

ARTICLE

‘Ces grandes dames, ça ne va pas souvent à l’église’: A corpus-based investigation of dislocation featuring co-referent clitic *ça*

Emily Linares

University of California, Berkeley

Corresponding author. Email: emily.linares@berkeley.edu

(Received 22 November 2017; revised 20 November 2018; accepted 12 March 2019)

ABSTRACT

Although previous research has documented dislocations in which a dislocated noun phrase (NP) occurs with the co-referent pronoun *ça* (e.g., Carlier, 1996; Jones, 1996), this study represents the first *corpus-driven* examination focused on this syntactic construction. In examining examples identified in both an oral and a written corpus, the investigation serves to nuance the association between dislocation and orality (McLaughlin, 2011). The research also interrogates the relationship between dislocations featuring co-referent clitic *ça* and genericity. The extent to which this construction more frequently signals a pejorative attitude towards certain groups of animate NPs than others is also explored. Finally, this study advances methodological considerations for research on rare syntactic forms.

1. INTRODUCTION

The examples that informed early research on dislocation were not consistently disclosed and were largely based on linguists’ intuition or that of their informants.¹ Blasco-Dulbecco (1999: 78) observes that, for this reason, the examples are somewhat unreliable. While there is evidence of data-driven approaches to dislocation (e.g., as adopted by McLaughlin, 2011), more corpus-based studies are needed to empirically document the numerous and varied forms that this construction may take (Blasco, 1997: 17; Blasco-Dulbecco, 1999: 10).

This article contributes to a gap in French syntactic research by undertaking the first corpus-based analysis of constructions in which a dislocated noun phrase (NP)

¹Blasco-Dulbecco (1999: 78) notes that, to her knowledge, Barnes (1985) was the only researcher to have used a numbered list of spoken examples.

occurs with the co-referent pronoun *ça*,² as in examples (1), (2) and (3), but may not display a straightforward, co-referential relationship,³ as in example (3).

- (1) **une femme, ça** vit dans le désordre, voyons! (Green, *Moïra*)
- (2) **la fin d'une vie, ça** ravigote (Beckett, *Malone meurt*)
- (3) **les petites vieilles ça** peut courir très vite (Cevin, 'Le Marchand,' in Carruthers, 2013)

As Beeching (1999: 82) observes, 'The beauty of *ça* lies in the flexibility of the syntactic connections which it can form'. In all three of the above examples, the clitic pronoun 'ne reprend pas le syntagme [NP] en tant qu'un individu; il réfère à la catégorie sémantique' (Blasco-Dulbecco, 1999: 122). In other words, dislocations that feature clitic *ça* are understood to impose a generic reading (Lambrecht, 1981; Riegel et al., 1994; Jones, 1996). Through the examination of dislocated constructions featuring clitic *ça* from two corpora (one oral, one written), this study questions the notion that all dislocated constructions featuring co-referent *ça* do in fact express an equal degree of genericity, with attention in particular to the presence of demonstrative determiners in dislocated NPs and the tense of verbs in the main clause. Additionally, the study explores the relationship between the expression of a pejorative attitude and the gender of referents in dislocated NPs encountered in the data. Finally, some observations concerning the register associated with this construction are presented.

2. LINGUISTIC BACKGROUND

Dislocation is generally defined as a syntactic phenomenon whereby an element is detached from the main clause but is simultaneously represented within the main clause by means of a clitic pronoun (Ayres-Bennett et al., 2001: 257), as illustrated in (4), where the dislocated NP and the co-referential clitic pronoun are bolded:

- (4) **L'auteur de l'Adonis, il** ne peut être qu'un esprit singulièrement attentif (Valéry) (Blasco-Dulbecco, 1999: 9; cited in McLaughlin, 2011: 2)

Research on this phenomenon is scattered throughout the literature, appearing under such discussions as word order (Ayres-Bennett et al., 2001), emphasis (Riegel et al., 1994), and the left clause periphery (Rowlett, 2007). In addition to being discussed under a range of topics, dislocation research has witnessed what Gadet termed a 'foisonnement terminologique' (1991: 119); a variety of names (e.g., dislocation, detachment, segmentation) have been used to describe 'dislocation', and the semantic overtones of this terminology has been a source of contestation. Blanche-

²I have chosen to exclude examples that feature *ce* combined with *être* (i.e., NP + *c'est*) due to the grammaticalization of *c'est* and 'the nearly obligatory use of the NP *c'est* LD in spoken language' (Barnes, 1985: 51).

