Les relatives en *dont* du français : études empiriques

Abeillé, Anne¹ & Hemforth, Barbara² & Winckel, Elodie^{3,a}

^{1,2}Laboratoire de Linguistique Formelle (LLF)

³Freie Universität Berlin

abeille@linguist.univ-paris-diderot.fr, barbara.hemforth@linguist.univ-paris-diderot.fr, elodie.winckel@fu-berlin.de

Résumé. Les relatives en *dont* ont fait l'objet de plusieurs études théoriques (Godard 1988) mais ont moins été étudiées d'un point de vue empirique (Gapany 2004). C. Blanche-Benveniste (1995) notait leur rareté dans des corpus oraux, où ils sont principalement compléments de certains verbes. Notre objectif est d'observer le comportement des relatives en *dont* en français d'un point de vue empirique. Pour ce faire, nous avons en premier lieu étudié les relatives dans deux corpus, le French Treebank (corpus écrit) et le Corpus de Français Parlé Parisien des années 2000 (corpus oral). Ces études nous ont permis de confirmer en partie les travaux de C. Blanche-Benveniste (1995) : l'usage de *dont* en français écrit diffère effectivement de son usage en français oral, mais *dont* est loin d'être réservé à l'oral à quelques expressions figées. Nous avons pu aussi confirmer l'étude de Godard (1980) : les *dont* complément de nom (extractions hors de SN) ne sont pas marginaux, et les plus fréquents sont ceux où le nom est sujet, ce qui considéré comme agrammatical par certaines théories linguistiques comme agrammaticales. De ce point de vue, les corpus nous ont permis de vérifier que les extractions hors du SN sujet ne posent pas de problèmes en français. Leur usage est même plutôt répandu, et inclut les extractions hors de « vrais sujets » de verbes transitifs, contrairement aux prédictions des hypothèses de la grammaire générative. En grammaire générative en effet, les relatives comme les autres structures à extraction sont soumises à des contraintes d'îles (Ross 1957). Selon la contrainte de l'îlot nominal, il est plus difficile d'extraire un complément de nom qu'un complément de verbe. La contrainte de l'îlot sujet exclut quant à elle l'extraction du complément du sujet, qui n'est possible que lorsque le sujet est dérivé (passif, moyen...) ou argument interne (Chomsky 2008). Or, nous constatons que les extractions hors de SN sujet sont plus fréquentes dans les corpus que les extractions hors de SN objet (en ce qui concerne les relatives en *dont*). Ce même fait est un argument en faveur de la théorie de la localité de Gibson, qui prévoit que l'extraction hors de SN sujet est non seulement bien formée mais moins coûteuse que celle hors de SN objet. D'autres observations sur les relatives en *dont* des deux corpus semblent corroborer cette théorie, comme par exemple la plus grande fréquence des sujets pronominaux ou inversés dans les relatives en *dont* dans lesquelles *dont* correspond à un complément du verbe. Dans un deuxième temps, nous avons mené une étude expérimentale de jugement d'acceptabilité afin de comparer les jugements de locuteurs natifs face à des relatives en *dont* correspondant à une extraction hors du SN sujet d'une part et à une extraction hors de SN objet d'autre part (la forme du sujet dans ces dernières, qui joue un rôle dans la théorie de la localité de Gibson, a été testée également). Les résultats de cette expérience sont eux aussi en accord avec les prévisions de la théorie de la localité de Gibson, les jugements étant significativement supérieurs pour les extractions hors de SN sujet par rapport aux extractions hors de SN objet.

Abstract. French relative clauses introduced by *dont* (from now on: *dont*-RC) have been well investigated on the theoretical level (Godard 1988) but have not lead to many empirical studies (Gapany 2004). Blanche-Benveniste (1995) observed that *dont*-RC are not very common in spoken French and restricted to the extraction of a few verb complements. The aim of this article is to describe *dont*-RC on the basis of three empirical studies: two corpus studies and one online experimental study. The corpus studies have been conducted on a written corpus (French Treebank) and on a spoken corpus (Corpus de Français Parlé Parisien des années 2000). The results confirm partly Blanche-Benveniste's claim: the written usage of *dont*-RC differs from its spoken use, but it is not restricted in the later to specific idioms.

^a L'activité d'E. Winckel dans le cadre des îlots syntaxiques est liée au projet financé par la DFG *Long-distance Dependencies in French: Comparative Analyses (HPSG and the Minimalist Program)*, G. Mensching, Göttingen [ME 1252/14-1] / S. Müller, Berlin [MU 2822/9-1].

Both corpora corroborate the work of Godard (1980) : *dont* is widely used as a complement on noun, extractions out of NP are not marginal, not even out of subject NP which are considered non-grammatical by some linguistic theories. The corpus studies show that this kind of extraction is widely-used in French and even concern “real subjects” of transitive verbs, which goes against the prediction made by many generative theories. In generative grammar, relative clauses (as well as any extraction) are subject to island constraints (Ross 1957). The nominal island constraint predicts that the extraction of a noun complement should be harder than the extraction of a verb complement. The subject island constraint rules out extractions out of subject NPs, which is claimed to be only acceptable when the subject NP is an underlying object (Chomsky 2008). However, in our corpora, *dont*-RC extractions out of NP subject are more frequent than extractions out of NP object. This observation confirms Gibson's Dependency Locality Theory (DLT), which predicts that the extraction from subject NP should have less cognitive costs than the extraction from object NP. Other observations corroborate this theory as well: for example, that *dont*-RC corresponding to the extraction of the complement of a verb favor pronominal or inverted subjects. Our experimental study is an online acceptability task. Its aim is to compare the judgments from native speakers related to *dont*-RC implying an extraction out of subject NP and object NP (the kind of subject involved in the later case has an impact of the DLT's predictions and have been tested as well). The results confirm again the predictions from the DLT: the judgments for the extraction out of subject NPs are significantly higher than for the extraction out of object NPs.

1 Introduction

Les relatives en *dont* ont fait l'objet de plusieurs études théoriques (Godard 1988) mais ont moins été étudiées d'un point de vue empirique (Gapany 2004). C. Blanche-Benveniste (1995) notait leur rareté dans des corpus oraux, où ils sont principalement compléments de verbes. Un premier objectif des études que nous présentons ici sera donc d'étudier les usages où *dont* correspond à un complément de nom, et de le contraster avec ceux où il correspond à un complément de verbe.

D'autre part, les relatives dans lesquelles *dont* est un complément de nom se révèlent être un cas intéressant dans l'analyse des phénomènes d'extraction en français. Il existe une littérature fournie consacrée aux extractions hors de syntagme nominal (SN) sujet (Deane 1991, Sabel 2002, Chaves 2013, ...) qui semble problématique dans certaines langues, comme l'anglais². Il existe plusieurs études empiriques sur l'anglais et sur d'autres langues (Polinsky et al. 2013, Sprouse et al. 2015), mais nous n'avons pas connaissance d'études similaires réalisées sur le français. Notre deuxième objectif sera par conséquent de contraster les cas où *dont* correspond au complément d'un nom sujet et ceux où il correspond au complément d'un nom objet.

