

LES SUJETS POSTVERBAUX EN FRANÇAIS PARLÉ : ENTRE FIGEMENT ET LIBERTÉ

Paul CAPPEAU
Université de Poitiers, FoReLL

RÉSUMÉ

Les données orales apportent un éclairage foncièrement différent à l'analyse des sujets inversés. Elles permettent notamment d'identifier les inversions figées, bloquées qui sont généralement de type formulaire (me semble-t-il, toujours est-il, ne serait-ce, ...) et de les distinguer des inversions réellement productives (c'est ce que disait ma mère). L'analyse prend appui sur un corpus de 450 000 mots environ (le CRFP) répartis selon 3 types de situations (oral privé, public et professionnel).

ABSTRACT

Data on oral speech production provide a radically different view of the phenomenon of subject inversion. Among other things, they enable the distinction between on the one hand frozen and fixed inversions, which usually belong to formulaic language (me semble-t-il, toujours est-il, ne serait-ce, ...) and on the other, genuinely productive inversions (c'est ce que disait ma mère). The present analysis is based on the CRFP, a 450,000-word corpus which comprises data on three types of situations (private, public and professional talks).

1. *IT AIN'T EASY*¹ : LE SUJET INVERSÉ AU FILTRE DES DONNÉES

L'intérêt pour le sujet inversé en français suscite, depuis bien longtemps, les commentaires des analystes de la langue française. Diverses grammaires de l'Ancien français au XVIII^e siècle (Buridant 2000, Gougenheim 1974, Haase 1914/1975, Seguin 1972) comportent d'ailleurs un chapitre sur l'ordre des mots (souvent le dernier) dans lequel sont formulées des indications plus ou moins développées sur la place du sujet. De nombreux travaux attestent aussi de l'intérêt que grammairiens et linguistes portent à ce fait de syntaxe

¹ Schnedecker (2005) a montré que David Bowie avait sa place dans un titre d'article de linguistique. J'ai tenté de faire de même avec des titres de ses chansons.

(pour plus de références on renvoie à Cappeau & Lahousse (2015)). Cette prolifération de regards présente néanmoins une forte homogénéité en ce qui concerne les données utilisées : les exemples sont essentiellement tirés de la langue littéraire.

1.1. *Fashion* : l'inversion telle que la présentent les linguistes

Sans surprise, on retrouve ce tropisme dans les travaux anciens comme Le Bidois (1952), mais aussi dans des ouvrages plus récents. Ainsi :

– Fuchs (1997) cite des exemples qui proviennent de *Frantext* et des exemples inventés ;

– Gardes Tamine (2003) puise dans la littérature classique (Hugo, Flaubert) pour illustrer ce qu'elle appelle les conditions textuelles. Elle complète son analyse à l'aide d'exemples construits ;

– Lahousse (2011) recourt à des exemples tirés de *Frantext*, à d'autres repris à des travaux antérieurs ainsi qu'à des énoncés forgés.

Aussi notre connaissance du sujet postposé n'est, pour l'essentiel, que celle des sujets postposés présents dans la littérature. On peut alors se demander si l'on ne pourrait pas étendre à bien d'autres auteurs la critique que formulait Chevalier (1985) : « F. Brunot ne décrit qu'une langue, la langue écrite des gens cultivés ».

Cet « oubli » de tout un pan de la langue (l'écrit non littéraire et la langue parlée) est-il négligeable, sans incidence sur la description et les analyses proposées ? En effet, ne peut manquer de se poser une question sur le caractère généralisable de descriptions échafaudées à partir de données aussi singulières et aussi homogènes : reflètent-elles ce qui se passe dans la grammaire de la langue française (dans sa totalité) ou ne constituent-elles que les possibilités que permet la langue littéraire mais que l'usage courant délaisse ?

1.2. *What's Really Happening* : l'éclairage des données orales

On peut signaler au moins trois déformations de l'objet d'étude, c'est-à-dire trois points de connaissance qui ne peuvent être conservés lorsque l'on passe de la langue littéraire à la langue parlée.

a) L'importance du phénomène. Tant qu'on s'intéresse à la langue écrite littéraire (ou de presse), la formule de Roig (ici même) : « L'inversion du sujet est une pratique relativement courante en français » est un fait aisément vérifiable. À l'oral, il convient de nuancer cette affirmation. Sur l'ensemble des 132 petits corpus qui composent le CRFP (*Corpus de Référence du Français Parlé*²), le nombre de sujets postposés (moins de 300 exemples relevés) est tout de même assez faible au regard de la quantité des verbes

² Pour une présentation détaillée de ce corpus voir DELIC (2004) ou Cappeau (2007).

utilisés³ et du nombre de mots (450 000 environ). Un inventaire plus précis fait ressortir la façon dont se répartissent les sujets postposés. On peut identifier deux grands ensembles :

– Dans 53 % des sous-corpus, on trouve soit 0 exemple, soit un seul cas.

- On ne trouve pas de sujet postposé dans 35 corpus (27 %).
- On en trouve une seule occurrence dans 34 corpus (26 %).

