	Antéposé	Postposé	Antéposé	Postposé
<i>Énorme</i>	57	66	102	37
<i>Immense</i>	15	15	124	14

- 65 Les exemples relevés montrent une indifférence de la place, à l'image de :

28/ *d'un village à l'autre il y a pas une énorme différence* [CORPAIX, 26SAVOIE]

29/ *il y a une différence énorme* [C-ORAL-ROM, ffamd125]

30/ *nous avons l'immense plaisir d'accueillir sur ce plateau Céline Dion* [CTFP, le_journal_tele]

31/ *c'est un plaisir immense* [CORPAIX, ANASTAS]

- 66 Pour ces deux adjectifs, aucune régularité ne semble se dégager, excepté la présence d'un modifieur qui bloque l'antéposition (du moins dans les que nous avons observées) :

32/ *ça c'est toute ma base on a un groupe assez énorme* [CRFP, PRI-PNO-1]

33/ ? *ça c'est toute ma base on a un assez énorme groupe*

4. Conclusion

- 67 A partir de l'analyse approfondie de neuf adjectifs alternants à l'oral, nous avons montré que la place d'un adjectif donné est généralement motivée. Les régularités constatées à l'oral non planifié ne se retrouvent pas nécessairement dans la presse écrite et plus généralement dans les écrits planifiés (cf. l'étude sur *prochain*). Si les contraintes préférentielles ne sont pas absentes des données orales (cf. Thuilier, dans le présent volume), les contraintes catégoriques (distribution complémentaire) semblent peser beaucoup plus fortement sur le français parlé non planifié. En ce qui concerne les adjectifs épithètes, il existerait donc ce que l'on pourrait appeler une distribution « privilégiée » respectant des paramètres extrêmement précis (forme du déterminant, sens, etc.) et se traduisant par une place « privilégiée ». Cette place « privilégiée » n'a pas été mise en évidence jusqu'à présent car sa visibilité est parasitée par des paramètres stylistiques qui pèsent plus fortement dans les écrits planifiés.
- 68 Toutefois, il existe quelques adjectifs pour lesquels la place est totalement libre et pour lesquels aucun paramètre ne permet de rendre compte d'une place ou d'une autre pour la majorité des occurrences. Il reste encore à dresser la liste des adjectifs à place libre en français parlé non planifié, ce que nous nous attacherons à faire dans une publication ultérieure.

- 69 En outre, les trois sources de données à la disposition du linguiste donnent à voir des fonctionnements différents. Les données orales non planifiées mettent en évidence un fonctionnement globalement régulier avec très peu d'adjectifs à alternance libre. Les données écrites planifiées, à cause du poids de la stylistique et du besoin de varier les tournures, brouillent ces régularités en étendant le champ des possibles mais en conservant toutefois des tendances fortes reflétées dans les fréquences d'emploi. Les données obtenues par introspection, quant à elles, vont avoir tendance à ne pas conserver une partie des exemples pourtant rencontrés dans la poésie ou la littérature. On comprend donc pourquoi certains linguistes en sont arrivés à la conclusion que l'alternance était une propriété de la quasi-totalité des adjectifs du français et pourquoi le mystère du fonctionnement des adjectifs n'a pas encore été percé. Cela met en évidence la question centrale de la grammaticalité, qui, au moins dans le cas de l'étude de l'adjectif, doit être maniée avec d'infinies précautions. Notre étude rend également visible les différents possibilités linguistiques reflétées dans les données observées.
- 70 A partir de ces constats, on peut émettre de stimulantes hypothèses concernant ce que l'on pourrait appeler les tournures grammaticales « privilégiées » (grammaire première) dans le français non planifié, domaine dans lequel il faudra approfondir les recherches. A l'écrit planifié, il y aurait une tendance à s'écarter de ces régularités « privilégiées » (grammaire seconde). Il reste encore à circonscrire de manière précise ces deux types de grammaire, à définir ce qu'elles nous apprennent sur la langue et à se doter d'outils théoriques et méthodologiques pour les observer. Notamment, la grammaire première se reflète clairement dans les fréquences de l'oral mais la grammaire seconde n'en est pas absente pour autant. Il y a donc une influence de la grammaire seconde, y compris sur l'oral non planifié. Et il est parfois difficile de discriminer avec précision les données façonnées par la grammaire première et celles façonnées par la grammaire seconde.
- 71 Toujours est-il que les études sur le français parlé non planifié ont un intérêt particulier en sciences du langage et que ces données devraient être privilégiées dans les travaux en psycholinguistique portant sur le fonctionnement du langage oral. De même, l'enseignement du français langue étrangère dans une perspective d'apprentissage de l'oral aurait tout à gagner à une systématisation des études sur le français parlé non planifié et aux règles mises en évidence dans ce cadre. Cela montre également la pertinence des études diachroniques portant sur les aspects socio-historiques des langues et l'utilité de travailler à partir de sources écrites (seuls témoignages restants) aussi proches que possibles de la langue spontanée. La description de l'évolution de la langue française s'en trouvera sans doute profondément modifiée.