³Researchers have remarked the power of *ce* and *ça* to produce a 'décalage' or 'décrochement référentiel' when they function as co-referent pronouns (Burston and Burston-Monville, 1981; Cadiot, 1988; Blasco-Dulbecco, 1999).

Benveniste (2006: 483) asserts that ‘dislocation’ is a poor label that ‘hints about movements and pathological results’. Some scholars have also criticized the use of directional adjectives to specify the position of the dislocated NP relative to the main clause (i.e., ‘left’ and ‘right’ dislocation) due to their association with the written medium (ibid.; Blasco-Dulbecco, 1999: 11). In order to avoid greater terminological confusion (as attested in discussions of borrowing, replication, code copying, transference, etc. in language contact research), I retain the term ‘dislocation’ due to its recognition in the field (Blasco-Dulbecco, 1999: 43; McLaughlin, 2011: 210), and I refer to instances of ‘left’ and ‘right’ dislocation for the same reason.

The classification of examples necessitates a consideration of the distinction between dislocations featuring double marking as opposed to hanging topics. In the most recent book monograph on French dislocation—and 40 years after Cinque (1977: 43) asserted that ‘in many cases, it is not simple to decide what construction we are observing’—De Cat (2007: 108) notes that ‘[t]here is however considerable disagreement in the literature as to the precise characteristics of each construction’. She asserts that the only dependable method to identify hanging topic left dislocations (HTLD) has been the absence of dependency markers (also see Cinque, 1977; Larsson, 1979; Delais-Roussarie, Doetjes and Sleeman, 2004): ‘The clearest case of HTLD would therefore be one in which the dislocated element could be introduced by a preposition’ (De Cat, 2007: 136) but is not. Applying this criterion, (a) would constitute a case of HTLD while (b) would represent left dislocation with double marking.

- (a) Son photographe, elle lui en a jamais parlé.
 - (b) À son photographe, elle lui en a jamais parlé.
- (examples reproduced from De Cat, p. 136)

This diagnostic was applied to the data for the present study in order to identify apparent examples of HTLD and exclude them from the analysis.

McLaughlin (2011) highlights another definitional problem that linguists studying dislocation have faced: the relationship between the dislocated element and the clitic pronoun with which it co-occurs. The type of dislocation under study features *clitic doubling*, or *double marquage*, whereby two grammatical categories occupy a single syntactic position (Blasco-Dulbecco and Caddeo, 2001): the clitic pronoun and NP exist in a relationship of co-reference (Blanche-Benveniste et al., 1987: 83; Blasco-Dulbecco, 1999). In order to underscore this relationship, the pronoun is referred to as ‘co-referent clitic *ça*’. In cases of double marking, Blasco and Cappeau (1993) assert that the clitic pronoun is first linked to the verbal construction, noting that a secondary, co-referential relationship exists between the clitic and the dislocated lexical item. Blasco-Dulbecco (1997: 2; 1999: 99) contends that strict co-referential relationships in terms of gender and number between a dislocated element and a clitic pronoun are not always respected, as evidenced by examples such as (5), in which agreement is not realized:

- (5) **Les gosses, ça** se lève tôt le matin (Queneau, *Zazie dans le metro*: 26, cited in Carlier, 1996: 134)

De Cat (2005; 2007) argues that French subject clitics do not function as agreement markers, concluding that the minimal similarities in behaviour between subject clitics and affixes or morphemes should be viewed as an ‘accident’. Examples such as (5) do, however, prompt consideration of the particular type of reading that they impose.

Jones (1996) notes that ‘*ça* (and more rarely *cela*) can be used, even though gender is evident, when a generic interpretation is intended (i.e., when the statement applies to an entire class of entities)’ (p. 260). The examples that he provides illustrate the generic effect produced by *ça*, namely:

(6) **Le café**, *ça* empêche de dormir [generic reading, *ça*]

(7) **Le café** que j’ai bu hier soir, *il* m’a empêché de dormir [specific reading, *il*]

Jones comments that when this construction is extended to human beings, as in ‘**les étudiants, ça ne travaille pas**’ (his example), it frequently communicates a ‘pejorative attitude’ towards the dislocated NP and may be regarded as ‘patronizing or sarcastic even when the comment is complimentary’ (p. 260). While such statements may be understood as advancing a criticism about a semantic category (e.g., *les étudiants* in Jones’s example), represented by *ça* in the main clause, Blasco-Dulbecco (1999: 122–123) highlights the possibility for ambiguous readings in terms of genericity when a *specific* group constitutes the dislocated NP, as in example (8).