– Dans 47 % des cas, on trouve plus d'un sujet postverbal. Mais ce nombre connaît une forte variation (de 2 à 19 occurrences par sous-corpus). Globalement le sujet postposé reste un phénomène assez marginal, faiblement présent dans la langue parlée.

b) Le caractère naturel des exemples utilisés. La langue littéraire conduit à raisonner sur des énoncés qui semblent, quelquefois, à la marge des usages ordinaires. Certains exemples littéraires avec sujet postposé peuvent-ils être considérés comme offrant une image réaliste de la langue ? On pense à l'exemple suivant : « *Au plaisir ajoute une sorte d'avant-goût héroïque l'incommodité du plaisir.* » (Gide, cité dans Jonare 1976 : 84). Quelle place accorder à un tel exemple dans une description ? Faut-il y voir un arrangement productif dans la syntaxe du français ou une tentative pour frôler les limites de ce qu'autorise la grammaire ? De fait, les analystes délaissent parfois les réserves ou les caractérisations que même des grammairiens traditionnels formulent : « Ces inversions ne sont plus du français spontané. Elles répondent à une intention stylistique [...] ». (Deloffre 1979 : 34). La fréquentation trop assidue des écrits littéraires ne finirait-elle pas par brouiller le sentiment que les linguistes portent sur la langue et l'usage courant ?

c) La répartition entre les différents types de sujets postposés. Les analyses conduites dans d'innombrables travaux sur des données convergentes n'ont-elles pas fait émerger des classifications désormais bien installées dans la représentation des grammairiens ? Certaines oppositions sont fondées sur des paramètres grammaticaux indiscutables. Ainsi, la distinction entre sujet clitique et sujet nominal instaure, quand le sujet est postposé, une coupure pertinente que l'on peut relier à des environnements syntaxiques discriminants. Pour autant, les données orales apportent un éclairage supplémentaire sur cette distinction (voir plus loin). Dans un autre cas, la langue littéraire donne à voir une réalité quelque peu distordue. Il s'agit de l'inversion locative (avec complément locatif / temporel en première position) : son importance dans la littérature scientifique ne peut manquer de surprendre quand on travaille sur la langue parlée...

L'intérêt de recourir à l'oral ne porte pas sur l'espoir de découvrir des exemples bizarres ou à la syntaxe déroutante. Blanche-Benveniste (2010) a souvent développé l'idée que la syntaxe de l'oral ne se différencie pas (ou

³ Pour donner une idée, on trouve presque 9800 exemples avec *c'est* et 6200 fois la forme *a* (dans la séquence *il y a*, mais aussi comme 3^e personne du verbe ou de l'auxiliaire *avoir*), auxquels il faudrait ajouter tous les autres verbes conjugués.

seulement à la marge) de celle de l'écrit. C'est au niveau des fréquences (de la représentation de certains phénomènes) que des écarts se manifestent le plus. Sinclair (1991) avait déjà signalé que les données orales donnent à voir des distributions souvent plus tranchées que l'écrit. En particulier, quelques faits majeurs sont bien représentés et émergent mieux de l'ensemble, alors qu'à l'écrit un certain lissage opère et les cas peu nombreux se multiplient. Benzitoun (2014) a illustré cette idée sur la place qu'occupent les adjectifs épithètes : la langue parlée permet de mieux séparer les grandes tendances des faits marginaux. Tel est l'objectif de ce travail : dégager des grandes tendances⁴ qui concernent le sujet postposé pour mieux discerner certaines contraintes.

Si l'on écarte les rares cas d'inversion complexe (qui ne seront pas analysés comme des sujets postposés), il ne reste plus que 291 cas qui peuvent être répartis en quatre types distincts :

	Sujet postverbal uniquement clitique	Sujet postverbal uniquement nominal	Sujet postverbal mixte	
			compatible avec <i>-t-il</i>	non compatible avec <i>-t-il</i>
Nombre d'exemples	38	136	108	9
Pourcentage	13 %	47 %	37 %	3 %
exemples	(1)	(2)	(3)	(4)

Tableau 1. – Répartition des sujets postverbaux dans le CRFP

Les exemples suivants illustrent les différents cas du Tableau 1 :

- (1) a. toujours est-il que je suis allé voir un projet qui + m'inquiète beaucoup⁵
b. * toujours est cette question que je suis allé voir un projet (inventé)
- (2) a. autour de ce micro environnement gravitent donc tous les acteurs qui sont en relation permanente avec l'entreprise
b. * autour de ce micro environnement gravitent-ils (inventé)
- (3) a. comment se passe l'opération

⁴ Cet objectif conduit à travailler sur un corpus fermé (qui permet de quantifier) et à ne pas retenir d'exemples pris à la volée (qui reflètent la sensibilité du chercheur à tel ou tel fait plutôt que l'usage). Certes le corpus utilisé est lui-même de taille modeste (450 000 mots) et par les caractéristiques qu'il possède, il ne peut prétendre fournir un reflet fidèle de la langue parlée. Mais quel corpus pourrait avoir une telle ambition ? L'appui sur le CRFP laisse tout de même espérer que l'on fournira une photographie des usages ou du moins de multiples usages oraux.

⁵ Sauf indication contraire (exemples cotés b et c), tous les exemples numérotés sont tirés du CRFP. Ils ne sont pas ponctués et le signe + indique une pause.

