

French wh-words + ça: a corpus study

Anne Abeillé (Université de Paris), Anastasiia Smirnova (Université de Paris), Jong-Bok Kim (Kyung Hee University)

One intriguing property of the French question discourse particle *ça* ('that') is its immobility (Bayer & Obenauer 2011). Obenauer (1974) and Cheng and Rooryck 2001 (C&R) note that *wh-ça* must be in situ, behaving like the opposite of *wh-the hell* in English. As noted by C&R *wh-ça* questions cannot serve as an echo-question as in (2). However, note that it can also occur in ellipsis, as in (3).

- (1) Tu as vu qui ça (cet après-midi) ? ('You have met who-ça (this afternoon)?')
- (2) *Qui ça as-tu vu (cet après-midi) ? ('Who-ça have you met?')
- (3) — J'ai vu quelqu'un. ('I saw someone.') — Qui ça ? ('Who-ça?')

Lefeuvre (2018) conducted a corpus study (Frantext 1700-2010) and found 780 occurrences of *wh-ça ?* with a higher proportion of *comment*, *où*, *qui* and *pourquoi*. She suggests that (3) is headed by the *wh*-element with *ça* as a pronominal subject, like *Pourquoi ce bruit ? Pourquoi cela ?* ('Why this noise ? Why this ?'). The analysis of *ça* as a pronoun is not possible in (1). It is also dubious in (3) since a nominal subject is infelicitous (?? *Qui cet homme ?* 'Who this man ?'). We thus consider *ça* as a discourse particle in (1) and (3) as well. The ungrammaticality of (2) challenges movement-and-deletion approaches where the remnant *wh-ça* is supposed to first move to the focus position and the remaining clause undergoes ellipsis (Merchant 2002).

This paper focuses on contemporary French, using Frantext after 1960 (1001 texts, 61,2 M words), searching for a *wh*-word + *ça* and the same *wh*-word followed by ?. *Ça* cannot occur with a weak (**Que ça?* 'what ça') or anaphoric pronoun (**lequel ça* 'which-ça') nor with a determiner (**Quel ça homme?* 'Which-ça man?'). After eliminating cases where *ça* is used as a pronoun ('that'), we have 481 occurrences (table 1), with a higher proportion of *comment ça* and *où ça*, and a lower proportion of *quoi ça* and *pourquoi ça*, and no *combien ça*, compared to the bare *wh*-word in final position (in situ or sluice). We annotate *wh-ça* for extraction/ in situ/ ellipsis, taking only 104 occurrences for *comment* and 100 for *où*. We thus have 334 occurrences, among which 330 ellipsis (table 2). Contrary to previous literature, we found 2 examples of *wh-ça* extracted (4), with nonstandard *que* ('that') insertion (4b), as well as 2 in situ examples (5):

- (4) a. Où ça qu'ils vont ces deux copains-là ? ('Where-ça that they are going these two friends?') (Sabatier 1985)
b. Et pourquoi ça vous feriez pendre mon papa, monsieur le duc ? ('And why-ça you would hang my dad, Mr Duke?') (Queneau 1965)
- (5) a. — Mon papa il est mort. — Ton papa il est mort, quand ça ? ('— My dad he died. — Your dad he died, when-ça?') (Mauriac 1961)
b. T'avais appris à tirer où ça ? ('You had learned to shoot where-ça?') (Boudard 1982)

We suggest that the unacceptability of (2) may be due to a register clash (clitic inversion is formal while *wh-ça* belongs to an informal register). The rarity of in situ uses and extracted cases cast doubt on the derivation of *wh-ça* sluices from non-elliptical sources. As for ellipsis, we distinguish between sluicing and stripping. We mostly found matrix sluices, especially in dialog, but also in a few monologs (6b), as well as 4 embedded sluices, which are limited to indirect speech (6a).

- (6) a. [...] me demandant pourquoi ci pourquoi ça. ('wondering why-ci why-ça.') (Rochefort 1961)
b. Ma femme ? Quelle femme ? Où ça, ma femme ? ('Which wife ? Where-ça my wife?') (Pennac 1995)

Following Ginzburg & Sag 2000, we distinguish between reprise-sluices and direct-sluices. In the corpus, the latter have an indefinite antecedent (7), an implicit antecedent (8a) or no antecedent (*sprouting*) (8b).