(8) mais je crois que **ces envaisseurs ça** a jamais existé (Jeanjean, 1985: 242–243).

In this example, the NP is marked by a demonstrative determiner, *ces*. The specificity expressed by the demonstrative coupled with the construction’s frequent association with generic readings call into question the pragmatic function (Ashby, 1988; Barnes, 1985) of this dislocation as an expression of genericity.

In addition to asserting that dislocation provides a means to provoke a generic reading, De Cat (2007: 83) notes that dislocated NPs that express the subject of a generic sentence are ‘obligatorily dislocated in spoken French’. De Cat relates the avoidance of dislocation to the written medium, to which she attributes greater formality. In response to the examination of dislocation as an oral structure, as posed by Gadet (1991: 110, i.e., ‘Le détachement: une structure de l’oral?’), McLaughlin (2011) contends that orality has been oversimplified. Her research nuances the link between dislocation and orality by proposing three sub-types of orality: ‘primary’ orality as found in the oral code, ‘representative’ orality as manifested in journalistic prose and ‘constructed’ orality as imagined in fiction. Secondly, McLaughlin challenges the notion that the exclusive stylistic function of dislocation in French fiction is to express ‘orality’.

Blasco (1997: 17) observes that types of dislocation are multiple and diverse in French: ‘il n’y a donc pas une dislocation en français mais plusieurs cas fort différents’. Although occasionally touched upon in discussions of non-prototypical cases of dislocation (Lambrecht, 1981; Blasco-Dulbecco, 1999; Jones,

1996), this article offers an initial corpus-based examination of a ‘rare syntactic form’ (Carruthers, 1999): dislocation with co-referent clitic *ça*. In drawing on both an oral and a written corpus, the investigation builds on McLaughlin’s (2011) refinement of ‘orality’ by illustrating the value of considering written data sources for the study of the dislocation under investigation as well as other syntactic constructions.

3. CORPORA AND METHODOLOGY

The *French Oral Narrative Corpus* (Carruthers, 2013), a collection of audio recordings and accompanying transcriptions of 87 stories told to an audience by 18 storytellers (almost 1000 minutes of ‘spontaneous’ speech) was initially selected as the corpus for the present investigation. This preexisting corpus is conveniently tagged for instances of detachment where ‘the replacing pronoun is not totally straightforwardly co-referential with the dislocated element, as in **les oignons ça fait pleurer**’ (Carruthers, 2013). In order to identify such cases of dislocation within the corpus, a search algorithm was created⁴ and an XML editor was used to search for these instances in each of the performed stories that comprise the corpus. Most of the results yielded by this search did not constitute examples of the syntactic construction under investigation and were eliminated accordingly (e.g., **ces cris-là** (ils) ne comprenaient pas **ce** que c’était). In total, the search produced only eight examples, which is unsurprising due to the focus of the present study on a rare syntactic form.

Carruthers (1999) suggests that researchers investigating rare syntactic constructions would benefit from a combination of two techniques, namely, a Labov-style interview, in which the researcher attempts to elicit the form being investigated, and oral and written questionnaires (e.g., asking whether participants use the construction in question, what sort of language it represents in their view, and with whom they associate it). These techniques were not feasible under the constraints of the present investigation. Consequently, it was deemed necessary to conduct a search in a broader corpus in order to generate a larger data sample for analysis, and FRANTEXT was selected.⁵ Given the association between avoidance of dislocation and the written medium (De Cat, 2007), the selection of FRANTEXT could at first be regarded as an unconventional choice since this written corpus would presumably contain fewer instances of dislocation than an oral corpus. Nevertheless, with the premise that orality has been oversimplified, this investigation seeks to further contribute to its refinement through an exploration of the third sub-category put forth by McLaughlin (2011): ‘constructed’ orality imagined in fictional written texts. The presence of written attestations of the dislocation under investigation in the

⁴i.e., XPATH: `//*[u//seg[contains(@ana,'#det')]]`

⁵Other French oral corpora are freely available online (e.g., C-Oral-Rom or Corpus de Français Parlé Parisien), and the consultation of a number of oral corpora of spontaneous speech could have offered a data sample roughly comparable in size to FRANTEXT. While a larger oral data sample could have enabled a comparison between ‘natural’ orality and ‘constructed’ orality in literary texts, the difficulty of formulating queries to identify the structure under investigation (in the absence of punctuation) justified the focus on a single oral corpus for the purposes of the present investigation.