- b. comment se passe-t-elle (inventé)
 - c. comment l'opération se passe-t-elle (inventé)
- (4)
- a. eh ben je vous attendais me me dit le président
 - b. eh ben je vous attendais me dit-il (inventé)
 - c. * eh ben je vous attendais le président me dit-il (inventé)

Cette présentation conforte la formulation utilisée dans le titre (*les sujets postverbaux*). La désignation au singulier (par exemple dans l'expression "le sujet inversé") est probablement trompeuse ou du moins peu satisfaisante car elle laisse pressentir une unité dont on peut considérer qu'elle n'existe pas. Il est préférable de parler des sujets postposés : ils n'ont pas la même composition paradigmatique, n'apparaissent pas dans les mêmes environnements et ne sont pas utilisés dans les mêmes proportions.

2. BLACKOUT : LES SUJETS POSTPOSÉS FIGÉS À L'ORAL

Si l'on se contente d'une approche quantitative, le sujet postposé uniquement clitique paraît relativement vivant mais un examen plus précis des données montre qu'en fait la quasi totalité des exemples renvoie à des séquences bloquées et utilisées par un nombre très réduit de locuteurs. Les faits essentiels sont regroupés dans le tableau ci-dessous (dans lequel le nombre d'exemples est indiqué entre crochets). Avec ces quelques rubriques, on décrit près de 90 % des cas de sujet postverbal uniquement clitique :

Clitique postverbal	Séquences observées	Nombre de locuteurs concernés
<i>ce</i> ⁶ [6]	<i>ne serait-ce</i> [4]	4
	lexique + <i>est-ce</i> [2]	1
<i>il</i> impersonnel [10]	toujours est-il [7]	5
<i>vous</i> [17]	<i>comprenez-vous</i> [6]	1
	<i>voyez-vous</i> [5]	2
	<i>avez-vous</i> [3]	1

Tableau 2. – Les principaux emplois du sujet clitique postverbal

Le caractère formulaire des structures est assez frappant dans cette partie. Ainsi, le clitique *il* impersonnel se rencontre majoritairement dans la sé-

⁶ La forme bloquée *est-ce que* n'a pas été comptabilisée.

quence bloquée *toujours est-il*⁷ et il en est de même avec le sujet *ce* qui apparaît dans *ne serait-ce*.

À plusieurs reprises, on peut identifier de probables tics langagiers liés à une situation ou à un locuteur particulier. Ainsi, *avez-vous* est utilisé par un guide dans la formule *avez-vous quelques questions*. Le « poids » du locuteur est particulièrement frappant pour le sujet *vous*, puisque toutes les occurrences de *comprenez-vous* dont on dispose sont réalisées par un seul locuteur. À quoi il faut ajouter quatre exemples (sur cinq) de l'expression *voyez-vous*. Ainsi, sur 17 emplois du sujet *vous* postposé, 10 se rencontrent chez un seul locuteur (âgé de 93 ans⁸). Globalement, ces 17 occurrences (qui représentent tout de même 45 % des cas du sujet clitique postverbal présents dans le CRFP) sont le fait de 3 individus différents et n'exploitent que 3 formulations verbales dont deux sont figées ! Un examen détaillé modifie profondément l'impression première appuyée sur un simple relevé quantitatif. Le corpus structuré constitue ici un garde-fou précieux qui évite d'accorder une place démesurée à un phénomène extrêmement marginal⁹. On se trouve, de fait, dans une observation de l'« infiniment petit¹⁰ » qui nécessite une infinie prudence...

La plupart des exemples déjà signalés relèvent d'une syntaxe figée : le clitique postverbal fait partie d'un bloc qui n'est pas réellement décomposable et dont les éléments ne semblent pas appartenir à des séries qui rendraient possibles des commutations. Ainsi, à partir de *toujours est-il*, on pourrait imaginer de rencontrer quelques séquences comme :

toujours	est-	il
...
...	était-	il

À partir de *avez-vous quelques questions*, on pourrait s'attendre à trouver *auriez-vous quelques questions*, *souhaitez-vous des compléments*, etc. qui ne se rencontrent pas. Voilà qui incite à parler de *figement* ou de *formule* pour caractériser ces tours qui ne connaissent pas, dans le corpus, de variation de temps ou de lexique.

Pour les clitiques *on* et *nous*, on peut reprendre l'expression de Moreau (1986) et envisager une association privilégiée (avec le verbe *dire* voire avec

⁷ Sur ces constructions à adjectif initial, on renvoie à Guimier (1997).

⁸ L'âge est un paramètre que le CRFP permet d'exploiter. Cela n'a pas été fait de façon systématique dans ce travail.

⁹ Coveney (2011) indique que des comptages sur corpus des interrogatives directes ont déjà conduit à ces conclusions. Ledegen & Quillard (2000) avançaient le même constat.

¹⁰ Selon une formule de Jean Véronis lors de sa communication (non publiée) aux 4^{ème} journées de la linguistique de Corpus (Lorient 2005).

le modal *pouvoir*) mais il faudrait, pour approfondir cette piste, disposer de plus d'exemples :

- (5) aussi + conséquence de ceci n'est-ce pas + pouvons-nous dire me semble-t-il que la critique
- (6) parce que euh au cours de notre X + peut-on dire maintenant de notre longue carrière

Finalement, le nombre d'exemples qui possèdent une syntaxe libre (c'est-à-dire qui n'est ni incluse dans une formule, ni limitée à un verbe particulier) est quasi insignifiant. Si l'on met à part le cas de *avez-vous* (qui fait partie de l'expression *avez-vous quelques questions*), il ne reste plus qu'un exemple avec *tu* :

- (7) et as-tu l'impression d'avoir les mêmes relations avec tous les clients

Mais c'est un exemple chimérique (sorte de compromis entre deux tensions contradictoires) qui est lié à la situation d'interview et où le locuteur doit combiner le lien amical avec l'interlocuteur et le caractère quelque peu formel de l'enregistrement d'où l'emploi (unique dans l'ensemble du corpus) de cette question avec inversion du sujet *tu*.

