



Linx

Revue des linguistes de l'université Paris X Nanterre

57 | 2007

Études de syntaxe : français parlé, français hors de France, créoles

Que peuvent nous apprendre en syntaxe des corpus oraux « anciens » ?

Paul Cappeau



Édition électronique

URL : <http://journals.openedition.org/linx/79>

DOI : 10.4000/linx.79

ISSN : 2118-9692

Éditeur

Presses universitaires de Paris Nanterre

Édition imprimée

Date de publication : 1 décembre 2007

Pagination : 15-25

ISSN : 0246-8743

Référence électronique

Paul Cappeau, « Que peuvent nous apprendre en syntaxe des corpus oraux « anciens » ? », *Linx* [En ligne], 57 | 2007, mis en ligne le 14 février 2011, consulté le 15 mai 2020. URL : <http://journals.openedition.org/linx/79> ; DOI : <https://doi.org/10.4000/linx.79>

Que peuvent nous apprendre en syntaxe des corpus oraux « anciens » ?

Paul Cappeau

Université de Poitiers – FoReLL et MoDyCo

La question du changement syntaxique lorsqu'elle prend appui sur des données orales est généralement traitée en terme de variation (Gadet : 1999). Le lien étroit entre les deux notions explique leur forte imbrication¹. Toutefois, parce qu'il s'agit de langue parlée, le relevé de formes porte fréquemment sur des phénomènes écartés de l'écrit. Dès lors, il devient presque impossible d'éviter que l'observation de la variation ne s'oriente vers un recensement de tournures condamnées par la norme :

La langue parlée a développé des mécanismes grammaticaux qui ne sont pas reçus comme des tournures correctes et qui sont, de ce fait, évitées par une grande partie des personnes qui écrivent. (Blanche-Benveniste : 2003)

Le relevé de quelques faits sensibles à une modification (la liaison, la négation, l'interrogation, les relatifs) dans des ouvrages qui traitent de l'histoire de la langue française (Antoine & Martin : 1995, Chaurand : 1999, Yaguello : 2003) confirme cette orientation. Pour le présent travail, l'angle adopté est légèrement différent pour deux raisons :

¹ Ce que la formulation de Marchello-Nizia (2006) synthétise bien : « toute variante est un changement possible ». Valli (1998) attire l'attention sur le caractère complexe de cette relation selon le domaine considéré : « Dans le domaine morphosyntaxique, les relations entre phénomènes de variation et de changement peuvent se montrer infiniment plus délicates à interpréter. »

- l'analyse est centrée sur la recherche de faits de langue présents ou absents dans des corpus oraux produits à des dates légèrement éloignées dans le temps². Par conséquent, la mise en relation avec la norme de l'écrit ne sera pas activée et les données écrites ne tiendront qu'une place marginale (lorsque cela paraîtra nécessaire) ;

- l'accent est mis sur la démarche (le « comment faire » et les difficultés rencontrées) et les faits signalés sont plutôt à voir comme des illustrations. Il ne faut donc pas attendre un relevé exhaustif des changements mais plutôt quelques indications qui tentent de répondre partiellement à la question posée dans le titre de cet article.

1. Les difficultés de l'entreprise

Quatre rubriques permettront de présenter plus avant le corpus utilisé et de défendre la démarche suivie dans ce travail.

1.1. Recul chronologique réduit

Le corpus *Phonothèque* qui est le support de ce travail est constitué d'interviews conduites par Francine Bloch. Il s'agissait pour elle d'interroger des artistes (en majorité liés au domaine musical comme G. Bacquier, N. Boulanger, G. Tailleferre, etc.) sur leur parcours. Ces entretiens étaient accessibles au public de la *Phonothèque française*³. Les enregistrements utilisés ont été collectés entre 1958 et 1979, soit un recul de 30 à 50 ans, ce qui soulève d'emblée la question du caractère « ancien ».

Deux réponses peuvent être apportées à cette première critique. Pour les corpus oraux, ce terme « ancien » ne peut avoir la même amplitude que pour les documents écrits. Travailler sur des données vieilles de 50 ans présente donc un intérêt certain⁴. De plus, un autre regard peut être apporté sur la question de la datation : plutôt que de s'en tenir à la seule date d'enregistrement, on peut s'intéresser à l'âge virtuel des interviewés (leur date de naissance ou l'âge qu'ils auraient s'ils avaient vécu jusqu'à aujourd'hui). Le corpus *Phonothèque* dévoile sa richesse puisque les âges s'étalent de 83 ans à 125 ans⁵.