Nous présentons ici les résultats de deux études de corpus et d'une expérience. La première étude de corpus est menée sur un corpus écrit, le *French Treebank* (FTB, Abeillé et al. 2003), un corpus d'environ 800 000 mots regroupant des articles tirés du journal *Le Monde* et publiés entre 1989 et 1995. La deuxième étude de corpus est, elle, menée sur un corpus oral, le *Corpus de Français Parlé Parisien des années 2000* (CFPP2000, Branca-Rosoff et al. 2012), un corpus d'un total de 37h45 d'enregistrement pour 29 interviews longues réalisées depuis 2005 auprès d'habitants de Paris et de sa périphérie. Ces études de corpus sont complétées par une expérience ayant comme variable dépendante des jugements d'acceptabilité. Cette expérience se concentre sur les cas où *dont* correspond à un complément du nom.

2 Trois hypothèses sur les relatives en *dont*

2.1 Les relatives en *dont* et l'usage

Les travaux de Blanche-Benveniste (1980, 1990, 1995) sur les relatives s'appuient sur des corpus oraux. Elle note que *dont* est un relatif unique dans les langues romanes, et qu'il est peu employé en français parlé, où l'usage non normatif le remplace par *que* :

(1) a. ! le livre que j'ai besoin

b. le livre dont j'ai besoin

Elle note que son usage oral est concentré sur peu de verbes (*parler, s'agir...*) ou dans des tours figés (*la façon dont, la manière dont...*). Elle mentionne toutefois quelques tours quantitatifs (*dont certains, dont beaucoup...*). Une hypothèse est que son emploi écrit soit dû à une pression artificielle de la norme.

2.2 Les relatives en *dont* et les contraintes de la grammaire générative

En grammaire générative, les relatives comme les autres structures à extraction sont soumises à des contraintes d'îles (Ross 1957). Selon la contrainte de l'îlot nominal, il est plus difficile d'extraire un complément de nom (2b) qu'un complément de verbe (2a). La contrainte de l'îlot sujet exclut quant à elle l'extraction du complément du sujet (2c).

- (2) a. This is the island we talked about _ last week.
 b. ?? This is the island we heard [a story about _] last week.
 c. * This is the island [a story about _] has caused a great scandal.

Ces contraintes d'îles ont été remises en cause, notamment pour le français où *dont* permet justement l'extraction d'un complément de nom comme dans l'exemple 3 (Godard 1980). Néanmoins, cette extraction est soumise à des contraintes sémantiques (Godard 1993, Godard et Sag 1996) et discursives (Erteschick-Shir 2007) : il faut notamment que le complément de nom extrait (ici l'antécédent de *dont*) puisse être un topique de discours.

- (3) a. un fruit dont j'aime le goût
 b. un fruit dont le goût est apprécié
 c. un fruit dont ressort le goût

La contrainte de l'îlot sujet a été formulée différemment selon les versions de la théorie générative (souvainance, barrière). Selon la théorie des phases (Chomsky 2008), la contrainte de l'îlot sujet n'est observée que lorsque le verbe construit un syntagme v^*P , c'est-à-dire lorsque le sujet est un argument externe. Les sujets internes (ou objets profonds) ne sont pas soumis à cette contrainte. D'après cette théorie, il est possible d'extraire le complément du sujet d'un verbe inaccusatif (3c), médiopassif ou passif (3b). Les syntagmes v^*P et CP représentent des phases pendant lesquelles l'information est transférée aux interfaces phonétiques et sémantiques.³ Dans le cas d'une extraction du complément du verbe (2a) ou hors d'un SN objet (2b), le complément auquel correspond *dont* se trouve dans le domaine de v^* qui peut le faire remonter en frontière de v^*P . Ainsi, le complément sera accessible pour C. A contrario, pour une extraction hors du SN sujet d'un verbe transitif, v^* ne peut faire remonter le complément auquel correspond *dont* et qui se situe dans (Spec, v^*P). Comme un transfert aura lieu au niveau de v^*P , C n'a pas accès à ce complément, et c'est pourquoi on observe d'après Chomsky une contrainte d'îlot sujet.

Ce type d'approche syntaxique prédit que les relatives en *dont* seront plus faciles si *dont* correspond à un complément de verbe, et pour les (rares) cas où il correspondrait à un complément de nom, plus faciles si ce nom est objet ou sujet interne.

Selon ce type d'approche, l'extraction d'un complément de nom sujet n'est pas grammaticale et ne pourra se faire que de façon parasite, quand elle correspond par ailleurs à un complément de verbe :

- (4) Il traduit aussi les difficultés d'un syndicat dont la base _ se sent exclue _.
 (FTB - flmf7ahlep-77)

2.3 Les relatives en *dont* et la théorie de la localité de Gibson

Gibson (1998, 2000) propose une théorie de la localité (*Dependency Locality Theory*, ou DLT) inspirée des travaux sur l'ordre des mots de Hawkins (1994). D'après cette théorie, la dépense de ressource cognitive pour traiter une phrase dépend de deux facteurs : (i) le stockage des prédictions de la structure au fur et à mesure qu'on avance dans la phrase (mémoire sollicitée) et (ii) l'intégration de chaque

nouveau mot à l'intérieur de la structure au fur et à mesure qu'on avance dans la phrase (énergie sollicitée). Pour simplifier, nous ne prendrons en compte que ce dernier facteur dans la suite de notre explication.

L'intégration de chaque nouveau mot dépend de l'accessibilité des différents référents dans le discours. Gibson propose⁴ que le coût de l'intégration de chaque nouveau mot soit la somme de deux coûts distincts : d'une part le coût impliqué lorsque ce nouveau mot introduit un nouveau référent (1 unité d'énergie) et d'autre part le coût impliqué par la distance relative des dépendants qui doivent être reliés entre eux, ou des dépendants par rapport à la tête dont ils dépendent (1 unité d'énergie par nouveau référent se situant entre les deux dépendants).⁵

Appliquée aux relatives de l'anglais, cette théorie explique des temps de lecture plus longs quand le pronom relatif correspond à l'objet (5a) que s'il correspond au sujet (5b).

(5) a. The reporter [who the senator attacked _] admitted the error.

b. The reporter [who _ attacked the senator] admitted the error.

Dans une optique où le pronom relatif est extrait et doit être mis en relation avec un gap pour être interprété, *who* est adjacent au gap sujet en (5b) tandis qu'il est séparé du gap objet par deux nouveaux référents en (5a) : le sujet et le verbe. Le coût impliqué par la distance relative des dépendants au niveau du gap est donc de 0 unité d'énergie en (5b) et de 2 unités d'énergie en (5a).