3.1 FEEL FREE : LES CONSTRUCTIONS LIBRES À L'ORAL

Les sujets postposés uniquement nominaux constituent le groupe le mieux représenté dans les données orales recueillies. On trouve des exemples de ce type dans 67 corpus (ce qui correspond à environ 2/3 des corpus du *CRFP* qui contiennent un ou plusieurs sujets postverbaux). Il s'agit donc aussi de sujets bien disséminés dans nos données. Il est utile d'examiner plus attentivement les contextes syntaxiques dans lequel apparaissent ces sujets.

Si l'on retient, pour faciliter la discussion, la notion traditionnelle de subordonnée (qui repose généralement sur la présence d'un morphème introducteur), on constate que 75 % des sujets nominaux postverbaux s'y trouvent. Mais comme on va le voir, il est préférable de procéder à une analyse plus fouillée pour mieux appréhender la distribution. On trouve assez majoritairement (près des 2/3) les sujets lexicaux postverbaux dans des relatives. Deux contextes méritent un examen plus précis :

3.1. *Too Dizzy* : Les structures les plus fréquentes

Les sujets nominaux postverbaux se rencontrent principalement dans 4 organisations pour lesquelles on peut repérer certains faits récurrents liés soit à la structure, soit au lexique.

a) Les relatives en *que*

Elles sont largement dominantes, ce qui recoupe des observations de Nordhal (1973) portant sur l'écrit. Cette distribution souligne que la formulation de nombreux auteurs, lorsqu'ils traitent des propriétés de construction des verbes qui possèdent un sujet postverbal, pêche bien souvent par imprécision. Ainsi, quand Gardes-Tamine (2003) écrit qu'il faut « que le cadre syntaxique ne propose que deux cases » (i.e. dans sa présentation un verbe et un sujet mais pas un complément direct) et qu'elle l'illustre dans l'exemple : * *Voit un garde le prisonnier* (cité p. 59). En fait, c'est plus la catégorie dans laquelle est réalisé le complément direct qui est pertinente que la position syntaxique elle-même. Les exemples avec des relatives soulignent fortement la nécessité de distinguer ces deux paramètres. La présence d'un complément direct n'est plus un obstacle dès lors qu'il est réalisé sous la forme du relatif *que*, c'est même le contexte le plus fréquent.

Si l'on tient compte des verbes utilisés dans ces environnements, il semble pertinent de différencier les cas où l'antécédent est un nom de ceux où il s'agit de la proforme *ce* :

– Avec un antécédent nominal, 3 verbes se partagent presque la moitié des emplois. Il s'agit de *avoir* qui est le plus employé, de *être* et de *faire*. Cet ordre d'apparition et cette importance quantitative ne sont toutefois pas vraiment surprenants puisqu'il s'agit là des verbes qui sont globalement les plus fréquents à l'oral :

- (8) entre les trois parties + de la formation que sont l'apprenti + les les les responsables de l'école et le maître d'apprentissage

– Lorsque l'antécédent est *ce*, le choix des verbes retient plus l'attention. En effet, on trouve d'abord *faire*, puis *être* et *dire*. Si l'on peut envisager qu'il y a une sorte de sélection, d'association privilégiée, il n'y a pas ici de syntaxe bloquée au sens employé précédemment. En effet, le verbe est utilisé sous des formes multiples (temps, personne, ajout de modal). Quelques exemples avec *dire* peuvent illustrer cette variété :

- (9) il y en a qui demandent à assister [...] à ce que dit maman
 (10) enfin c'est ce que disait ma mère d'ailleurs très précisément
 (11) ah ben pour l'Indien de base moi ce que m'en ont dit les Français qui étaient là-bas

Il est bien difficile à partir des données recueillies de rattacher la séquence *ce que + verbe + sujet* à une situation d'enregistrement particulière. Le corpus permet de rechercher si le caractère plus formel de l'échange,

induit à partir de la situation de parole publique (PUB) ou professionnelle (PRO) pouvait jouer¹¹. Le tableau suivant recense les cas rencontrés :

	Nombre total d'occurrences	Nombre dans PRO & PUB
<i>ce que + dire</i>	6	3
<i>ce que + être</i>	8	4
<i>ce que + faire</i>	10	2
<i>ce que + autre verbe</i>	2	0

Tableau 3. – Répartition des constructions en *ce que*

Même si cela reste fragile et très hypothétique, il est possible d'esquisser, en l'état, les tendances suivantes :

– *ce que + dire* ou *être* pourraient être des marqueurs de langue formelle, à relier éventuellement à des genres professionnels (Blasco-Dulbecco & Cappeau 2012) ;

– *ce que + faire* ne semble pas posséder la même propriété et serait donc plus “banal”, il ne constituerait pas un marqueur de genre.