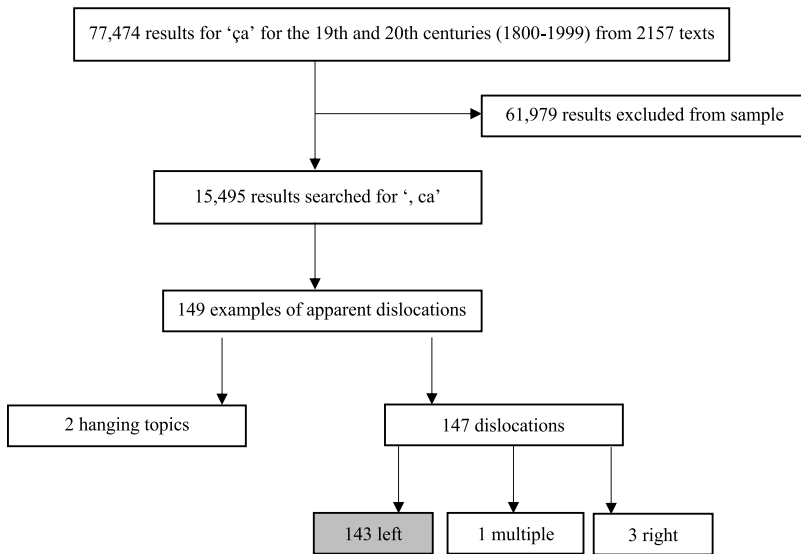


Figure 1. Delimitation of the FRANTEXT Data Set

FRANTEXT corpus highlights the importance of broadening our conception of orality in our research practices, as McLaughlin suggests.

FRANTEXT was selected as an appropriate corpus due to its size and the wide range of texts that it contains. So as to limit the scope of the project and allow for greater depth in analysis, the search was restricted to the nineteenth and twentieth centuries.⁶ A preliminary search for all instances of *ça* in the entire period under examination, namely, 1800–1999, yielded 77,474 occurrences (Figure 1). These attestations were stratified into four half-century periods (i.e., 1800–1849, 1850–1899, 1900–1949, 1950–1999) according to search parameters available in the FRANTEXT database. The first fifth of the results for ‘ça’ were reviewed for each 50-year period. However, this approach did not yield a significant number of NP (,) *ça* [verb] or *ça* [verb] (,) NP dislocation examples. Consequently, it was deemed appropriate to instead search for ‘, ça’: (NP) ‘, ça’ [verb]. Although this could cause certain examples without a comma (NP \emptyset *ça* [verb]) as well as instances of right dislocation (*ça* [verb] (,) NP) to be excluded from the present study, this represented the only feasible way to search and analyze thousands of instances of *ça* with the aim of encountering the dislocation variant under examination.⁷ Even so, examples of right dislocation could still be encountered

⁶The FRANTEXT corpus, which is comprised of novels, poetry, theatre, journalism, essays, correspondence and treatises (but excludes translations), contains a total of 2,157 texts for these two centuries combined.

⁷As McLaughlin (2011) observes, generative analyses commonly make a distinction between subject doubling and dislocation, predominately on the basis of prosody, whereby ‘*Mon chat il dort tout le temps*’ represents subject doubling, in contrast with ‘*Mon chat, il dort tout le temps*’, classified as a dislocation (Rowlett, 2007: 137; also see McLaughlin, 2011: 3). Following McLaughlin (2011), in this

Table 1. Left Dislocations by Half-Century (FRANTEXT)

Corpus	1800–1849	1850–1899	1900–1949	1950–1999
FRANTEXT	24	49	52	18

due to the way in which some dislocations were embedded in the surrounding narrative discourse, as in example (9).