Dans une analyse sans catégorie vide (Fodor et Sag 1996), le pronom relatif doit être mis en relation avec la tête, ici le verbe, dont il dépend. Le résultat est similaire : *who* est adjacent au verbe dont il est sujet en (b) et séparé du verbe dont il est objet en (a) par le sujet, donc par un référent.

Cette théorie a été à notre connaissance appliquée aux relatives en *que* du français (Pozniak et Hemforth 2015) mais pas à celles en *dont*. S'il s'agit de minimiser la distance entre *dont* et le gap (ou la tête dont il est un dépendant), elle prédit que les relatives en *dont* seront plus faciles si *dont* est complément d'un nom (ou pronom) sujet (6a) ou d'un verbe à sujet inversé (6b) (0 nouveau référent), un peu moins si *dont* est séparé du verbe dont il dépend par un sujet (6c) (1 nouveau référent), et plus difficile si *dont* dépend d'un objet dont il est séparé par le sujet et le verbe (6d) (2 nouveaux référents).

(6) a. des élèves dont certains sont turbulents

b. le livre dont parle le professeur

c. le livre dont le professeur parle

d. le poivron dont le garçon aime le goût

Enfin, il est important de noter que si chaque nouveau référent introduit dans le discours consomme des ressources, les pronoms sont considérés comme déjà présents dans le discours. Par conséquent, l'exemple 7 sera moins coûteux en unité d'énergie que l'exemple 6d, car le sujet de 7 n'introduit aucun nouveau référent dans le discours.

(7) le poivron dont tu aimes le goût

Dans cette théorie, l'extraction hors d'un complément nominal est grammaticale, que le nom soit sujet ou objet, et son acceptabilité sera graduelle.

Pour résumer, nous avons avec l'approche générativiste de Chomsky et l'approche cognitive de Gibson deux approches qui font appel à la notion de processing. Cependant, si d'après la première approche les contraintes de processing sont fortes et que leur violation entraîne l'agrammaticalité, l'approche cognitive voit dans ces contraintes une accumulation de coûts qui peuvent être réduits en adaptant le contexte. De plus, les deux modèles ont des prédictions opposées en ce qui concerne les extractions hors du SN sujet

en français. Le modèle minimaliste prévoit en effet que les extractions de vrais sujets (c'est-à-dire de sujets qui ne sont pas des arguments internes) sont bloquées et moins acceptables que les extractions hors du SN objets. La théorie de la localité prévoit au contraire pour le français une plus grande facilité à extraire d'un SN sujet que d'un SN objet car la distance sera plus courte entre le pronom relatif et le sujet de sa relative qu'entre le pronom relatif et l'objet de sa relative. Cette même théorie prévoit une réduction des coûts de processing sur l'extraction hors du SN objet si le sujet de la relative est prénominal et n'introduit pas de nouveau référent.

3 Etude sur corpus écrit

On dénombre dans le *French Tree Bank* (FTB) 384 relatives en *dont* avec verbe impliquant un gap.⁶ Nous décrivons dans un premier lieu les différents compléments auxquels sont associés les occurrences de *dont*. Puis, nous examinerons plus attentivement le cas où le *dont* est une extraction hors d'un SN, pour voir quelle fonction est privilégiée dans la langue écrite. Enfin, en nous concentrant sur les extractions hors du SN sujet, nous examinerons quels verbes donnent lieu à ces extractions.

3.1 Les relatives en *dont* du corpus écrit

Le Tableau 1 donne la fonction des gaps auxquels sont associés les *dont* que nous avons pu identifier dans le corpus.⁷ Nous examinons leur répartition en fonction du type de sujet de la relative : relative avec sujet préverbal et non-pronominal, relative avec sujet préverbal pronominal et relatives avec sujet postverbal.

		Cplt de Verbe ⁸	Cplt de Nom	Cplt d'adjectif ⁹	Ajout ¹⁰	Total
Sujet préverbal nominal		7	235	2	5	249
Sujet pronominal préverbal¹¹	clitique	32	40	3		
	autre	1	18	1		
	Total	33	58	4	4	99
Sujet inversé	nominal inversé	32	2	0		
	pronominal inversé	1	0	0		
	Total	33	2	0	1	36
Total		73 (19%)	295 (77%)	6	10	384

Tableau 1 : Répartition des occurrences de *dont* du FTB selon leur fonction

(8) Exemples de *dont* complément de nom du corpus :

a. L'arrivée du Peso ne signifie pas l'élimination automatique des Australs, **dont** le *remplacement* devrait se faire progressivement, en moins de six mois. (FTB - flmf7aa1ep-58)

b. Une mutation est en marche, qui ne s'arrêtera sans doute pas, mais **dont** on se garde bien de préciser, du côté des directeurs de ressources humaines, les *tenants et aboutissants*. (FTB - flmf7ag1exp-245)

c. Le premier n'avait pas les faveurs de France Télécom, **dont** Matra Communication est l'un des tout premiers *fournisseurs*. (FTB - flmf7ag1exp-449)

(9) Exemples de *dont* complément de verbe du corpus :

a. Le gouvernement n'avait ni écrit ni choisi cet accord **dont** nous avons *hérité*. (FTB - flmf7af2ep-602)

b. Les problèmes politiques, par exemple, **dont** *souffrait* l'écu, semblent maintenant résolus. (FTB - flmf7ad1co-11)

(10) Exemple de *dont* complément d'adjectif du corpus :

Ils ôtent à la politique monétaire ses références chiffrées **dont** les membres du directoire sont *friands*. (FTB - flmf7aa2ep-513)

On peut observer que dans une écrasante majorité des cas, *dont* correspond à un complément de nom (8). Cela confirme qu'en français il n'existe pas de contrainte sur l'îlot nominal (Godard 1988). Par ailleurs, quand *dont* correspond à un complément de verbe, il s'agit dans 90% des cas d'une relative à sujet pronominal (9a) ou inversé (9b). Ce fait semble confirmer la théorie de la localité de Gibson (1998, 2000) : lorsque le sujet introduit un nouveau référent dans le discours, la position postverbale permet de réduire d'une unité la distance entre *dont* et son gap. De même, les relatives à sujet pronominal favorisent l'extraction du complément du verbe. Ces différences ont été confirmées avec un modèle linéaire généralisé (régression logistique) qui montre significativement que les relatives en *dont* correspondent plus à des compléments de verbe lorsque leur sujet est pronominal que lorsqu'il est nominal préverbal ($z=7.7$, $p<.001$) et que lorsqu'il est inversé ($z=6.6$, $p<.001$). Encore une fois, cela corrobore la théorie de la localité car dans ces cas-là, le sujet n'implique pas de nouveau référent.