² Wagner (1953) : « Quand on compare une langue avec elle-même, sous les formes qu'elle présente en deux points éloignés du temps, ses formes ne se recouvrent pas ».

³ Je tiens à exprimer tous mes remerciements à Pascal Cordereix (conservateur à la BnF) qui m'a grandement facilité l'accès à ces documents.

⁴ Dans les *Archives de la parole*, Brunot nous donne accès à des documents sonores certes plus anciens (début du 20^{ème} siècle) mais il s'agit de parole récitée.

⁵ Une réflexion plus approfondie devrait être engagée sur ce paramètre et pourrait valoir aussi pour l'écrit. Pour Bernard Combettes (c.p.), il pourrait être utile de ne pas se limiter à la date de parution d'une œuvre et de retenir aussi l'âge de l'auteur lorsque le texte est produit. Dans le domaine de l'oral, le projet ESLO (*Enquête Sociolinguistique à Orléans*), parmi d'autres, permettra de comparer les productions d'un même locuteur à 40 ans de distance.

1.2. La taille des données

Le corpus transcrit comprend douze interviews, ce qui représente un total de 66 000 mots et correspond à 5h15 de son. Il s'agit de toute évidence d'un petit corpus. Mais, là encore, divers arguments conduisent à porter un regard plus nuancé sur la limitation des données.

Inutile de rappeler que les corpus oraux sont souvent d'une taille réduite (qui tient pour une bonne part au lourd travail de transcription et de vérification qu'ils nécessitent). La taille indiquée doit donc être évaluée à cette aune : dans les projets de corpus qui intègrent des données écrites et orales, l'oral ne représente au mieux que 30 % de l'ensemble. De plus, le corpus *Phonothèque* n'est qu'une partie (certes essentielle) des données exploitées. D'autres sources orales ont été utilisées : des corpus oraux constitués par quelques membres de l'ancienne équipe du GARS⁶ (*Groupe Aixois de Recherches en Syntaxe*) et le CRFP – un *Corpus de Référence du Français Parlé* (cf. RSFP 18). Ce dernier corpus présente une diversité intéressante (notamment géographique et en trois situations de parole : privée, professionnelle et publique). L'une de ses limites tient, encore une fois, à sa taille réduite (440 000 mots)⁷. Pour tenter d'affiner certaines hypothèses, il a été utile d'étendre la recherche à des corpus écrits, comme cela sera précisé par la suite. Bref, les données orales exploitables n'ont pas été confinées à celles du « petit » corpus *Phonothèque* mais portent sur un ensemble de 2 millions de mots.

1.3. Domaine observé a priori peu sensible

Si dans le cadre de la variation, le thème de l'hétérogénéité grammaticale est fréquemment traité (Gadet : 1999), il semble moins facile de vouloir identifier des changements syntaxiques dans les limites (de recul chronologique et de taille des données) précédemment évoquées. De fait, les modifications dans la syntaxe semblent demander du temps, un temps plus important que celui qu'offrent les données consultées. Blanche-Benveniste (1995) considère qu'une période de 30 ans ne constitue pas une durée suffisamment étendue :

Il est difficile de dire s'il y a eu, entre 1914 et 1945, de réels changements dans la syntaxe du français. C'est un laps de temps un peu court pour y voir des évolutions qui, en ce domaine, sont habituellement lentes.

Gadet (1999) formule le même constat quand elle aborde la langue du 20^{ème} siècle, même si elle évoque une piste (concernant des variations dans la fréquence) qui sera explorée par la suite (cf. 2.2.) :

La morphologie constitue une dimension peu mobile, où s'expriment des tendances de longue durée. Il en va de même pour la syntaxe, avec peut-être des modifications de fréquence, possibles signes avant-coureurs de changements.

⁶ En particulier M. Blasco-Dulbecco (*Clermont Ferrand*), M. Savelli (*Grenoble*) et moi-même.

⁷ Divers corpus sont accessibles sur le net (cf. bibliographie) comme PFC (*Phonologie du Français contemporain*), mais ils sont peu utiles dans le cadre de ce travail.