Pour approfondir cette piste, il faudrait aussi entrer dans le détail des enregistrements classés comme parole privée pour vérifier le caractère formel ou non de l'échange, ce qui n'a pas été fait. L'observation d'un corpus d'interviews politiques (de 300 000 mots) confirme l'hypothèse envisagée. Le Tableau 4 présente la répartition de certaines séquences :

	Nombre total d'occurrences	Pourcentage
<i>ce que + dire</i>	34	32 %
<i>ce que + être</i>	31	29 %
<i>ce que + faire</i>	14	13 %
<i>ce que + autre verbe</i>	28	26 %

Tableau 4. – Les constructions en *ce que* dans un corpus d'interviews politiques

La séquence *ce que + dire* se rencontre de façon assez importante dans des corpus de langue formelle, et dans le cas du corpus d'interviews politiques, elle indique un mode de citation des paroles d'autrui. Cette observation sur les sujets postverbaux fait partie d'un ensemble plus vaste qui

¹¹ Les sous-corpus de parole privée (PRI) représentent 63 % du corpus, ceux de parole professionnelle (PRO) 18 % et ceux de parole publique (PUB) 19%. Ces éléments permettent de mieux apprécier les indications quantitatives fournies dans le Tableau 3.

corrobore la piste suivie. Si l'on regarde ce qui se passe pour l'ordre *sujet + verbe + dire*, on relève ainsi que la séquence *ce que + sujet + [...] dit* est rare dans le CRFP (11 exemples) et bien plus répandue dans le corpus d'interviews politiques (31 cas) qui est pourtant un corpus de taille moindre (d'environ 1/3). Cela laisse penser que les modes de citation, les formules introductrices des paroles rapportées constituent un paramètre sensible qui peut être relié à des caractéristiques formelles ou de genre des productions dans lesquelles ces tours apparaissent.

b) Les relatives en *où*

De façon plus inattendue, le relatif *où* n'est pas associé à une série précise et répétée de verbes, alors qu'il aurait pu sembler plausible que le sémantisme locatif du pronom soit lié à quelques verbes récurrents. La seule régularité qui émerge des données (mais qui reste pour l'instant impossible à interpréter) concerne les sujets qui sont très majoritairement au pluriel (8 sur 10). Là encore, le nombre d'exemples conduit à rester prudent :

- (12) ici vous trouvez l'ensemble + des des lieux où vont se stocker les métaux dans les différents organes du corps humain
- (13) les grandes caves où se trouvaient d'immenses cuves

Pour conclure cette partie sur les relatives, le Tableau 5 permet de constater que les relatives pertinentes sont bien représentées et ne sont pas regroupées dans un nombre trop limité de corpus mais sont bien dispersées dans l'ensemble des données utilisées.

Introduceur	Nombre d'exemples	Nombre de corpus
N <i>que</i>	35	25
<i>ce que</i>	28	25
N <i>où</i>	10	9
autres cas	15	14

Tableau 5. – Emplois d'un sujet lexical postverbal après un relatif

c) Les clivées

Après les relatives, les constructions clivées constituent la deuxième structure la plus fréquente. Un constituant est placé entre *c'est... que* et le sujet postposé se réalise dans la deuxième partie de la construction (celle qui suit *que*) :

- (14) c'est avec la la la régionalisation que se sont développés euh ces les les les médias euh natio- euh nationaux

Là encore, le verbe *être* est celui que l'on rencontre le plus fréquemment (dans la deuxième partie de la construction) :

- (15) c'est là qu'est notre responsabilité + politique
- (16) c'est là qu'est toute la difficulté

d) Les structures en *comme*

Immédiatement (du point de vue quantitatif) après cette organisation, on trouve une structure dans laquelle on peut relever un phénomène d'association privilégiée. Il s'agit de constructions en *comme* dans lesquelles le verbe *dire* apparaît à 7 reprises (sur les 10 cas relevés). Toutefois, il ne s'agit pas d'une structure figée puisque, une nouvelle fois, le verbe se présente sous des formes multiples :

- (17) alors donc comme l'a dit Danielle on est toutes les deux euh étudiantes
- (18) voilà comme disait madame elle avait raison
- (19) comme diraient les moralistes

Il se confirme avec ces derniers exemples que certains verbes se rencontrent plus fréquemment avec des sujets postverbaux, mais on voit aussi que ce paramètre doit être affiné et adapté aux contextes syntaxiques observés. Ainsi pour *dire*, un décompte global serait peu intéressant (vu le nombre très élevé d'occurrences de ce verbe à l'oral) alors qu'un relevé ciblé sur les structures telles que *comme* + verbe + sujet ou *ce que* + verbe + sujet fait ressortir une distribution originale de ce verbe.