3.2 Les cas d'extraction hors du SN du corpus écrit

Le Tableau 2 recense les cas d'extraction hors du SN (sans les 3 cas de double gap) :

		Cplt de sujet	Cplt d'objet	Cplt d'attribut ¹²	Total
Sujet nominal préverbal		200	22	10	235
Sujet pronominal préverbal	Clitique	-	25	15	40
	Autre	17	1	0	18
	Total	17	26	15	58
Sujet nominal inversé ¹³		0	0	2	2
Total		217 (74%)	48 (16%)	27	295

Tableau 2 : Répartition des occurrences de *dont* compléments de nom du FTB selon la fonction du SN

Contrairement à ce que laisserait supposer l'approche générativiste (voir section 2.2), les extractions hors du SN sujet (8a) sont très fréquentes et représentent près des trois quarts des extractions hors du SN. Cela remet fortement en cause la contrainte d'îlot sujet et semble confirmer la théorie de la localité. Un autre fait va dans le sens de cette dernière : il n'existe aucune extraction hors du SN sujet où le sujet soit inversé (l'inversion du sujet rajouterait une distance supplémentaire non-nécessaire).

3.3 Types de verbes en cause dans les extractions

Nous l'avons vu, l'extraction hors du SN sujet est considérée par l'approche minimaliste comme facilitée lorsque ces sujets sont des objets profonds, ou des sujets « internes », c'est-à-dire lorsqu'ils sont sujets de passifs, de médiopassifs ou d'inaccusatifs (Chomsky 2008). Le Tableau 3 donne les verbes (hors verbe *être*) mis en cause dans les extractions hors du SN sujet et les extractions de compléments du verbe relevés dans le corpus.

		<i>Dont cplt de nom sujet</i>		<i>Dont cplt de verbe</i>	
Verbe		Nombre d'occurrences	Pourcentage	Nombre d'occurrences	Pourcentage
Sujet argument interne	passif	53	33%	3	4%
	inaccusatif ¹⁴	26	16%	10	14%
	mediopassif	5	3%	2	3%
	total	84	53%	15	21%
Vrai sujet	Inergatif	19	12%	43	59%
	transitif direct	56	35%	15	21%
	Total	75	47%	58	79%
Total		159		73	

Tableau 3 : Répartition des types de verbes mis en cause dans les extractions hors du SN sujet et extraction de complément du verbe dans les relatives en *dont* du FTB (hors verbe *être*)

On remarque qu'une forte proportion des extractions hors du SN sujet correspond en effet à des sujets de verbes passifs (11a), inaccusatifs (11b) et médiopassifs. En comparant les verbes à sujet externe aux verbes à sujet interne un modèle linéaire généralisé (régression logistique) montre que l'extraction du sujet est en effet significativement plus présent pour les verbes à sujet interne ($z=4.4$, $p<.001$). Le verbe *être* (58 occurrences) est fortement présent (11c).

(11) a. Cette vente, dont le montant est gardé secret, concrétise les rumeurs courant sur Biba depuis l'automne.
(FTB - flmf7al1ep-464)

b. Les entreprises dont l'effectif se situe entre 50 et 400 salariés acquittent leurs cotisations.
(FTB - flmf7ab2ep-853)

c. La fumée contenait en effet du oxyde cyanhydrique dont les effets sont immédiats.
(FTB - flmf3_08500_08999ep-8561)

Cependant, les verbes transitifs directs (12a) et inergatifs (12b) représentent 47% des cas relevés. Ce pourcentage n'est pas négligeable. L'extraction hors de sujets de verbes transitifs directs n'est pas exceptionnelle, et l'on ne peut par conséquent pas parler ici d'agrammaticalité.

(12) a. Les premiers étaient des coopératives dont les membres exploitaient sous forme privée des lopins de terre.
(FTB - flmf7ab1co-440)

- b. Une Allemagne encore plus puissante est en train de naître, dont le poids risque de déplacer vers l'Est le centre de gravité de l'Europe. (FTB - flmf7ah1ep-296)

Une hypothèse plus sémantique peut aussi expliquer la forte proportion des verbes non-agentifs (passifs, inaccusatifs, médiopassifs). Il est en effet intéressant de constater que seulement 8% des sujets de relatives en *dont* où *dont* est complément de nom sujet sont animés (18 occurrences, comme dans l'exemple 12a).¹⁵

On observe également à l'écrit qu'il existe une grande diversité de verbes qui donnent lieu à une extraction de leur complément. Les verbes inergatifs forment une majorité (13a) mais le verbe *parler* n'y est pas dominant (seulement 4 cas). On retrouve des verbes transitifs (13b) et des inaccusatifs (13c). Les passifs (13d) et les médiopassifs (13e) sont plus rares. Cela montre que le corpus ne se comporte pas comme les corpus oraux étudiés par C. Blanche-Benveniste (1995).

- (13) a. On avançait une série de remèdes contre les « maladies » dont souffre le Parlement.
(FTB - flmf3_03500_03999ep-3611)

- b. Un accord de principe - dont on sait encore peu de chose - devrait être conclu entre Avia et le groupe industriel allemand. (FTB - flmf7aa2ep-812)

- c. Ils ne sont pas légion, ceux dont on sort plus heureux qu'en entrant.
(FTB - flmf3_08500_08999ep-8591)

- d. De son côté, la COB enquête sur la date exacte de cette cession dont elle a été informée a posteriori. (FTB - flmf7ae1ep-244)

- e. Il est probable que [...] l'« accord à vivre » n'a pas été étranger au redressement des résultats commerciaux de Renault comme à l'amélioration de la qualité des véhicules dont se félicite la marque au losange. (FTB - flmf7am1ep-373)

4 Etude sur corpus oral

On dénombre dans le Corpus de Français Parlé Parisien des années 2000 (CFPP2000, Branca-Rosoff et al. 2012) 107 relatives en *dont* avec verbe impliquant un gap.¹⁶ Comme précédemment, nous allons présenter les chiffres issus du corpus en nous basant d'abord sur les différents compléments auxquels sont associés les occurrences de *dont*, puis sur la fonction des SN dont ils sont extraits et enfin sur les verbes qui donnent lieu à une extraction hors du SN sujet et à une extraction du complément du verbe.

4.1 Les relatives en *dont* du corpus oral

Nous présentons dans le Tableau 4 les relatives en *dont* selon leur fonction par rapport au syntagme dont ils sont extraits. Ce tableau est à mettre en relation avec le Tableau 1.