BIBLIOGRAPHIE

Abeillé, A., L. Clément & F. Toussanel (2003) Building a Treebank for French, in A. Abeillé (ed.), *Treebanks: Building and Using Parsed Corpora*, Dordrecht: Kluwer, p. 165-187.

- André, V. & E. Canut (2010) Mise à disposition de corpus oraux interactifs : le projet TCOF (traitement de corpus oraux en français), *Pratiques*, 147/148, p. 35-51.
- Benzitoun, C., K. Fort & B. Sagot (2012) TCOF-POS : un corpus libre de français parlé annoté en morphosyntaxe, Actes de la conférence conjointe JEP-TALN-RECITAL, volume 2 : TALN, Grenoble, 4-8 juin, p. 99-112.
- Benzitoun, C. & L. Bérard (2010) Mutualisation et uniformisation de ressources de français parlé, in Azzopardi S. (coord), *Corpus, Données, Modèles, Cahiers de Praxématique*, 54-55, PULM : Montpellier, p. 175-188, article paru en 2013.
- Blanche-Benveniste, C. (1990) Grammaire première et grammaire seconde : l'exemple de *en*, *Recherches sur le français parlé*, 10, p. 51-73.
- Blanche-Benveniste, C. (1999) Constitution et utilisation d'un grand corpus, *Grands corpus : diversité des objectifs, variété des approches*, *Revue Française de Linguistique Appliquée*, 4/1, p. 65-74.
- Blanche-Benveniste, C., C. Rouget & F. Sabio (2002) *Choix de textes de Français parlé, 36 extraits*, Paris, Champion, Collection Les français parlés.
- Blanche-Benveniste, C. (2010) *Le français - Usages de la langue parlée*, Louvain-Paris : Peeters.
- Blasco-Dulbecco, M. & P. Cappeau (2004) Quelques remarques sur l'adjectif à l'oral, in J. François (dir.), *L'adjectif en français et à travers les langues*. Caen : Presses Universitaires de Caen, Bibliothèque de Syntaxe & sémantique, p. 413-428.
- Blasco-Dulbecco, M. & P. Cappeau (2005) Ce que les corpus oraux nous apprennent sur les adjectifs, in G. Williams (dir.), *La linguistique de corpus*, Rennes : Presses Universitaires de Rennes, p. 69-80.
- Branca-Rosoff, S., S. Fleury, F. Lefevre & M. Pires (2012) *Discours sur la ville. Présentation du Corpus de Français Parlé Parisien des années 2000 (CFPP2000)* - <http://cfpp2000.univ-paris3.fr/CFPP2000.pdf>
- Branca-Rosoff, S., S. Fleury, F. Lefevre & M. Pires, adresse du site internet du CFPP : <http://cfpp2000.univ-paris3.fr/>
- Cappeau, P. (2002) Entre l'auxiliaire et le participe passé, *Recherches sur le français parlé*, 17, p. 11-28.
- Cresti, E. & M. Moneglia (2005) *Integrated Reference Corpora for Spoken Romance Languages*, *Studies in corpus linguistics*, 15, John Benjamins.
- Delomier, D. (1980), La place de l'adjectif en français : bilan des points de vue et théories du XX^e siècle, *Cahiers de lexicologie*, 37/2, Jacques & Demontrond, p. 5-24.
- Durand, J., B. Laks & C. Lyche (2002) La phonologie du français contemporain : usages, variétés et structure, C. Pusch & W. Raible (eds.), *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache/Romance Corpus Linguistics - Corpora and Spoken Language*, Tübingen, Gunter Narr Verlag, p. 93-106.
- Durand, J., B. Laks & C. Lyche (2005) Un corpus numérisé pour la phonologie du français, G. Williams (éd.), *La linguistique de corpus*, Rennes, Presses Universitaires de Rennes, p. 205-217.
- Equipe DELIC (2004) Présentation du 'Corpus de référence du français parlé', *Recherches sur le français parlé*, 18, p. 11-42.
- Forsgren, M. (1978) *La place de l'adjectif épithète en français contemporain Etude quantitative et sémantique*, Acta Universitatis Upsaliensis, Studia Romanica Upsaliensis, 20, Uppsala.