Toutefois certains travaux antérieurs (comme Yaguello : 1998) ont pu mettre en lumière des évolutions intéressantes (ainsi l'emploi de *genre* qui devient dans la langue contemporaine un outil grammatical dans un exemple tel que *ça fait genre trois cents kilomètres*). Il semble donc que, même sur une période courte, certains faits puissent être repérés. D'autre part, la problématique adoptée atténue cette réserve : il ne s'agit pas de rechercher des nouveautés mais plus largement des modifications ou des disparitions dans la syntaxe.

1.4. Questions sur l'homogénéité du corpus

La question sur le statut du corpus demanderait à être bien plus développée notamment parce qu'elle est d'actualité à une époque où se multiplient les constitutions de gros corpus et où la différence entre *corpus* et *archive* (bien identifiée en théorie cf. Rastier : 2005 ou McEnery et alii : 2006) tend à être négligée ou abandonnée dans la réalité des pratiques (Cappeau & Gadet : 2007). Si de nombreux paramètres conduisent à considérer que les enregistrements rassemblés ont bien droit à l'appellation de corpus (une seule intervieweuse, des personnalités souvent proches du milieu musical, des buts probablement voisins lors du recueil, des entretiens généralement faits à domicile... et surtout une identification suffisamment précise des productions collectées), il n'en reste pas moins que l'homogénéité que le titre *Phonothèque* laisse attendre n'est que très partielle. Ainsi les interviews peuvent se présenter sous forme de dialogues fréquemment entrecoupés ou laisser au locuteur la possibilité de prises de parole longues (proches de narrations). Les deux extraits suivants donnent une idée des différences formelles entre les interviews rassemblées :

Extrait A : Bacquier

- L1 et à huit heures je me jetais à l'eau + huit heures et demie parce que c'était huit heures et demie à cette époque-là
L2 oui
L1 je me jetais à l'eau + et je me jetais à l'eau avec la vaillance de la jeunesse vous comprenez
L2 oui
L1 /et puis, 0/ qui n'a rien à perdre et tout à gagner
L2 {rire}
L1 vous comprenez

Extrait B : Sauguet

- L1 une barbe blanche un binocle des petits gants des gants gris hum hum qui n'étaient pas en peau mais en filoselle je crois euh euh un costume gris foncé un col à coins cassés il parlait très doucement euh ses yeux flamboyaient tout de même derrière ses lunettes un petit air malicieux mais fort courtois fort aimable se gardant bien de dire quoi que ce soit qui fût euh pe- qui pût être pris par nous pour une sorte de de de de de puissance

L'homogénéité semble plus reposer sur des critères externes que sur des paramètres internes (qui demanderaient d'ailleurs à être identifiés). Cette difficulté à préciser, parmi une multitude de paramètres, ceux qui sont pertinents, fait partie des problèmes auxquels le chercheur sur corpus est confronté (Gadet : 2006).

On aurait pu imaginer que comparer des productions du même « genre » (soit des interviews d'artistes à 50 ans d'intervalle) serait la solution la plus rigoureuse. Mais ce serait négliger le fait que la situation d'interview s'est modifiée entre ces deux périodes⁸ et que la langue utilisée à l'époque est certainement plus formelle que celle pratiquée de nos jours dans des circonstances considérées comme équivalentes.

2. Quelques pistes

Trois aspects différents liés à la syntaxe ont été observés : les tournures (ou constructions), la distribution de certains éléments, l'organisation syntaxique de passages longs.

2.1. Les tournures

Plusieurs limites rendent l'analyse délicate : avec un corpus de taille modeste, on ne peut s'attendre à rencontrer qu'un faible nombre d'attestations d'une même tournure, il faut alors se demander si la présence ou l'absence d'une tournure est réellement significative ; mais surtout se pose la question du partage entre lexique et syntaxe, lorsque l'on veut faire des recherches dans les données contemporaines.

a) Le *partitif*

L'exemple (1) illustre bien les questions soulevées ci-dessus :

(1) *j'ai rencontré de vos élèves de vos amis* (Phonothèque)

C'est le seul cas de cette construction rencontré dans le corpus *Phonothèque*. Mais la répétition observée du complément (*de vos élèves de vos amis*) incite à considérer que la structure n'est pas due à une maladresse du locuteur. Cette construction évoque la description que Bonnard (1978) propose pour le *partitif*. Il indique que deux formes ont été en concurrence : *mangier pain* et *mangier de pain*. Une concurrence analogue semble pouvoir être posée entre *rencontrer vos amis* et *rencontrer de vos amis*.