- (9) Voyez-vous, me dit Joseph, **ça** mange beaucoup, **ces sortes d'animaux** (Dumas, *Le Capitaine Pamphile*)

The first fifth of the results for each 50-year period (15,495 occurrences in total for the four half-centuries) were examined for homogenous coverage of both centuries, and all instances of dislocation in which a dislocated NP co-occurred with ‘, ça’ were extracted for analysis (Table 1).

The data sample obtained from the FRANTEXT search initially yielded 149 examples of apparent dislocations marked by clitic *ça*. Two of these results were found to represent apparent cases of hanging topics according to the aforementioned diagnostic.⁸ Unsurprisingly, almost all of the examples of dislocation (143 out of 147 examples, or 97 per cent) represent instances of left dislocation, which is likely a consequence of the adopted methodology. Despite the expected predominance of left dislocations, one instance of repeated dislocation (McLaughlin, 2011) was encountered:

- (10) une femme qui est mariée... **une femme qui est mariée...** **ça** peut tout entendre, **ça** peut tout voir, **une femme qui est mariée...** (Goncourt, *Sœur Philomène*)

Additionally, three instances of right dislocation were obtained from the search, as in (11).⁹

- (11) Eh bien, **ça** parle, **un avocat!** (Sand, *L'Homme de neige*).

Since previous syntactic research has demonstrated the purely surface-level homogeneity of left and right dislocations (Lambrecht, 1981, 2001; Fradin, 1988), the present study focused exclusively on the 143 examples of left dislocation.

study involving written data, I conflate these phenomena but expect to find more examples involving a comma as a consequence of my methodology.

⁸The following were classified as hanging topics: ‘Tous ces gens en nage, ça ne doit pas sentir très bon...’ (Colette, *Claudine à l'école*) in which the situation (i.e., [avec] *tous ces gens en nage*) is presumably represented by *ça* rather than the swimmers themselves, and ‘Des ruines dans le feuillage, ça fait très bien’ in which the *inclusion* of ruins in the landscape contributes to a visual effect (i.e., [avec] *des ruines dans le feuillage*).

⁹One example from the Carruthers (2013) corpus (example 18) featured right dislocation.

Table 2. Grammatical Number of Dislocated NPs (Carruthers)

Corpus	Quantifier	Indefinite	Partitive	Definite
Carruthers (2013)	∅	3	1	3

Table 3. Demonstrative and Possessive Adjectives Modifying Dislocated NPs (Carruthers)

Corpus	Demonstrative	Possessive
Carruthers (2013)	∅	1

4. RESULTS AND DISCUSSION

The examples from the oral Carruthers (2013) corpus and written FRANTEXT corpus are first presented separately. After analysing the data, I reinforce the value of employing oral and written corpora in syntactic research and encourage the consideration of written data in future investigations of rare syntactic forms.

4.1 Carruthers (2013)

Carruthers' (2013) corpus yielded only eight examples, example (3) and examples (12)–(18), consisting of seven left dislocations and one right dislocation (example 18). Table 2 summarizes the grammatical number (i.e., quantifier, or indefinite, partitive, definite article) of the dislocated NPs, and Table 3 summarizes the use of demonstrative or possessive determiners in the eight examples. The latter two types of determiners were categorized separately since they are not associated with expressions of genericity.

- (3) **les petites vieilles ça** peut courir très vite (Cevin, 'Le Marchand'¹⁰)
- (12) **une femme en colère ça** marche vite (Kiss, 'Les Baies d'amour')
- (13) **une vache ça** coûte cher quand même (Kiss, 'Le Paysan et la paysanne')
- (14) **une femme aussi belle ça** n'existe pas (Kiss, 'Titete et Ticorps')
- (15) ah **des fleurs roses des fleurs blanches ça** embaumait! (Kiss, 'Les Baies d'amour')
- (16) **les oignons ça** faisait autant pleurer (Kiss, 'Le Paysan et la paysanne')
- (17) **ses larmes ça** donnait du courage aux gens qui en manquaient (Cevin, 'Grand-Mère Mensonge - 1')
- (18) **ça** sent longtemps dans les casseroles **l'odeur de brûlé** (Kiss, 'Le Paysan et la paysanne')