3.2. *Don't Be Afraid* : Les structures peu observées

Il n'y aurait pas forcément lieu d'évoquer les constructions rares puisque l'on a choisi de mettre en avant les faits les plus représentés. Mais ici la surprise tient au décalage entre les analyses linguistiques qui ont abondamment traité de certaines tournures et leur rareté dans notre corpus.

a) L'inversion absolue

On ne rencontre que 4 cas dans lesquels le verbe n'est précédé d'aucun constituant régi comme dans :

- (20) donc euh justement euh viennent d'arriver euh Christian Faniard + vice-président d'EDA qui habite Aubry et Gérard Debreyll par qui le dossier est arrivé

Dans cet exemple, l'adverbe *justement* n'est pas régi mais doit être analysé comme un associé, ce qui est le fonctionnement habituel de cette forme (Bilger 2004). On rencontre souvent des verbes dits de survenance (*arriver, rester*) et les sujets sont développés (Abeillé 2002). Sans entrer dans le

détail, on peut noter que, dans bien des cas, la structure du sujet postverbal est « simple » (un déterminant + un nom) mais que, dans certains contextes comme celui étudié ici, des structures plus étoffées (liste de syntagmes, sujet avec relative) peuvent être observées :

- (21) eh bien euh s'ensuit un mouvement de de succion de va-et-vient des des crochets

b) L'inversion locative

Pour les exemples avec syntagme nominal en tête, ils sont un peu plus nombreux même si leur total reste faible (9 cas rencontrés). Un facteur semble émerger et concerne la valeur locative du complément plus exploitée (2/3 des emplois) que la valeur temporelle. Mais il faudrait disposer de données plus nombreuses pour pouvoir avancer dans cette voie qui n'est peut-être due qu'aux thèmes traités dans les corpus sources :

- (22) à côté d'ailleurs + ce- + du lieu où où j'avais le travail + se trouvait la clinique des Glycines dans laquelle d'ailleurs sont nés mes trois enfants

Comme on l'a signalé précédemment, l'importance que prend cette inversion dans les analyses linguistiques étonne au vu de la place réduite qu'elle occupe dans l'usage oral.

4. DAY IN DAY OUT : LES CAS DE SUJET POSTVERBAL MIXTE

Cette distribution correspond à deux classes distinctes qui peuvent être différenciées, en ce qui concerne les propriétés formelles, grâce à l'inversion complexe. Ces deux grands types possèdent aussi des valeurs sémantiques différentes : certaines interrogatives en 4.1 et les incises en 4.2.

4.1. Les interrogatives partielles

Lorsque la structure autorise que le sujet postverbal soit clitique tout autant que nominal, on constate tout de même une forte disparité des sujets effectivement réalisés. Les sujets nominaux représentent alors 70 % des cas.

En ce qui concerne les sujets clitiques, les observations que l'on peut faire vont dans le même sens que celles indiquées précédemment lorsque seuls les sujets clitiques étaient envisageables : dans de nombreux cas, les énoncés sont figés, les locuteurs emploient en fait des formules toutes faites comme *comment dirais-je, que voulez-vous*¹². Restent tout de même des cas

¹² Il n'est peut-être pas inutile de rappeler que ces énoncés sont classés dans cette rubrique des inversions mixtes, même si seul un clitique est attesté, parce que la structure se prête à une alternance entre clitique et lexique dès lors que l'on recourt à des formes qui s'y prêtent. Ainsi *comment dirais-je* est dans cette classe des postverbaux mixtes parce que s'y trouvent aussi *comment dirait-il* et *comment dirait ton frère*.

de syntaxe libre. Tout dépend en fait du clitique qui se trouve en position postverbale :

– *je* n'a été rencontré que dans la séquence bloquée *comment dirais-je*. Les 12 exemples recueillis ont été produits par seulement 5 locuteurs :

(23) on a euh m- comment dirais-je préféré des des poètes très marqués

– le pronom *il* est assez souvent relié à l'interrogatif *que*. C'est d'ailleurs presque exclusivement le pronom impersonnel qui est employé dans cet environnement :

(24) mais enfin que reste-il en ce temps là hein

– les autres formes sont assez peu nombreuses, ce qui empêche de discerner des régularités.

Si l'on regarde les interrogatifs, les deux formes les plus fréquentes sont *comment* (surtout à cause de *comment dirais-je*) et *que* qui a la particularité d'apparaître avec la plupart des clitiqes. Enfin, *qui* (rencontré à 3 occasions) est toujours utilisé avec le verbe *être* :

(25) et votre public euh + qui est-il et comment ça se passe

Lorsque le sujet est nominal, la structure la plus développée est celle qui combine l'interrogatif *quel*¹³ et le verbe *être*. Plus de la moitié des exemples se rattachent à cette organisation. La forme des sujets elle-même est assez spécifique. Comme l'indique Blanche-Benveniste (2002), le sujet doit être défini mais en outre on observe que c'est dans ce contexte que se regroupe la majorité des sujets dits « lourds ». Comme on l'a signalé, le sujet est le plus souvent un syntagme nominal binaire (composé d'un déterminant et d'un nom), ce qui n'est pas surprenant puisque c'est la structure du SN la plus fréquente (Hug 1989). Rien de tel dans la configuration *quel est* : près de 70 % des sujets ont une structure plus étoffée. Les sujets qui incluent une relative sont notamment bien représentés :

(26) quels sont les les grands noms que vous aimeriez faire connaître un peu

(27) quelles sont vos les les branches qui vous qui vous intéressent le plus

Il faut noter que cette structure est très liée à la situation d'interview et qu'elle est souvent employée par l'intervieweur. Elle relève donc d'une situation très précise.