Avant de se prononcer sur un changement éventuel dans la construction du verbe *rencontrer*, deux remarques doivent être formulées :

- Cette construction n'est pas recensée dans des ouvrages descriptifs des constructions verbales comme Caput & Caput (1969) ou Dubois & Dubois-Charlier (1997).

- Parmi les quelque 430 occurrences du verbe *rencontrer* que comporte notre corpus oral, aucun exemple de cette construction n'est attesté. Cette situation illustre l'un des problèmes que pose le recours aux corpus : quand la syntaxe repose sur des items lexicaux précis, on sait que seuls des corpus de très grande taille (plusieurs

⁸ Sur cette question du genre, voir notamment Adam & Heidmann (2006).

dizaines de millions de mots) permettent d'espérer rencontrer des attestations. Le résultat obtenu sur les bases arrêtées initialement (recherche dans des corpus oraux à cinquante ans d'intervalle) est à la fois peu probant et frustrant : il ne permet pas de savoir si la construction partitive avec *rencontrer* est un hapax ou si elle témoigne d'un changement. Une issue envisageable est de se tourner vers le web malgré les réserves que cette archive aux limites mal cernées peut susciter à juste titre (Kilgarriff & Grefenstette : 2003). On a toutefois choisi de tester la construction partitive avec les deux verbes cités (*manger* et *rencontrer*). Le type de requête autorisée avec Google nous a conduit à comparer quatre séquences lexicales très précises :

Construction directe		Construction « partitive »	
<i>manger ces gâteaux</i>	28 cas	<i>manger de ces gâteaux</i>	4 cas
<i>rencontrer vos amis</i>	13 400 cas	<i>rencontrer de vos amis</i>	0 cas

Tableau 1. Occurrences de séquences lexicales attestées sur le web

Ces résultats ne fournissent pas de certitude (telle n'était pas l'attente d'ailleurs) mais soulèvent des questions intéressantes :

- La construction partitive semble ne plus exister avec le verbe *rencontrer* (on pourrait dire qu'elle persiste pour les gâteaux avec le verbe *manger* – encore que le nombre d'occurrences soit étonnamment faible – mais qu'elle n'a pas survécu avec du lexique humain).

- Cette construction a-t-elle été vraiment utilisée ou s'agit-il d'un idiolecte ? Une recherche étendue à l'écrit littéraire (par le biais de *Frantext* ou d'autres sources) a permis de trouver quelques cas, certes peu nombreux (2)-(3), mais qui montrent que cet emploi est attesté :

(2) *Ainsi vous ne risquerez pas de rencontrer de vos camarades.* (Montherlant)

(3) *Il avait la réputation justifiée d'être un très habile tireur, et, comme, à défaut de revenants, on pouvait rencontrer de ces rôdeurs qui battent les frontières, ou, à défaut de rôdeurs, quelque ours mal intentionné, il n'était que prudent d'être en mesure de se défendre.* (J. Verne. Le château des Carpathes)

Même si la démonstration a dû prendre certains détours imprévus, l'exploitation du corpus *Phonothèque*, telle qu'elle avait été envisagée, s'avère utile pour observer un type de changement bien réel dans la construction partitive d'un verbe précis. Moins que le résultat, modeste, c'est la piste qui est intéressante : elle prouve que sur une période de temps réduite, certains changements peuvent être repérés et que les corpus disponibles apportent des ressources indispensables à la démonstration.

b) Le participe présent

Si le point précédent a porté sur un type de construction (et a finalement été réduit à un lexème verbal), un phénomène syntaxique plus étendu concerne la construction des participes présents. Le corpus *Phonothèque* présente un emploi surprenant pour un lecteur de corpus oraux contemporains (4) :

- (4) *il doit certainement aimer la musique car ma mère **aimant** la musique terriblement elle était très très bonne musicienne mais /on devrait le voir, XXX/ + et elle m'a fait apprendre la musique* (Phonothèque – Lévy)

Pour s'en tenir aux exemples rencontrés dans le CRFP, si l'on met à part des emplois figés (concernant quelques verbes particuliers tels que *s'agissant...*) ou l'emploi de type adjectival (que l'on aurait dans *c'est frappant*), on rencontre alors des exemples tels que (5)-(6) :