Examples (15)–(17) deviate from expectations for generics in two respects. Generic statements are habitually expressed in the present tense since they communicate generalizations—truths from the perspective of the speaker—about

¹⁰(Storyteller, 'Story title')

a particular class of objects or animate beings. The examples in question involve incongruous use of the past tense with imperfect aspect. Further, the use of a possessive adjective (example 17) is not typically associated with generic readings due to its expression of specificity. These observations motivated attention to grammatical determiners as well as tense in investigating the expression of genericity in the larger FRANTEXT corpus. The gender of human referents is also noteworthy: all of the human NPs are women in these examples. This prompted further exploration of whether men also featured as NPs in the FRANTEXT data and, if so, whether a pejorative attitude was expressed towards them with similar frequency. That all examples from the Carruthers corpus were produced by only two storytellers also reinforced the necessity of consulting a larger corpus to eschew generalizations on the nature of dislocations with co-referent *ça* based on the speech of only two informants.

4.2 FRANTEXT

The analysis of the 143 left dislocations identified in the FRANTEXT data offered insights for discussions of genericity (Section 4.2.1), the differential expression of a pejorative attitude according to gender (Section 4.2.2) and the use of a particular register in the construction under study (Section 4.2.3).

4.2.1 Genericity

Generics are generally defined as statements that refer to an entire class or subclass of objects or individuals (Galmiche, 1983: 30; Riegel et al., 1994: 284; Jones, 1996: 260). Generic readings can be achieved with or without recourse to dislocation. For example, both example (19) and example (20) impose generic interpretations, with the dislocated element referring to a semantic category (in this case, *les maris*).

(19) **Les maris** durent toujours (*my example*)

(20) **Les maris, ça** dure toujours (Scribe, *Le Mariage de raison*)

Galmiche (1983: 30–31) defines ‘les articles génériques’ (i.e., articles used to introduce generics) as *un(e)*, *le/la*, *les*, and *des*. He emphasizes that only phrases presenting ‘un syntagme générique’ can elicit a generic reading (p. 30). The presence of an ‘élément restrictif’ limits the scope of the dislocated NP and results in an ambiguous, ‘ungrammatical’ statement (p. 31), as in:

(21) ***Les gosses de mon voisin, ça** se lève tôt le matin¹¹ (cited in Carlier, 1996: 134).

Non-generic determiners that constitute an ‘élément restrictif’ (i.e., demonstrative and possessive adjectives, e.g., *ses gosses*) violate properties of generic statements since they restrict the scope of the dislocated element that they modify and denote specificity (Blasco-Dulbecco, 1999: 123). In order to test

¹¹This example is contrasted with a ‘grammatical’ utterance taken from Queneau’s *Zazie dans le métro*: ‘Les gosses [ø], ça se lève tôt le matin’.

Table 4. Determiners for Dislocated NPs (FRANTEXT)

Corpus	Quantifier	Indefinite	Partitive	Definite	Demonstrative	Possessive
FRANTEXT	3	51	7	56	24	2

the association in the previous literature between dislocation featuring co-referent clitic *ça* and genericity, the data were analysed with regard to the types of determiners modifying dislocated NPs (Table 4).

The number of dislocated NPs introduced by an indefinite article (51 examples, or 36% of 143) is similar to the number introduced by a definite article (56 examples, or 39% of 143). Most striking is the high presence of demonstratives (24 examples, or 17% of 143), as illustrated by examples (22) and (23), due to their restrictive quality (Blasco-Dulbecco, 1999: 123).

- (22) **Ces gosses**, **ça** avale des gobelets d'eau froide, c'est détestable pour la santé (Colette, *Claudine à l'école*)
- (23) **ces gens-là**,¹² **ça** n'a ni famille ni patrie, **ça** n'a de commerce qu'avec le démon, et **ça** pille, et **ça** tue pour rien, pour le plaisir de faire le mal... (Zola, *Vérité*)¹³

In such instances, the construction represents a commentary on a more restricted group of individuals—not a 'class' of individuals in a broad sense—and consequently prompts us to question the notion that *ça* always functions as a marker of genericity in this type of dislocation.

Although researchers agree that references to a particular object or individual (or group of objects or individuals) do not constitute generics (Galmiche, 1983; Riegel et al., 1994; Carlier, 1996; Jones, 1996; Blasco-Dulbecco, 1999), example (24) convincingly demonstrates that *ça* may be used to refer to both a specific, restricted group *and* a particular individual.