		Cplt de Verbe ⁶	Cplt de Nom	Cplt d'adjectif ⁷	Ajout ⁸	Total
Sujet nominal préverbal		0	16	0	1	17
Sujet pronominal préverbal⁹	Clitique	65	11	3	3	82
	Autre	0	3	0	2	5
	Total	65	14	3	5	87
Sujet nominal inversé¹⁷		2	0	0	1	3
Total		67 (63%)	30 (28%)	3	7	107

Tableau 4 : Répartition des occurrences de *dont* du CFPP2000 selon leur fonction

(14) Exemples de *dont* complément de nom du corpus :

a. puis j'ai un voisin un copain aussi qui euh dont c'est le **dont** *le père* s'occupe de cette association qui m'a promis de de de me les présenter (CFPP2000[13-03])

b. c'est un peu le quartier village **dont** tu connais *tous les commerçants* et à peu près *tous les habitants* (CFPP2000[13-02])

c. (locuteur 1 :) puis à côté de ça oui j'ai j'ai vu deux trois choses pas très (mm) pas très sympathiques + des agressions des trucs comme ça donc euh (locuteur 2 :) ah oui que vous n'av- (locuteur 1 :) voilà (locuteur 2 :) que vous avez pu **dont** vous avez été *témoin* (CFPP2000[20-01])

(15) Exemples de *dont* complément de verbe du corpus :

a. ça c'est un + c'est l' trajet **dont** j' m' *souviens* le mieux en fait (CFPP2000[12-04])

b. parce que la première chose **dont** me *parlent* les mes enquêtés c'est du plaisir d'aller dans les petits commerces (CFPP2000[07-06])

(16) Exemple de *dont* complément d'adjectif du corpus :

il y a d'autres exemples mais **dont** je me sens pas très très *proche* vraiment je ne vis pas la même chose (CFPP2000[18-01])

Nos résultats confirment les observations de Blanche-Benveniste (1980, 1990, 1995) : Les cas où *dont* correspond à un complément de verbe (15) représentent en effet une large majorité dans ce corpus oral. Cependant, bien que moins fréquents que dans le corpus écrit, les cas où *dont* correspond à un complément de nom (14) représentent 28 % des relatives en *dont* et restent par conséquent courants.

Même s'il n'existe que très peu de relatives introduites par *dont* contenant un sujet inversé, ce sont les extractions du complément du verbe que ces relatives privilégient (15b). L'extraction du complément du verbe semble quant à elle facilitée par un sujet pronominal : aucune des extractions du complément du verbe n'a lieu dans un cadre où le sujet est à la fois préverbal et nominal. L'effet du sujet inversé n'a pas pu être confirmé statistiquement de manière significative sur ce corpus seul, sans doute en raison du peu de données. Mais un modèle linéaire mixte type régression logistique appliqué aux données des sujets nominaux préverbaux et sujets pronominaux confirme une différence significative ($z=3.95$, $p<0,001$).¹⁸ Ces observations confirment avec encore plus de contraste ce que nous avons observé dans le corpus écrit. Elles semblent corroborer elles aussi la théorie de la localité.

4.2 Les cas d'extraction hors du SN

Le Tableau 5 récapitule les cas d'extraction hors du SN d'après la fonction du SN dans la phrase. Il est à mettre en relation avec le Tableau 2 en 3.2. Il n'y a pas de cas de sujet inversé.

		Cplt de sujet	Cplt d'objet	Cplt d'attribut ¹²	Disloquée à droite ¹⁹	Total
Sujet préverbal nominal		16	0	0	0	16
Sujet pronominal préverbal	Clitique	-	9	1	1	11
	Autre	3	0	0	0	3
	Total	3	9	1	1	14
Total		19 (63%)	9	1	1	30

Tableau 5 : Répartition des occurrences de *dont* compléments de nom du CFPP2000 selon la fonction du SN

Tout comme les résultats du corpus écrit, les fréquences du corpus oral vont à l'encontre des prédictions de l'approche générativiste (voir section 2.2). Les extractions hors du SN objet ne sont pas plus fréquentes que les extractions hors du SN sujet. Cela confirme la théorie de la localité. Cette dernière prédit également que les extractions hors du SN objet sont facilitées si le sujet est pronominal (et donc n'introduit pas de nouveau référent dans le discours), ce que confirment en effet nos données.

4.3 Types de verbes mis en cause dans les relatives en *dont*

Le Tableau 6 recense les verbes qui donnent lieu à une extraction hors d'un SN sujet, le verbe *être* excepté :

		<i>Dont cplt de nom sujet</i>		<i>Dont cplt de verbe</i>	
Verbe		Nombre d'occurrences	Pourcentage	Nombre d'occurrences	Pourcentage
Sujet argument interne	passif	0		0	
	Inaccusatif ¹²	3	25%	14	21%
	mediopassif	1	8%	1	1%
	total	4	33%	15	22%
Vrai sujet	inergatif	4	33%	48	72%
	transitif direct	4	33%	4	6%
	total	8	67%	52	78%
Total		12		67	

Tableau 6 : Répartition des types de verbes (hors verbe *être*) dans les extractions hors du SN sujet et extraction de complément du verbe dans les relatives en *dont* du CFPP2000

On constate que dans le corpus oral une majorité de verbes transitifs directs (17a) et inergatifs (67%) donne lieu à une extraction hors du SN sujet. Cela va à l'encontre des prédictions de Chomsky (2008). Les verbes passifs, médiopassifs et inaccusatifs (17b) sont rares, ce qui est certainement dû à la grande proportion de sujet animés (agentifs) dans le corpus. En effet, 68% des sujets de relative en *dont* où *dont* est complément de SN sujet sont animés (13 occurrences). On peut donc supposer que seule la différence dans les thématiques des deux corpus explique la variation de proportion de verbes transitifs. Les verbes inergatifs (17c), et entre autre le verbe *être* (17d), restent fréquents dans ce corpus.

(17) a. moi j'étais en classe avec des certainement des gens dont les parents avaient des grosses fortunes et d'autres qui n'en avaient pas (CFPP2000[07-05])

b. donc on il reste plus qu'une famille là-bas enfin d'ailleurs une euh un couple maintenant dont les enfants sont sont partis depuis quelques années (CFPP2000[13-02])

c. alors est ce qu'il y a des animaux auxquels euh + dont vous + dont la présences vous plaît dans Paris ou au contraire vous déplaît fortement (CFPP2000[05-01])

d. c'est vraiment un langage euh + qui est + bon dont les codes sont différents dont les codes sont différents (CFPP2000[07-05])

Ce corpus confirme également ce que notait Blanche-Benveniste sur une quasi-fixation de l'usage de *dont* à l'oral. En effet, le seul verbe *parler* représente plus de 71% des occurrences d'extraction de complément de verbe (47 occurrences, par exemple 18a). Le deuxième verbe le plus courant, *se souvenir*, représente quant à lui 15% des cas d'extraction de complément de verbe (10 occurrences, par exemple 18b). Ces deux verbes représentent donc à eux seuls 85% des cas. Il faut nuancer ce constat cependant, car nous avons pu voir avec les extractions hors du SN sujet que les relatives en *dont* sont relativement productives à l'oral.