- (5) *il s'agit là du 3ème attentat **s'abattant** sur la capitale* (CRFP)
(6) *en tant que fille **aimant** s'amuser draguant eh elle essaie de voir au fait qui il est* (CRFP)

La différence essentielle entre les emplois « anciens » et contemporains porte sur le lien de dépendance entre le nom et le participe présent. Il est lâche en (4) puisque l'on a une construction détachée qui permet d'antéposer le participe, ce que les emplois liés ne permettent pas (4')-(6') :

- (4') *car aimant la musique terriblement ma mère (elle) était très très bonne musicienne*
(5') ** s'abattant sur la capitale il s'agit là du 3ème attentat*
(6') ** en tant qu'aimant s'amuser fille elle essaie de voir qui il est*

Mais la portée de ces observations par rapport à la question du changement syntaxique reste ténue. En effet, la construction détachée du participe présent a été et reste observée à l'écrit (7) :

- (7) *Nous allions, ne songeant à rien.* (Cité par Grévisse et Goose, § 923)

Il s'agit donc, au mieux, d'un emploi qui se raréfie à l'oral ou, comme signalé précédemment, qui montre que la situation d'interview suscitait le recours à une langue plus formelle que ce que l'on observe dans les productions contemporaines.

2.2. La distribution de certaines unités semi-lexicales

L'entrée par les lexèmes est envisageable pour peu que ceux-ci aient une diffusion suffisante. Les mots lexicaux se prêtent mal à la démonstration (à cause de leur faible fréquence) et les mots grammaticaux, plus fréquents, sont *a priori* moins sensibles au changement. Reste alors à se tourner vers les mots situés à mi-chemin. Des formes comme *chose* ou *gens* entrent bien dans ce cadre. C'est cette dernière forme qui a été retenue à travers ses emplois dans 3 sous-corpus de taille comparable : le corpus *Phonothèque*, la partie *parole professionnelle* et la partie *parole publique* du CRFP. La variation du nombre d'occurrences ne singularise pas le corpus *Phonothèque* :

Corpus	<i>Phonothèque</i>	<i>Parole professionnelle</i>	<i>Parole publique</i>
Nombre total de mots	66 000	68 000	66 000
Nombre d'occurrences de <i>gens</i>	51	56	189

Tableau 2. Occurrences de *gens*

Mais ce résultat brut n'est pas réellement pertinent pour dégager des changements portant sur la syntaxe de ce terme. D'où la nécessité de mieux cerner les faits de distribution. Si l'on tient compte du déterminant qui précède *gens* ou de la fonction qu'occupe cet élément, le corpus *Phonothèque* retrouve une forte singularité et se distingue des tendances observées dans les deux corpus contemporains :

Corpus	<i>Phonothèque</i>	<i>Parole prof.</i>	<i>Parole publique</i>
Proportion de <i>les gens</i>	46 %	68 %	61 %
Emploi sujet	6 %	18 %	14 %

Tableau 3. Distribution de *gens*

Au niveau du type de complément, on peut observer une autre modification dans la syntaxe. Dans le corpus *Phonothèque*, *gens* est plus souvent traité comme un nom plein et comme tel suivi de compléments (*gens de maison*) ou d'adjectifs (*gens très importants*, *gens absolument extraordinaires*, *gens spéciaux*...). Ces emplois même s'ils ne sont pas écartés des corpus oraux contemporains y sont moins visibles car il semble que ce terme soit en voie de grammaticalisation (Cappeau & Deulofeu : 2006) et se comporte plus comme une sorte de pronom (humain et pluriel). La différence dans l'emploi en position sujet (position nettement plus fréquente dans les données contemporaines) va dans le même sens. On serait donc dans un cas de changement de fréquence d'emploi (plus que de réelle nouveauté) qui, comme le suggérait Gadet (1999), est peut-être annonciateur d'une modification plus importante à venir.