Autre surprise, qui tient probablement au contenu des échanges, sur les sujets qui figurent dans cet environnement : les formes au pluriel sont très représentées (40 %). C'est un cas à part dans l'ensemble des données recueillies et cette proportion est aussi plus élevée que le rapport habituel entre singulier et pluriel dans les corpus (de l'ordre de 2/3 de formes au singulier).

¹³ Il s'agit du pronom *quel*. Le déterminant n'est utilisé qu'à 3 reprises.

Deux autres interrogatifs sont fréquents dans nos données avec un sujet postverbal nominal : *comment* et *où* (qui connaît une variante *d'où*). Pour l'interrogatif *comment*, on pourrait, si l'on disposait de plus de données, améliorer la description et distinguer entre la question directe (utilisée par l'intervieweur) et l'emploi en structure régie (un peu moins fréquente dans nos exemples) qui peut aussi être employée par l'interviewé¹⁴ :

- (28) et comment est structurée l'agence
 (29) je pense que Bernard vous a montré un peu comment étaient euh organisés les lieux

Dans le reste des exemples, on peut extraire la structure avec *en quoi* qui est toujours utilisée avec le verbe *consister*. Mais sur un total de 3 exemples, le constat doit rester mesuré :

- (30) en quoi consiste mon travail et comment je le vis

4.2. Les incises

Les constructions qui possèdent ces deux caractéristiques (sujet postverbal mixte et absence d'inversion complexe) correspondent aux incises :

- (31) eh ben je vous attendais me me dit le président je j'attendais votre réaction

Ces structures sont très peu représentées dans notre corpus. Ce n'est pas véritablement surprenant puisque les verbes introducteurs de parole sont très fréquemment placés en tête d'énoncé (ce qui écarte la possibilité d'un sujet postverbal) :

- (32) j'arrive et je je je te dis ben j'ai mal au dos

Il est difficile d'établir une distribution avec un nombre très réduit d'exemples. Sur cette rubrique, la différence est forte entre ce qu'indique un corpus écrit littéraire et ce qu'apporte un corpus oral. Toutefois le choix des verbes (souvent impersonnels) montre que la distinction entre ce groupe et celui des sujets uniquement clitiques demanderait à être approfondie.

La distribution obtenue à partir du paradigme des formes sujets postverbales pourrait, en fait, être modifiée en passant des réalisations virtuelles aux réalisations observées. Pour l'instant, il semble préférable de conserver les quatre classes dégagées car elles renvoient à d'autres propriétés couplées. Ainsi, on observe bien une différence entre les structures qui ont un sujet clitique postverbal :

¹⁴ La coupure opérée par la norme concernant le sujet clitique postposé – autorisé en principale et écarté en subordonnée – ne semble pas très pertinente lorsque l'on travaille sur des données attestées (aussi bien orales qu'écrites). Elle n'a donc pas été retenue ici.

– Lorsque seul le clitique peut être postposé, l'énoncé prend une valeur non assertive (très souvent interrogative) :

(33) y a-t-il une transparence

(34) ensuite a-t-elle mis en œuvre ses engagements

– Rien de tel lorsque le sujet postposé maintient une bonne relation entre clitique et lexique (et que l'inversion complexe est écartée), la modalité non assertive est absente. Cette caractéristique est conservée même quand le sujet est impersonnel :

(35) parce que Charlemagne euh + paraît-il ne buvait que des rouges

(36) il y a eu semble-t-il plusieurs constructions

5. REALITY : CONCLUSION

Au terme de ce rapide examen de la totalité des sujets postverbaux présents dans le corpus *CRFP*, toutes les pistes envisagées n'ont pas été suivies. Un certain nombre d'observations ont tout de même pu être dégagées. Elles sont reprises, pour l'essentiel, dans le tableau suivant :

Sujet postverbal clitique [16 %]		Sujet postverbal nominal [84 %]	
Pour l'essentiel une syntaxe bloquée ou limitée à quelques verbes particuliers		Syntaxe libre même si l'on note des associations privilégiées	
clitique seul	postposition mixte (clitique / lexique) – [t-il]	+ [t-il]	SN seul
13 %	3 %	37 %	47 %
Syntaxe bloquée. Des tournures figées : <i>toujours est-il</i> <i>ne serait-ce</i>	Syntaxe réduite limitée à quelques verbes : <i>dire...</i>	– les clitiques entrent dans une syntaxe réduite – les SN font partie d'une syntaxe libre	– la structure dominante est la relative en <i>que</i> (<i>N que</i> ou <i>ce que</i>)

Tableau 6. – Contraintes sur les divers types de sujets postverbaux (dans le *CRFP*)

Cet inventaire offre, semble-t-il, une meilleure vision du sujet postverbal en français et permet de faire converger des analyses multiples aux tendances parfois disparates. L'une des impressions les plus frappantes est que cohabitent des formes figées dans lesquelles on peut d'ailleurs se demander s'il est utile de se livrer à un examen interne et d'identifier un sujet (comme dans *toujours est-il*), des structures qui se rencontrent fréquemment avec une liste réduite de verbes (*comme + dire* par exemple) et des structures dans

lesquelles le choix des verbes est plus ouvert et où la syntaxe semble libre (modulo des contraintes sur, notamment, la catégorie réalisée comme sujet). C'est autour de ces trois degrés que semble organisé le sujet postposé. On peut ainsi mieux faire ressortir qu'il existe plusieurs cas de sujets postverbaux et qu'on se trouve en présence de plusieurs systèmes en grande partie spécifiques.