(18) a. c'est pas celui dont on parle dans l'quartier les gens ne vont jamais au marché bio
(CFPP2000[07-05])

b. et ils emploient des mots que je dont je me souviens plus mais je suis capable de tenir uune conversation parfaite voilà
(CFPP2000[18-01])

5 Expérience sur les relatives en *dont*

Les deux études de corpus présentées précédemment ont donné plusieurs arguments en faveur de la théorie de la localité de Gibson et contre la théorie syntaxique des flots de la grammaire générative. Afin de compléter la confrontation entre ces deux approches, nous avons également mené une expérience de jugement d'acceptabilité qui nous a permis de confronter les prédictions opposées des deux modèles théoriques en ce qui concerne l'extraction hors du SN sujet et objet.

Les items testés lors de ces expériences sont au nombre de 30, déclinés d'après 3x2 conditions. La première variation porte sur la distance séparant *dont* du gap auquel il est associé : *dont* est associé à un complément de nom sujet (19a) ou à un complément de nom objet, et dans ce dernier cas le sujet est pronominal (19b) ou introduit un nouveau référent dans le discours (19c). Une version de contrôle sans extraction est faite pour chacune de ces distances (20).

(19) a. Ils entament une discussion dont l'importance surprend la préfecture.

(Extraction hors SN sujet)

b. Ils entament une discussion dont je ressens l'importance.

(Extraction hors SN objet + sujet pronominal)

c. Ils entament une discussion dont la préfecture ressent l'importance.

(Extraction hors SN objet + sujet nominal)

(20) a. Ils entament une discussion et son importance surprend la préfecture. (contrôle de 19a)

b. Ils entament une discussion et je ressens son importance. (contrôle de 19b)

c. Ils entament une discussion et la préfecture ressent son importance. (contrôle de 19c)

Le nom dont *dont* est le complément est pour les 30 items un nom de qualité (*importance, beauté, propriété...*). Tous les verbes sont des verbes psychologiques transitifs (impliquant un complément d'objet direct).

L'approche générativiste, telle que nous l'avons exposée en 2.2, prédit une meilleure acceptabilité des conditions d'extraction hors du SN objet que hors du SN sujet. En effet, nous avons affaire dans toutes les phrases testées à des verbes transitifs construisant un syntagme v*P. Par conséquent une phrase comme (19a) devrait être moins bien formée et plus difficile à traiter que (19b) et (19c). Des tests similaires ont

été effectués par Polinsky et al. (2013) pour l'anglais et le russe.²⁰ Cette même approche ne fait aucune prédiction sur l'influence de la forme du sujet dans les extractions hors du SN objet (et ne prévoit donc aucune différence notable d'acceptabilité ou de processing entre 19b et 19c).

A contrario, la théorie de la localité de Gibson (1998, 2000) prévoit qu'il est moins coûteux en unité d'énergie de relier un pronom relatif à son gap si aucun nouveau référent ne se trouve entre eux, ce qui est le cas lors d'une extraction hors d'un SN sujet situé directement après le pronom relatif. Les Tables 7, 8 et 9 ci-dessous calculent les coûts d'intégration totaux pour chacun des mots de la relative dans les 3 conditions d'extraction que nous avons exemplifiés dans les exemples 19. Comme l'illustrent ces tables, la théorie de la localité prévoit une hausse de la difficulté de processing sur le dernier mot de la relative entre l'extraction hors du sujet et l'extraction hors de l'objet, et dans ce dernier cas entre un contexte de sujet pronominal et un contexte de sujet nominal. La prédiction est donc celle d'une difficulté croissante $19a < 19b < 19c$.

	dont	l'	importance	surprend	la	Préfecture
Coût pour un nouveau référent	0	0	1	1	0	1
Coût impliqué par la distance	0	0	0	0	0	0
Coût total (coût de l'intégration de chaque nouveau mot)	0	0	1	1	0	1

Table 7 : Prédiction mot par mot de la théorie de la localité pour la relative de 19a

	dont	je	ressens	l'	importance
Coût pour un nouveau référent	0	0	1	0	1
Coût impliqué par la distance	0	0	0	0	1
Coût total (coût de l'intégration de chaque nouveau mot)	0	0	1	0	2

Table 8 : Prédiction mot par mot de la théorie de la localité pour la relative de 19b

	dont	la	préfecture	ressent	l'	importance
Coût pour un nouveau référent	0	0	1	1	0	1
Coût impliqué par la distance	0	0	0	0	0	2
Coût total (coût de l'intégration de chaque nouveau mot)	0	0	1	0	0	3

Table 9 : Prédiction mot par mot de la théorie de la localité pour la relative de 19c

L'expérience a été menée en ligne sur la plateforme Ibex²¹ entre janvier et février 2016. Les 57 participants dont les jugements ont pu être pris en compte étaient de langue maternelle française et avaient entre 19 et 73 ans. Leur participation était volontaire et non-rémunérée.

Les résultats de l'expérience ont été mis à disposition de la communauté scientifique et peuvent être consultés librement.²² Le graphique ci-dessous représente les jugements d'acceptabilité attribués par les

Graphique 1: Jugements d'acceptabilité

participants à l'ensemble des items selon les conditions. Un modèle linéaire généralisé (régression logistique) confirme une différence significative entre l'extraction hors du SN objet avec un sujet de la forme [det + N] (jugement d'acceptabilité moyen : 7,9) et l'extraction hors du SN objet avec un sujet pronominal (jugement d'acceptabilité moyen : 8,2). Une différence significative peut être confirmée de la même manière entre l'extraction hors du SN objet avec un sujet de la forme [det + N] et l'extraction hors du SN sujet (jugement d'acceptabilité moyen : 8,4). Quant aux phrases de contrôle, elles ne présentent pas de différence

significative entre elles.

Les résultats obtenus par cette expérience vont dans le même sens que les résultats des deux corpus pour ce qui concerne les relatives en *dont* : il existe en français une légère préférence pour les extractions hors du SN sujet par rapport aux extractions hors du SN objet. Pour ce qui concerne les relatives en *dont* impliquant une extraction hors du SN objet, une légère préférence pour les sujets pronominaux par rapport aux sujets qui introduisent un nouveau référent a été observée. Cela était prédit par la théorie de la localité. Cette dernière prévoyait en sus une différence entre les structures de 19a et 19b en faveur de la première. Cette différence n'a pas pu être observée de manière significative par les jugements d'acceptabilité. Des études supplémentaires (par exemple sur les temps de lecture) sont donc encore nécessaires.