- (7) Quelqu'un me l'a dit. — Qui ça ? ('Someone told me that. — Who-ça?') (Vargas 2008)
 (8) a. — Bien sûr, je copie. — Qui ça ? — Damascius, *Traité des Premiers Principes*. ('Of course, I copy. — Who-ça? — Damascius') (Roubaud 2002)
 b. — Mlle Dreyfus ne travaille plus ici. Elle nous a quittés. — Quand ça ? ('She left us. — When-ça?') (Gary 1974)

The interpretation requires contextual information: a non-elliptical question may be possible (*Elle vous a quittés quand ça ?* 'She left you when-ça?') in (8b) or not (**Qui ça vous l'a dit ?* 'Who-ça told you that?') in (7). Contrary to Yoshida et al (2015), wh-stripping is not limited to *why* and is very common with *comment ça* ('how ça'). Like reprise sluice, wh-stripping questions the form which has been used. The reprise fragment can be of any category: an adjective (10a), a noun, a preposition (10b), even a weak pronoun (11a), an imperative (11b) or a whole sentence:

- (10) a. — C'est curieux, fis-je en tournant mon regard vers la ville. — Pourquoi ça, curieux ? ('Why ça curious?') (Del Castillo 1981)
 b. — Avant, au moins, il y avait les mariages, les naissances, les vraies crises, les guerres, les révolutions, [...]. — Avant ? Quand ça, avant ? ('Before? When-ça, before?') (Sollers 1987)
 (11) a. — Ils ont arrêté Paul. [...] — Qui ça, ils ? ('Who-ça, they?') (Clavel 1968)
 b. — Laisse, dis-je. — Comment ça, laisse ? ('Leave (it), I said - How-ça, leave ?') (Pennac 1995)

There can be vehicle change (12a) or not (12b). Notice also the change of interrogative form in (12a) (a *si*-question to reprise a questioning declarative) and the indefinite change in (12c).

- (12) a. — Vous croyez ? dit le type, — Comment ça si je crois ? ('You believe (so)? the guy said - How-ça if I believe?') (Benoziglio 1974)
 b. B : Qui est là ? D : C'est moi, c'est moi. B : Qui ça, moi ? ('Who-ça, me?') (Ionesco 1963)
 c. — Elles ont emporté un certain nombre de caisses. — Quoi ! Comment ça, des caisses ? ('They took a certain number of boxes - What! How-ça, some boxes?') (Garat 2010)

In most cases, a non elliptical counterpart is not available. We conclude that matrix sluice is the main use of wh-ça and argue for a direct interpretation approach for the uses of wh-ca in a variety of elliptical constructions including sluicing as well as stripping (Ginzburg & Sag 2000, Kim and Abeillé 2019).

References

- Cheng L., & Rooryck J. (2001). Licensing wh-in-situ. *Syntax* 3 (1), 1-19.
 Ginzburg J., & Sag I. (2000). *Interrogative Investigations: The Form, Meaning and Use of English Interrogatives*. CSLI Publications.
 Kim, J-B, Abeillé A. (2019). Why-stripping in English: A corpus-based perspective. *Linguistics Research* 36 (3), 365-387.
 Lefeuve F. (2018) Les interrogatives partielles averbales : le cas de « proforme interrogative + ça / cela. in Béguelin M-J. et al. (eds) *L'Interrogative en français*, Peter Lang, 183-208.
 Merchant J. (2001). *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford U Press
 Yoshida M., Chinzuru N., Ortega-Santos I. (2015). The syntax of why-stripping. *NLLT* 33(1), 323-370.

Table 1. Wh-ça /wh-? in Frantext (1960-)

Wh-word	Quoi 'what'	Qui 'who'	Où 'where'	Quand 'when'	Pourquoi 'why'	Comment 'how'	Combien 'how many'	Total
Wh-word+ça	3	91	116	13	23	235	0	481
wh-word+?	4566	1379	939	570	2474	850	254	11032
Total	4569	1470	1055	583	2497	1085	254	11513

Table 2. Annotated uses of wh-ça in Frantext (1960-)

Wh-ça	Quoi ça 'what'	Qui ça 'who'	Où ça 'where'	Quand ça 'when'	Pourquoi ça 'why'	Comment ça 'how'	Total
Extracted	0	0	1	0	1	1	3
In situ	0	0	1	1	0	0	2
Sluice	3	65	86	11	21	38	223
Stripping	0	26	12	1	1	65	106
Total	3	91	100	13	23	104	334