Forsgren, M. (1997) Un classique revisité : la place de l'adjectif épithète, in *Etudes de linguistique française, médiévale et générale offertes à Robert Martin à l'occasion de ses 60 ans*, G. Kleiber & M. Riegel (éd.), Champs linguistiques, Duculot, Louvain-la-Neuve, p. 115-126.

Fox, G. (2012) *L'acquisition des modifieurs nominaux : le cas de l'adjectif du français*, Thèse de doctorat, Université Paris 3.

Glatigny, M. (1967) La place des adjectifs épithètes dans deux œuvres de Nerval, *Le français moderne*, 35, p. 201-220.

Goes, J. (1999) *L'adjectif. Entre nom et verbe*, Duculot, Paris – Bruxelles.

Larsson, B. (1994) *La place et le sens des adjectifs épithètes de valorisation positive*, *Etudes romanes de Lund*, 50, Lund, University Press.

Reiner, E. (1968) *La place de l'adjectif épithète en français, théories traditionnelles et essais de solution*, Vienne, Stuttgart, W. Braumüller.

Thuilier, J. (2012) *Contraintes préférentielles et ordre des mots en français*, Thèse de doctorat, Université Paris-Diderot.

Tognini-Bonelli, E. (2001) *Corpus Linguistics at work*, John Benjamins Publishing Company, Amsterdam and Philadelphia.

Wagner R.-L. & J. Pinchon (1962) *Grammaire du français classique et moderne*, Hachette.

Wilmet, M. (1980) Antéposition et postposition de l'épithète qualificative en français contemporain, *Travaux de linguistique*, 7, p. 179-201.

Wilmet, M. (1981) La place de l'épithète qualificative en français contemporain. Etude grammaticale et stylistique, *Revue de linguistique romane*, 45, p. 17-73.

NOTES

1. Pour l'instant, nous nous attachons uniquement à la forme, sans distinguer les adjectifs qui n'ont pas le même sens. Nous justifions ce choix un peu plus loin.
2. Les données sont celles présentes sur le site en juin 2010.
3. Il s'agit uniquement des entretiens libres présents sur le site de PFC en mars 2010.
4. Le fichier d'apprentissage servant de paramètre au logiciel *TreeTagger* est disponible à l'adresse suivante : <http://www.cnrtl.fr/corpus/perceo/>. Le protocole ayant permis cet entraînement est détaillé dans Benzitoun et al. (2012).
5. <http://lipn.univ-paris13.fr/~audibert/lox.php>
6. C'est ce que se propose de faire le projet ORFEO, dirigé par Jeanne-Marie Debaisieux et financé par l'Agence Nationale de la Recherche.
7. Wilmet (1980) ne précise pas la taille totale des données qu'il a utilisées. En revanche, il signale que son corpus est composé des 50 premières pages de 80 ouvrages, ce qui fait un total de 4.000 pages. Notre calcul estimatif est basé sur le chiffre moyen de 300 mots par page.
8. A noter que Thuilier (2012) a ôté de ce décompte toutes les occurrences d'adjectifs comportant un modifieur postposé (ex : *un défaut propre à son frère*) étant donné que cela oblige l'adjectif à se postposer au substantif auquel il se rapporte.
9. Ces dix adjectifs sont (par ordre de fréquence décroissante) : *petit, autre, grand, même, premier, bon, gros, certain, dernier, beau*.
10. Pour une étude plus détaillée, le lecteur pourra se reporter à Benzitoun et al. (2010).
11. A moins de postuler une différence de sens, comme dit plus haut.

12. Nous utilisons la dénomination vague de « cas à part » pour désigner les exemples que nous ne sommes pas parvenu à intégrer à la description. Ce terme a l'avantage de ne pas être stigmatisant même si nous sommes conscient qu'il manque de précision.

RÉSUMÉS

La place de l'adjectif épithète en français parlé n'a pas encore fait l'objet d'une étude systématique. Nous avons donc entrepris de combler ce manque en nous concentrant, dans un premier temps, sur les adjectifs alternants. Nous avons plus particulièrement cherché à savoir si la place était majoritairement motivée ou au contraire relativement libre. Pour ce faire, 9 adjectifs ont été observés de manière précise et nous avons essayé de dresser une première cartographie des différents fonctionnements. De nos premiers résultats, nous en tirons la conclusion que la motivation de la place est majoritaire mais qu'il existe cependant quelques adjectifs à place totalement libre en français parlé. Nous apportons des éléments supplémentaires permettant d'étayer les hypothèses de C. Blanche-Benveniste sur la dualité entre grammaire première et grammaire seconde, par l'intermédiaire d'une comparaison avec des données issues de la presse écrite.