- (24) **ces hommes...** **celui-là surtout**, **ça** ne se doute de rien, **ça** ne pense à rien (Scribe, *Le Mariage de raison*)

The singling out of a particular individual (i.e., *celui-là surtout*) from an already restricted group (i.e., *ces hommes*) increases specificity. 'Ces hommes' could only support a generic interpretation if the NP were read as '*ce genre d'hommes*'. As previously mentioned, De Cat (2007: 83) asserts that dislocated NPs expressing the subject of a generic sentence are 'obligatorily dislocated in spoken French'. The presence of *ça* in this line, taken from Scribe's play, could therefore be understood as a feature of the oral code that is represented graphically (i.e., ROG, Representation of the Oral in the Graphic, see McLaughlin, 2011: 7). The

¹²The restrictive scope of this NP is reinforced by the use of a deictic, *là*.

¹³An examination of the text surrounding example (23) reveals that 'ces gens-là' is a pejorative reference to a group of Jews. The statement can be interpreted as a generic to the extent that the individuals in question are understood as representative of a larger class of people.

assertion made in (24) that the men being referred to, and that one man in particular, *ne se doute de rien, ne pense à rien* would seem to support the notion that detachment featuring co-referent *ça* may convey a pejorative attitude and may represent spoken French but need not express a consistent degree of genericity.

Carlier (1996: 142) asserts that *ça* cannot refer to a particular individual, however an individual is singled out by name in example (25).

- (25) **le riche banquier Leuwen**, avec Mademoiselle Des Brins, **ça** ne triche pas... (Stendhal, *Lucien Luwen*)¹⁴

A generic reading could only be produced if ‘le riche banquier’ were used to refer to a class of bankers (e.g., le riche banquier, *ça* ne triche pas), but this is not the case. For this reason, the use of detachment with *ça* can be interpreted as characterizing spoken French.

A different ‘generic particle’ introduces a member of a particular family in example (26):

- (26) mais **un Mauprat**, vois-tu, **ça** sait lire et écrire, et **ça** n’en est que plus méchant. (Sand, *Mauprat*)

This example further attests to the extent to which the scope of a NP was found to vary within the data. Here, a particular family represents a ‘class’ of individuals, and generalizations are made about members’ characteristics and abilities. As with all generics, the judgement made about ‘un Mauprat’ is expressed as an unequivocal truth (Galmiche, 1983: 29).

Galmiche (1983: 32) points out that, paradoxically, while generic statements express ‘truths’, the dislocated elements cannot be introduced by quantifiers expressing ‘universality’ (i.e., tout(e)(s), n’importe quel(le), chaque). In other words, although generic statements can be made about *a* Mauprat, they cannot be made about *all* Mauprats. He provides example (27) as an illustration of this point.

- (27) ***Tous les canaris**, **ça** chante bien. (Galmiche’s example)

Nevertheless, one example expressing universality was identified in the written corpus, namely (28).

- (28) **tous les journaux**, **ça** ne sert à rien (Murger, *Scènes de la vie de Bohème*)¹⁵

¹⁴With regard to critical reference, the context surrounding the example was examined, and neither a pejorative attitude nor sarcasm appears to be expressed towards *le riche banquier*. Rachel Shuh, who has published on *Lucien Luwen* (Shuh, 1999), confirmed an absence of sarcasm or criticism in the statement made about the banker: ‘The Kortis character is supposed to be sort of a “naïf” with a healthy respect for rich bankers and opera singers’ (personal correspondence).

¹⁵Although (28) does not represent an example of a hanging topic, one can be found in the surrounding context: The speaker criticizes all newspapers, concluding: ‘tous les journaux, c’est des menteries’. This statement features a hanging topic (i.e., in the absence of the preposition *dans*), however it was not identified in the creation of the corpus since it does not feature clitic *ça* but *c[e]’est* (see note 2).