6 Conclusion

Le présent article avait pour objectif d'observer le comportement des relatives en *dont* en français d'un point de vue empirique. Pour ce faire, nous avons étudié les relatives dans deux corpus, le *French Treebank* (corpus écrit) et le *Corpus de Français Parlé Parisien des années 2000* (corpus oral). Ces études nous ont permis de confirmer en partie les travaux de C. Blanche-Benveniste (1995) puisque si l'usage de *dont* en français écrit diffère effectivement de son usage en français oral, *dont* est loin d'être réservé à l'oral à quelques expressions figées. Nous avons pu aussi confirmer l'étude de Godard (1980) puisque les extractions hors du SN n'étaient pas marginales, pas même celles hors du SN sujet qui sont considérées par certaines théories linguistiques comme agrammaticales.

De ce point de vue, les corpus nous ont permis de vérifier que les extractions hors du SN sujet ne posent pas de problèmes spécifiques en français. Leur usage est même plutôt répandu, et inclut les extractions hors de « vrais sujets » de verbes transitifs, contrairement aux prédictions des hypothèses générativistes classiques (celles de la théorie des phases, entre autre). Il est intéressant par exemple de constater que les extractions hors du SN sujet sont plus fréquentes dans les corpus que les extractions hors du SN objet, alors que ces dernières devraient avoir la préférence d'après une approche minimaliste. Ce même fait est un argument en faveur de la théorie de la localité de Gibson, qui prévoit qu'en français l'extraction hors du SN sujet est non seulement bien formée mais moins coûteuse que celle hors du SN objet. D'autres observations sur les relatives en *dont* des deux corpus semblent corroborer cette théorie, comme par exemple les préférences marquées pour des sujets pronominaux ou inversés dans les relatives en *dont* dans lesquelles *dont* correspond à un complément du verbe.

Ces observations sur corpus demandaient à être confirmées par une étude expérimentale. Nous avons par conséquent mené une étude expérimentale de jugement d'acceptabilité afin de comparer les jugements de locuteurs natifs face à des relatives en *dont* correspondant à une extraction hors du SN sujet d'une part et à une extraction hors du SN objet d'autre part. Les résultats de cette expérience sont eux aussi en accord avec les prévisions de la théorie de la localité de Gibson, les jugements étant significativement supérieurs pour les extractions hors du SN sujet par rapport aux extractions hors du SN objet.

Références bibliographiques

- Abeillé, A. ; Clément, L. ; Toussanel, F. (2003). Building a treebank for French. Dans Abeillé, A. *Treebanks*. Dordrecht : Kluwer.
- Blanche-Benveniste, C. (1980). Divers types de relatives en français parlé. *TA Informations*, 2, 16-25.
- Blanche-Benveniste, C. (1990). Usages normatifs et non normatifs dans des relatives en français, espagnol, portugais. Bechert, J. ; Bernini, G. ; Buridant, C. *Towards a typology of European languages*, 317-333. Berlin : de Gruyter.
- Blanche-Benveniste, C. (1995). De la rareté de certains phénomènes syntaxiques en français parlé. *Journal of French Language Studies*, 5(1), 17-29.
- Branca-Rosoff, S. ; Fleury, S. ; Lefevre, F. ; Pires, M. (2012). *Discours sur la ville. Présentation du Corpus de Français Parlé Parisien des années 2000 (CFPP2000)*. URL : <http://cfpp2000.univ-paris3.fr/>.
- Chaves, R. (2013). An expectation-based account of subject islands and parasitism. *Journal of Linguistics*, 49(2), 285-327.
- Chomsky, N. (2008). On phases. Freidin, R.; Michaels, D.; Otero, C. P.; Zubizarreta, M. L. *Foundational Issues in Linguistic Theory : Essays in honor of Jean-Roger Vergnaud*, 133-165. Cambridge : MIT Press.
- Deane, P. (1991). Limits to Attention: A Cognitive Theory of Island Phenomena. *Cognitive Linguistics*, 2(1), 1-63.
- Erteschick-Shir, N. (2007). *Information Structure : The Syntax-Discourse Interface*. Oxford : Oxford University Press.
- Fodor, J. ; Sag, I. (1996). Une analyse sans catégorie vide des phénomènes d'extraction. *Langages*, 122, 8-31.
- Gapany, J. (2004). *Formes et fonctions des relatives en français*. Berne : Peter Lang.

- Gibson, T. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1–76.
- Gibson, T. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. Miyashita, Y. ; Marantz, A. ; O’Neil, W. *Image, Language, Brain*, p. 95-126. Cambridge, MA : MIT Press.
- Godard, D. (1988). *La syntaxe des relatives en français*. Paris : CNRS edition.
- Godard, D ; Sag, I. (1996). Quels compléments de nom peut-on extraire en français ? *Langages*, 122, 60-79.
- Godard, D. (1992). Extraction out of NP in French, *Natural Language & Linguistic Theory*, 2, 232-277.
- Hawkins, J. A. (1994). *A Performance Theory of Order and Constituency*. Cambridge, UK : Cambridge University Press.
- Laurens, F. ; Bîlbîie, G. (2010). Towards a nonelliptical analysis of verbless relative adjuncts. *Bucharest Working Papers in Linguistics (Bucharest Working Papers in Linguistics)*, 1, 51-67.
- Legendre, G. ; Sorace, A. (2003). Auxiliaires et intransitivité en français et dans les langues romanes. Godard, D. *Les langues romanes*, 185-233. Paris : CNRS Editions.
- Polinsky, M. ; Gallo, C. G. ; Graff, P. ; Kravtchenko, E. ; Milton Morgan, A. ; Sturgeon, A. (2013). Chapter 13. Subject islands are different. Sprouse, J ; Hornstein, N. *Experimental Syntax and Island Effects*. Cambridge, UK : Cambridge University Press.
- Pozniak, C. ; Hemforth, B. (2015). Processing subject and object relative clauses in French and Mandarin Chinese. Poster à CUNY, Los Angeles.
- Ross, J. R. (1968). *Constraints on variables in syntax*. Indiana University Linguistics Club.
- Sabel, J. (2002). A minimalist analysis of syntactic islands. *The Linguistic Review*, 19, 271-315.
- Sprouse, J. ; Caponigro, I., Greco, C., Cechetto C. (2015). Experimental syntax and the variation of island effects in English and Italian. *Natural language and linguistic theory* 34(1), 307-344.

¹ L’activité d’E. Winckel dans le cadre des îlots syntaxiques est liée au projet financé par la DFG *Long-distance Dependencies in French: Comparative Analyses (HPSG and the Minimalist Program)*, G. Mensching, Göttingen [ME 1252/14-1] / S. Müller, Berlin [MU 2822/9-1].