The positioning of the attributive adjective is a subject that has produced a substantial literature in French linguistics. In most articles and books, it is considered that the adjective may as well be before or after its head noun. However, previous studies have focused almost exclusively on written French and / or invented examples. There are few studies on spoken French and no broad description. We therefore choose to launch a systematic description of the position of adjectives in spoken French (mostly unplanned), starting by limiting to adjectives appearing on both right and left of the noun to which they relate. These adjectives are described as alternating. This is the case in examples like:

*une **sublime** jeune femme / une jeune femme **sublime*** ('a gorgeous young woman').

We focus on adjectives that can be pre-nominal or post-nominal in similar or identical contexts, without apparent distinctive setting (as in the example above).

To carry out this study, we adopted a corpus-driven approach (Tonigni-Bonelli, 2001), restricted exclusively to occurrences met in our data. This paper presents a first exploratory study in which we mention only the first results. To highlight the influence on the results of the nature of the data taken into account, a comparison will be conducted with data of the same size from newspaper. It will be focusing specifically on the cases for which the trends are not congruent between oral and newspaper. So, we show that the detection of regularities can be strongly affected by the data taken into account.

Our study is based on an oral resource of approximately 2.3 million words coming from a combination of various corpora. File formats and transcription conventions were standardized. The entire resource was automatically tagged into parts of speech and lemmatized with a tool specially trained to work on spoken French. Regarding the written data for comparisons, we also combined various newspaper corpora. The size of the newspaper part is comparable to the size of the oral part. Queries were launched on the tagged corpus to observe the main trends and to extract the list of alternating adjectives. Then we read concordances to identify and correct detection errors. However by this method, we realized that we missed a small part of relevant examples. In a second step, to have the guarantee to work on the entirety of relevant data and to

control a maximum of parameters, we queried directly the raw transcriptions to lead detailed analysis.

On the tagged spoken corpus, we identified a total of 33,397 occurrences being divided as follows: 16,037 pre-nominal adjectives (48%) and 17,360 post-nominal adjectives (52%). Alternating adjectives are not frequent in terms of different forms (only 162 of 2,146), but they are the majority in terms of occurrences (19,130). This phenomenon is related to the presence, in this category, of the most common adjectives like *grand* ('large'), *petit* ('little'), *autre* ('other'), etc. Furthermore, we saw that the pre-nominal position, among alternating adjectives, are large majority and that majority is mostly related to the 10 most frequent adjectives in this position.

Regarding the various adjective usages, we highlighted four cases. First, we observed non spontaneous alternation. They are adjectives that alternate in writing but very rarely in the oral data that we viewed. The only examples are encountered in recordings of planned spoken French:

*Frodo jouit pleinement de son **actuelle** suprématie en l'absence de Freud* [c-oral-rom] ('Frodo fully enjoys his current supremacy in the absence of Freud' [tv report on chimpanzees])

or in situations where there is an effect of parallelism:

*c'est un **long** travail et un **dur** travail* [CRFP] ('it is a long work and a hard work').

The second case includes adjectives which the position is in complementary distributions. In this category, we met only discriminating contexts. This is the case of the adjective *prochain* ('next' in English). All contexts in which *prochain* is in pre-nominal position are different from those where it is in post-nominal position.

The third case is similar to the previous but it contains some rare examples that are not in complementary distribution. The adjective *propre* ('own' in English) belongs to this category. We have found two post-nominal examples that are identical to the pre-nominal context (presence of a possessive determiner):

*le patois était le euh leur ben leur langage **propre*** [Corpaix] ('the dialect was their own language').

Finally, there are free alternate adjectives, that is to say adjectives can be either pre-nominal or post-nominal. This last category seems relatively marginal compared to others. But only a thorough work on the 162 alternate adjectives will demonstrate this hypothesis.

We draw the tentative conclusion that the position of adjectives in unplanned spoken French is mostly intended. This point highlights the existence of a spontaneous grammar, hard to see in written data especially for stylistic reasons.

INDEX

Mots-clés : Adjectifs épithètes, adjectifs alternants, ordre des mots, grammaire spontanée, français parlé

Keywords : Attributive adjectives, alternating adjectives, words order, spontaneous grammar, spoken French

AUTEUR

CHRISTOPHE BENZITOUN

Université de Lorraine & ATILF CNRS
Christophe.Benzitoun@univ-lorraine.fr