

***Dont* and *de qui* relatives in written French**

De qui ('of whom') relatives alternate with *dont* ('of which') relative clauses (RC) for animate antecedents. They have been considered to obey the subject island constraint (Tellier 1991, Sportiche 1998) contrary to *dont* RC (Godard 1988, Abeillé, Hemforth & Winckel 2016)(1). In this corpus study, we show that it is not the case and provide other factors for the alternation.

De qui also differs from *dont* since it can be used in non-finite RC, free RC and questions. Furthermore, it can be used with pied-piping (*auprès de qui* 'close to whom') whereas *dont* cannot (2). Using the literary corpus Frantext (<http://www.frantext.fr/>, here texts from 2000), we found 449 occurrences of *de qui*, among which 201 (45%) RC, 89 (20%) free RC and 128 (29%) interrogatives, plus some noise. In the same subcorpus, we found 13644 *dont* RC and annotated a random subset of 500 we found 143 RC with an animate antecedent, and compared them with *de qui* RC.

After taking out 8 RC without gap (or with 2 gaps), we have 193 *de qui* RC: 15% relativize a complement of Verb and 68%, a complement of Noun (Table 1). Among the relativizations of a complement of Noun, we found 56 extractions out of subject NP (42%)(3), 23% out of object NP, and 6% out of predicative NP. In our subset of *dont* RC, 20% relativize a complement of Verb and 76%, a complement of Noun (Table 2). Looking at *dont* extractions out of NP, we found 63% out of subject (4), 32% out of object and 4% out of predicative NP. We conclude that French does not rule out extraction out of subject: with both *dont* RC and *de qui* RC, extraction out of subject NP is attested and more frequent than extraction out of object NP. These extractions out of subjects do not only occur with unaccusative or passive verbs: 56% have transitive verbs with *de qui* and 32% with *dont*.

Concerning the alternation, we found a specialization of certain verbs for *de qui* or *dont* (*parler* 'talk' for *dont*; *tenir* 'hold' and *apprendre* 'learn' for *de qui*). Furthermore, 79% of *de qui* RC are appositives, but only 56% of our *dont* RC. We annotated each occurrence for subject position (postverbal/preverbal), type of verb (transitive or not), definiteness of the antecedent, number of the antecedent, and RC interpretation (appositive or restrictive). A generalized logistic regression analysis using these factors as predictors for the relativizer choice (with *dont* 0, *de qui* 1) show that both the RC interpretation and the verb type appeared significant: appositive RC are 2.67 times more likely to be introduced by *de qui* than *dont* ($p < .01$) and relatives with a transitive verb are 2.14 times more likely to be introduced by *de qui* than *dont* ($p < .05$) (table 3). We implemented a Hosmer and Lemeshow goodness of fit test which finds no significant difference between our regression model and the observed data ($p = .8951$).

We also looked at *de qui* questions in the same subcorpus. After taking out 42 verbless questions, we had 87 *de qui* questions, 41 direct and 46 subordinate (table 3). *De qui* can be extracted or in situ. We found more extractions of a complement of verb (66%, compared to 15% for the RCs) and fewer of a complement of noun (32%, 68% among RCs). Strikingly, we did not find any extraction out of subject NP (26% out of object NP). We thus conclude that the subject island constraint only holds for *wh*-questions. This is compatible with discourse-based approaches to islands (Erteschik-Shir 2007, Kuno 1982, Goldberg 2013) which analyze them as a result of discourse infelicity: since the extracted element is a focus in *wh*-questions, there is a mismatch with the discourse status of the subject noun (which is a topic); such a conflict does not arise with RC.

(1) a. un linguiste dont/ *de qui les parents ont déménagé à Chartres (Tellier 1991:90) ‘a linguist of whom the parents have moved to Chartres’

b. un linguiste dont/ de qui vous avez rencontré les parents (Tellier 1991:89) ‘a linguist of whom you have met the parents’

(2) la femme près de qui /*dont je dors (*Voyage*, Moustaki) ‘the woman next to who I sleep’

(3) [...] la jeune fille, de qui la silhouette [...] dépassait [...] les enfants assis en rond. (Pense à Demain, Garat, 2010) ‘the young woman, of who the silhouette passed the children seated in circle’

(4) [...] quelqu’un dont la fille a une copine de classe qui a une sclérose en plaques [...] (Le corps incertain, Gault, 2006) ‘someone of which the daughter has a friend of class who has a multiple sclerosis’

Table 1. *De qui* Relatives in Frantext (after 2000)

Cplt of verb	Cplt of noun	Cplt of adjective	Cplt of prep	Total
29 (15%)	132 (68%)	2 (1%)	30 (16%)	193

Table 2. *Dont* Relatives in Frantext after 2000 (sample with animate antecedent)

Cplt of verb	Cplt of noun	Cplt of adjective	Total
25 (20%)	96 (76%)	5 (4%)	126

Table 3. Results of the multiple regression model

Factors	Coefficient	Odds ratio	Standard error	z value	Pr(> z)
(Intercept)	-0.2668	1.305779	0.3909	-6.83	0.49483
RC interpretation (restrictive)	-0.9817	2.668990	0.3034	-3.235	0.00122
Subject position (postverbal)	0.1895	1.208645	0.6372	0.297	0.76617
Verb type (transitive)	0.7626	2.143843	0.2977	2.561	0.01043
Definiteness of the antecedent (definite)	0.4881	1.629218	0.3104	1.572	0.11585
Numerus of the antecedent (singular)	0.1192	1.126595	0.3410	0.350	0.72666

Table 4. *De qui* questions (direct and indirect).

Extr.: cplt of verb	Extr: cplt of noun	Extr.: Cplt of adjective	Extr.: Cplt of prep	in situ: cplt of verb	in situ: cplt of noun	in situ: cplt of adjective	Total
52	21	1	2	5	5	1	87

References

- Abeillé, Anne, Barbara Hemforth & Elodie Winkel. 2016. Les relatives en dont du français: études empiriques. In F. Neveu, G. Bergounioux, M.-H. Côté, J.-M. Fournier, L. Hriba & S. Prévost (eds.), *5e Congrès Mondial de Linguistique Française* (SHS Web of Conferences), vol. 27. http://www.shs-conferences.org/articles/shsconf/pdf/2016/05/shsconf_cmlf2016_14011.pdf.
- Erteschik-Shir, Nomi. 2007. *Information Structure: The Syntax-Discourse Interface*. Oxford University Press. Oxford.
- Godard, Danièle. 1988. *La syntaxe des relatives en français* (Science du langage). Paris: Eddu Centre national de la Recherche Scientifique.
- Goldberg, Adele. 2013. Backgrounded constituents cannot be extracted. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental Syntax and Island Effects*. Cambridge University Press.
- Kuno, Susumu. 1982. The structure of the Japanese Language. *Foundations of Language* 13(3). 421–447.
- Sportiche, Dominique. 1998. *Partitions and atoms of clause structure: Subjects, agreement, case, and clitics*. London and New York: Routledge.
- Tellier, Christine. 1991. *Licensing theory and French parasitic gaps* (Studies in Natural Language and Linguistic Theory). Vol. v. 26. Dordrecht, The Netherlands and Boston: Kluwer Academic Publishers.