² Chaves 2013 note que des contraintes discursives (l’élément extrait doit être le thème de la proposition) et processives (un sujet complexe est plus difficile à traiter qu’un objet complexe) sont à l’œuvre pour que les extractions hors du sujet en anglais soient rares, mais pas impossibles :

(i) A question to which [the only acceptable answer] is “Not anymore.”
(COCA: 2010 May Esquire Magazine)

³ Une fois ce transfert effectué, le domaine de la tête de la phase (le domaine de v* dans le cas qui nous occupe) n’est plus accessible aux phases supérieures. (Phase-Impenetrability Condition, ou PIC)

⁴ Nous exposons ici la version simplifiée, qui est proposée par Gibson (2000).

⁵ Nous considérons la distance linéaire (en nombre de référents de discours). Une distance structurale (qui tient compte du nombre de nœuds syntaxiques) est aussi pertinente : si l’on adopte une structure syntagmatique assez plate, sans têtes fonctionnelles, il n’y a pas de différence entre les deux dans les cas qui nous concernent.

⁶ Les relatives partitives à *dont* sans verbe (ii) ont été étudiées pour ce même corpus par Laurens et Bîlbîie (2010). Nous les écartons dans la présente étude. Nous avons exclu aussi un cas de relative à *dont* sans sujet où le verbe est *voici* (iii) (*dont* correspond au complément de nom objet). Nous laissons de côté les relatives dites à *dont*-pronom (cf. Godard 1988, p.21) dans lesquelles le *dont* ne correspond pas à un gap (iv).

(ii) En Amérique latine, 23 journalistes ont trouvé la mort, dont 9 en Colombie et 7 au Pérou.
(FTB - flmf7aa1ep-328)

(iii) Au terme de ces travaux, le service de presse de la présidence de la république a diffusé un communiqué dont voici les principaux extraits. (FTB – flmf3_01000_01499ep-1168)

- (iv) Un bel effort, dont l'avenir dira s'il persuade les consommateurs, s'il suscite des imitateurs, le tout, naturellement, pour la plus grande gloire de ce monument : foie gras.
(FTB - flmf7ak1ep-272)

⁷ Les cas de double gap du corpus avec extraction du complément du sujet (3 occurrences exclues du tableau par soucis de lisibilité) ne sont pas tous des gaps parasites : v est un cas d'extraction du complément du sujet et du complément de l'ajout comparatif. Les exemples vi et 4 sont les deux autres cas que nous avons pu identifier.

- (v) Les particuliers ont ouvert en 1991 un plus grand nombre de comptes et de plans d'épargne-logement qu'en 1990, ce qui a évité une aggravation du déséquilibre du régime, **dont** les ressources nouvelles progressent moins vite que les dépenses, selon une note que vient de publier le ministère de l'économie.
(FTB – flmf7ah2ep-769)
- (vi) Les techniques utilisées par les pilotes de la guerre du Golfe pour se diriger en "voyant" sur leur visière de casque leurs objectifs peuvent être transposées à l'automobiliste de demain, **dont** le pare-brise deviendrait le tableau de bord en image.
(FTB - flmf7ab2ep-662)

⁸ Nous avons considéré comme complément de verbe les expressions [avoir/faire N de...] (*avoir besoin, faire preuve, faire partie*). Pour les expressions de la forme [être N de...] (*être l'objet, être question*), nous avons considéré *dont* comme correspondant à un complément de nom attribut.

⁹ Voir exemple 10 pour une illustration.

¹⁰ Les cas où *dont* correspond à l'extraction d'un syntagme prépositionnel (SP) ajout sont répertoriés dans cette catégorie. Il s'agit principalement de compléments circonstanciels de manière (*de cette manière...*) :

- (vii) Rien n'a été décidé au sujet de la manière dont elle se procurera les fonds.
(FTB – flmf7ad1co-39)
- (viii) de façon plus générale comment vous réagissez: à la mondialisation à la façon dont ça touche Paris et éventuellement la banlieue
(CFPP2000[11-02])

¹¹ Ont été catégorisés comme sujets pronominaux : les pronoms clitiques, les pronoms forts, les SN sans nom (*certain, beaucoup*) et les noms de quantité ou de proportion non-référentiels (*la moitié, une dizaine*). Il s'agit de proformes anaphoriques, déictiques ou partitives qui n'introduisent pas de nouveaux référents de discours (pas de pronoms indéfinis comme *quelqu'un* ou *quelque chose* en fonction sujet dans les relatives en *dont* des deux corpus).

¹² Il s'agit seulement d'attributs nominaux.

¹³ Nous n'avons pas trouvé dans le corpus de sujet pronominal inversé pour les *dont* correspondant à un complément de nom.

¹⁴ Notre distinction entre verbes inaccusatifs et verbes inergatifs suit la distinction établie par Legendre et Sorace (2003). Nous avons toutefois jugé utile de distinguer le verbe *être* des autres verbes inergatifs.

¹⁵ Ce pourcentage s'élève à 11% en prenant en compte les entreprises et autres entités inanimées susceptibles d'être personnifiées (5 occurrence supplémentaires).

¹⁶ Le nombre total d'occurrences de *dont* dans le CFPP2000 est de 134. Les relatives en *dont* sans verbe (ix) ont été exclues, ainsi que les relatives à *dont*-pronom (x) :

- (ix) et oui les lignes qui passent ici hein bien sûr + euh dont celles qui sont bloquées par la mono-activité textile
(CFPP2000[11-03])
- (x) (locuteur 1 :) avec des drogués (locuteur 2 :) oui voilà voilà (locuteur 1 :) dont on sait jamais comment ils vont
(CFPP2000[12-03])

De plus, plusieurs occurrences dont le sens n'était pas reconstituable (car inachevées ou trop incohérentes) ont été écartées.

¹⁷ Nous n'avons pas trouvé, dans le corpus oral, de sujet pronominal inversé.

¹⁸ Cette analyse a demandé l'ajout d'erreur aléatoire (smoothing) dans la condition sujet nominal préverbal où toutes les occurrences étaient des compléments de verbe.

¹⁹ On recense un cas de disloque à gauche qu'il nous semble intéressant de noter ici :

(xi) j'dirais qu'y a environ une quarantaine de familles + dont la moitié on on les connaissait à peine +
(CFPP2000[07-04])

²⁰ Leurs expériences diffèrent cependant de la nôtre sur un point qui nous semble crucial : les extractions testées par Polinsky et al. (2013) sont des questions directes et non des relatives. Or, les extractions mettant en cause des questions impliquent des contraintes liées à la structure informationnelle des énoncés (l'élément interrogatif doit être un focus), en plus des contraintes liées à la syntaxe des extractions à proprement parler. A partir du moment où le sujet est plus souvent un topique, on peut penser que l'extraction hors du sujet est plus difficile dans une interrogative directe que dans une relative, et que cela n'a rien à voir avec la syntaxe.

²¹ La plateforme Ibex Farm (<http://spellout.net/ibexfarm/>) permet d'héberger des expériences en ligne.

²² Les résultats sont disponibles à l'adresse suivante : <http://islands.hypotheses.org/69>