2.3. L'organisation de passages longs

Ce dernier point permet de s'intéresser aux phénomènes de planification de la production. Les locuteurs produisent des listes étendues qui excèdent (en nombre d'éléments) ce que l'on observe dans les productions actuelles. Voici un exemple qui comporte une liste de 4 séquences en *le plus* + Adjectif (8) :

- (8) *il ne jouait pas du tout à l'homme majestueux au maître pas du tout il était vraiment l'homme le plus courtois l'homme le plus aimable l'homme le plus charmant l'homme le plus discret et avec tout cela je vous dis cette espèce de mystère qui l'entourait et qui est devenu de plus en plus grand plus en plus à mesure que je le connaissais* (Phonothèque)

Dans les productions contemporaines, les exemples du même modèle semblent se limiter à 3 éléments (9) :

- (9) *et vous avez + dans l'Espagne du siècle d'or + hein l'auteur à mes yeux le plus original le plus + anti-féministe le plus misogynne de toute littérature qui est Quevedo* (CRFP)

Toujours dans l'organisation des listes, l'exemple (10) présente une énumération de 6 éléments (facilement identifiables par l'amorce *que ce soit...*) dont on n'a pas trouvé d'équivalent dans les productions orales contemporaines :

- (10) *je crois qu'un compositeur un compositeur d'aujourd'hui surtout doit être disponible à tous les doit pouvoir répondre présent à tous les appels il doit avoir un métier assez large assez souple assez ren- assez renouvelé pour pouvoir se plier à toutes les disciplines que ce soit la musique d'orchestre que ce soit la musique de théâtre que ce soit la musique de film que ce soit la musique radiophonique que ce soit la musique concrète que ce soit la musique religieuse nous devons pouvoir répondre nous avons nous devons pouvoir répondre présent à tous les appels* (Phonothèque)

Ce type de passage développé possède une organisation interne qui semble pouvoir être reliée à la formation reçue (Chervel : 1998). Il est probable que des modèles rhétoriques sous-jacents influencent la maîtrise de ces régularités. L'une des différences avec les réalisations contemporaines porte sur le souci de ne pas reproduire la même séquence au-delà de trois ou quatre occurrences (en changeant par exemple l'élément introducteur). Ainsi, dans le passage (11) la liste en *comment* comporte 4 occurrences même si elle se déploie dans une liste plus large :

- (11) *finalement on redécouvre Balzac on redécouvre l'inventeur de ce genre romanesque qui veut tout saisir à la fois l'histoire la philosophie la société les destins individuels comment ce Balzac est entré à partir d'un moment donné en concurrence alors sur le plan de la vente de ses propres livres avec des auteurs plus populaires comme Eugène Sue et Dumas comment il en a souffert comment à côté de cette œuvre publiée sous différents noms il y a toute une correspondance avec madame Hanska et comment aussi la fin de de Balzac est tragique* (France Culture – 20-11-05)

Liste :

*on redécouvre ...
comment ce Balzac est entré...
comment il en a souffert
comment à côté ...
et comment aussi la fin de de Balzac est tragique*

3. Conclusion

Ce bref aperçu visait à esquisser des pistes de description que les corpus oraux anciens offrent. Les résultats ne correspondent pas entièrement aux prévisions et confirment, ce qui n'est pas une surprise, que les données modifient (ou même doivent modifier) les perspectives tracées *a priori*. Ainsi, l'idée de s'en tenir aux seules productions orales n'a-t-elle pas toujours été tenue : cela aurait conduit à se contenter d'indices trop fragmentaires et peu probants alors que l'extension à des corpus plus amples (donc écrits) apporte, quelquefois, des preuves plus tangibles. Sur certains faits néanmoins, l'hypothèse qu'un corpus oral « ancien » présente un grand intérêt pour aborder les changements en syntaxe s'est avérée et plaide en faveur d'une exploration plus développée du patrimoine (comme les données disponibles à l'INA) et un retour vers