Le fait d'avoir utilisé un corpus relativement bien délimité offre l'opportunité de repérer des récurrences d'emploi et de les relier à des situations de parole ou à des contextes plus particuliers. Il devient possible d'envisager (certes avec prudence) que certaines tournures sont des marqueurs de genres (hypothèse qui demanderait à être développée). D'autre part, en travaillant sur un corpus assez développé et varié, la description obtenue s'approche beaucoup plus de l'usage courant¹⁵ (appréhendé via l'oral) que des observations accumulées sur des exemples de provenances diverses ou trop homogènes.

BIBLIOGRAPHIE

- ABEILLÉ A. (2002). *Une grammaire électronique du français*. Paris : CNRS Éditions.
- BENZITOUN C. (2014). La place de l'adjectif épithète en français : ce que nous apprennent les corpus oraux. CMLF. Téléchargeable à : <http://dx.doi.org/10.1051/shsconf/20140801066>.
- BILGER M. (2004). Quelques données sur les adverbes en *-ment* dans le corpus de référence du français parlé. *Recherches sur le français parlé* 18, 63-82.
- BLANCHE-BENVÉNISTE C. (2002). « Le mot *QUEL* ». *Verbum* XXIV-4, 363-373.
- BLANCHE-BENVÉNISTE C. (2010). *Le français : Usages de la langue parlée*. Leuven : Peeters.
- BLASCO-DULBECCO M., CAPPEAU P. (2012). « Identifier et caractériser un genre : l'exemple des interviews politiques ». *Langages* 187, 27-40.
- BURIDANT C. (2000). *Grammaire nouvelle de l'ancien français*. Paris : Sedes.
- CAPPEAU P. (2007). Constituer et exploiter un gros corpus oral : l'exemple du CRFP. In : M. Ballard, C. Pineira-Tresmontant (éds), *Les corpus en linguistique et en traductologie*. Arras : Artois Presses Université, 145-158.
- CAPPEAU P., LAHOUSSE K. (2015). Le sujet postposé. Fiche de l'EGF. http://encyclogram.fr/not/005/005_Notice.php

¹⁵ Bien que les facteurs peu pris en compte (comme l'âge) jouent certainement un rôle important.

- CHEVALIER J.-C., (1985). Les grammaires françaises et l'histoire de la langue. In : G. Antoine, R. Martin (éds), *Histoire de la langue française 1880-1914*. Paris : Éditions du CNRS, 577-600.
- COVENEY A. (2011). L'interrogation directe. *Travaux de linguistique* 63, 112-145.
- DELIC (2004). Présentation du Corpus de référence du français parlé. *Recherches Sur le Français Parlé* 18, 11-42
- DELOFFRE F. (1979). *La phrase française*. Paris : Sedes.
- FUCHS C. (éd.). (1997). *La place du sujet en français contemporain*. Louvain-la-Neuve : Duculot.
- GARDES TAMINE J. (2003). La place du sujet en français. In : J.-M. Merle (éd.), *Le sujet*. Paris / Gap : Ophrys, 51-62.
- GOUGENHEIM G. (1974). *Grammaire de la langue française du 16^e siècle*. Paris : Éditions Picard.
- GUIMIER C. (1997). La place du sujet clitique dans les énoncés avec adverbe initial. In : C. Fuchs (éd.), *La place du sujet en français contemporain*. Louvain-la-Neuve : Duculot. 43-96.
- HAASE A. (1914). *Syntaxe française du XVII^e siècle*. Paris : Delagrave. (Nouvelle édition par M. Obert, 1975)
- HUG M. (1989). *Structures du syntagme nominal français*. Paris / Genève : Champion / Slatkine.
- JONARE B. (1976). *L'inversion dans la principale non-interrogative en français contemporain*. Stockholm : Almqvist & Wiksell.
- LAHOUSSE K. (2011). *Quand passent les cigognes. Le sujet nominal postverbal en français moderne*. Paris : Presses Universitaires de Vincennes.
- LE BIDOIS R. (1952). *L'inversion du sujet dans la prose contemporaine (1900-1950)*. Paris : d'Artrey.
- LEDEGEN G., QUILLARD V. (2000). Quelle référence pour l'interrogation totale ? Réponses syntaxiques et pragmatiques. *Cahiers de l'institut linguistique de Louvain* 26, 1-4, 297-312.
- MOREAU M.-L. (1986). Les séquences préformées : entre les combinaisons libres et les idiomatismes. Le cas de la négation avec ou sans *ne*. *Le Français Moderne* 3-4, 137-160.
- NORDHAL H. (1973). Inversion et progression dans la subordonnée relative en français moderne. *Le Français Moderne* XLI-2, 113-130.
- ROIG A. (2017). L'inversion (pro)nominal du sujet : un indice syntaxique ? *Verbum* (ce numéro).
- SCHNEDECKER C. (2005). Quand "un certain David Bowie repoussait les limites du bon goût..." ou que modifie *certain* dans les séquences *un + certain + nom propre* ? *Langue française* 146, 99-113.
- SEGUIN J.-P. (1972). *La langue française au XVIII^e siècle*. Paris : Bordas.
- SINCLAIR J. (1991). *Corpus Concordance Collocation*. Oxford : Oxford University Press.