le passé, dans le cadre de projets de corpus oraux, est une direction qu'il ne faudrait pas négliger. Enfin, même si cette piste n'a été qu'effleurée, la difficulté à isoler les paramètres pertinents dans la description d'un corpus demanderait aussi à être mieux dégrossie. Ainsi, la notion d'*interview* ne semble pas suffisamment stable pour pouvoir constituer un critère fiable dès lors que l'on change d'époque.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ADAM, J.-M. & HEIDMANN, U. (2006). « Six propositions pour l'étude de la généricité ». *La Licorne*. 79 : 21-34.
- BLANCHE-BENVENISTE, C. (1995). « Quelques faits de syntaxe » in Antoine, G. & Martin, R. (éds). *Histoire de la langue française 1914-1945*. Paris. CNRS-Editions. 125-147.
- BLANCHE-BENVENISTE, C. (2003). « La langue parlée », in Yaguello, M. (éd.). *Le grand livre de la langue française*. Paris. Seuil. 317-344.
- BIBER, D., JOHANSSON, S., LEECH, G., CONRAD, S. & FINEGAN, E. (1999). *Longman Grammar of Spoken and Written English*. London and New York. Longman.
- BONNARD, H. (1978). « Article » in *Grand Larousse de la langue française*. Paris. Larousse. 258-260.
- CAPPEAU, P. & DEULOFEU, J. (2006). « Les “indéfinis” en relation avec la position sujet dans trois types de constructions prédicatives », in Corblin, F., Ferrando, S. & Kupferman, L. (dir.). *Indéfini et prédication*. Paris. PUPS. 125-138.
- CAPPEAU, P. & GADET, F. (2007). « L'exploitation sociolinguistique des grands corpus. Maître mot et pierre philosophale ». *Revue Française de Linguistique Appliquée*. XII-1 : 99-110.
- CAPUT, J. & CAPUT, J.-P. (1969). *Dictionnaire des verbes français*. Paris. Larousse.
- CHAURAND, J. (1999). *Nouvelle histoire de la langue française*. Paris. Le Seuil.
- CHERVEL, A. (1998). *La culture scolaire. Une approche historique*. Paris. Belin.
- DUBOIS, J. & Dubois-Charlier, F. (1997). *Les verbes français*. Paris. Larousse.
- GADET, F. (1999). « La langue française au XX^e siècle – I. L'émergence de l'oral », in Chaurand, J. (éd.). *Nouvelle histoire de la langue française*. Paris. Seuil. 583-674.
- GADET, F. (2006). « Hier comme aujourd'hui. Quelques faits de variation en syntaxe » in Guillot, C. & alii (éds). *A la quête du sens. Etudes littéraires, historiques et linguistiques en hommage à Christiane Marchello-Nizia*. Paris. ENS Editions. 191-197.
- KILGARRIFF, A. & GREFENSTETTE, G. (2003). « Introduction to the Special Issue on the Web as Corpus ». *Computational Linguistics*. 29-3 : 333-348.
- MARCHELLO-NIZIA, C. (2006). *Grammaticalisation et changement linguistique*. Bruxelles. De Boeck.
- MCENERY, T., XIA, R. & TONO, Y. (2006). *Corpus-Based Language Studies. An advanced resource book*. New York. Routledge Applied Linguistics.
- RASTIER, F. (2005). « Enjeux épistémologiques de la linguistique de corpus » in Williams, G. (éd.). *La linguistique de corpus*. Rennes. PUR. 31-45.
- Recherches Sur le Français Parlé* n° 18. (2004). « Autour du corpus de référence du français parlé ».

Que peuvent nous apprendre en syntaxe des corpus oraux « anciens » ?

- VALLI, A. (1998). « Réflexion sur la notion de variation en syntaxe », in Bilger, M., van den Eynde, K. & Gadet, F. (Eds). *Analyse linguistique et approches de l'oral. Recueil d'études offert en hommage à Claire Blanche-Benveniste*. Louvain / Paris. Peeters. 371-380.
- WAGNER, R. L. (1953). *Grammaire et philologie*. Cours de la Sorbonne. Paris. CDU.
- YAGUELLO, M. (1998). *Petits faits de langue*. Paris. Le Seuil.
- YAGUELLO, M. (éd.). (2003). *Le grand livre de la langue française*. Paris. Seuil.

ACCÈS À DES CORPUS ORAUX

Site PFC : <http://www.projet-pfc.net/>

Site CLAPI : <http://clapi.univ-lyon2.fr/>

Site corpus d'Orléans (ESLO1) : <http://bach.arts.kuleuven.be/elicop>

Site Corpus de la parole : <http://www.corpusdelaparole.culture.fr/>

Site CFPP2000 : <http://ed268.univ-paris3.fr/syled/ressources/Corpus-Parole-Paris-PIII/Corpus.html>

Site CRDO (Paris) : <http://crdo.risc.cnrs.fr>

Site CRDO (Aix) : <http://crdo.up.univ-aix.fr/>