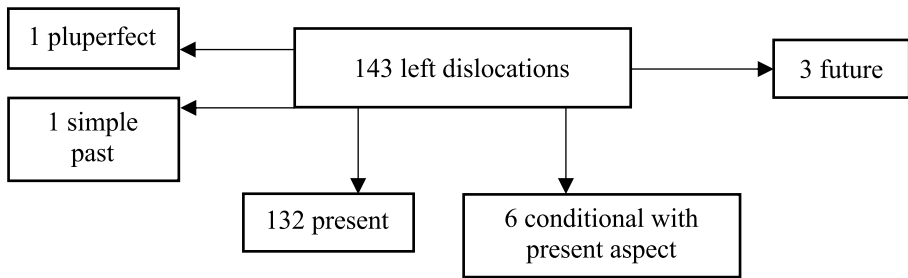


Figure 2. Tense of Left Dislocations (FRANTEXT)

Similarly, Galmiche notes that other quantifiers (i.e., **beaucoup de*, *quelques*, *plusieurs*, *certains*, *deux* [*nombre*]) are equally unacceptable in generic statements, but example (29) from the FRANTEXT corpus contains the quantifier ‘cent’.

- (29) **Cent francs**, même à la condition d’en rendre cent vingt, **ça** ne se lâche pas comme un chien dans une église... (Balzac, *Les Petits Bourgeois*)

Given the specificity of an amount of money, in this instance *ça* could hardly be interpreted as a marker of genericity. While this example could be understood as another instance of *ROG* (McLaughlin, 2011: 7), one could also posit that the dislocated element is resumed by *ça* to compensate for the interjection of a prepositional phrase.

In addition to considering restrictions in the scope of NPs imposed by particular determiners, it is important to examine limitations on generic readings resulting from tense selection. Given the association between generic statements and truths, we would expect to encounter verbs in the present tense in the main clause of examples. However, the attestation of examples in the past tense in the Carruthers corpus motivated an examination of the FRANTEXT examples with attention to tense (Figure 2).¹⁶

Unsurprisingly, most of the FRANTEXT examples contain a verb in the present indicative (132 or 92% of 143). Although the use of the simple past closely approximates the effect of examples in the present, namely, the communication of a constant state or ‘truth’ (example 30), the inclusion of the indicative pluperfect in example (31) expresses specificity on two levels: the employed tense and the presence of a demonstrative adjective illustrate that dislocations featuring co-referent clitic *ça* do not systematically impose generic interpretations.

- (30) **Les femmes**, **ça** n’a jamais su prendre les hommes. (Zola, *Travail*)

- (31) **ces** **jeunesses**, **ça** s’était parlé sans se rien dire. (Lamartine, *Geneviève*)

¹⁶Also see example 8 from Jeanjean (1985), which is in the past tense.

This observation is reinforced by the two examples in the future tense, which are also characterized by the presence of a demonstrative, as in:

- (32) **ces beaux cheveux blonds, ça tirera l'oeil.** (Sand, *Le Marquis de Villemer*)

and a dislocated NP modified by a possessive adjective:

- (33) **mes vers, ça durera toujours!** (Claudel, *Poésies diverses*)

Although four of the statements including a present conditional support a generic reading, there are two exceptions in which such an interpretation is less straightforward: although to a lesser degree than in cases involving a demonstrative, specificity is expressed through a comparison (i.e., a particular individual is positioned as representative of a kind of individual) in example (34). Specificity is also communicated by means of a possessive adjective in example (35):

- (34) **un gaillard membré comme vous, ça gagnerait de l'or dans not[re] pays.** (Moselly, *Terres Lorraines*)

- (35) **Vos boniments au mariage, ça pourrait bien faire du vinaigre.** (Aymé, *Clérambard*)

4.2.2 Pejorative attitudes

As previously mentioned, the presence of exclusively-female animate NPs in the oral corpus prompted further investigation into differences in the use of the dislocation under study according to gender.

In the FRANTEXT data, animates¹⁷ are seen to outnumber inanimates (81 animate examples, or 57% of 143). Of the animate referents encountered, 73 are human (90% of 81),¹⁸ and it would seem important to explore Jones's (1996: 260) observation that 'the generic use of *ce* and *ça* extends to human beings' and 'often conveys a pejorative attitude towards the persons concerned'. Examples expressing a pejorative attitude are classified as 'critical references' in this analysis.

In order to investigate the frequency with which different human referents are criticized, these NPs were divided into four sub-categories: children, men, women and ambiguous mixed-gender groups. The number of examples for each group is summarized in Figure 3.

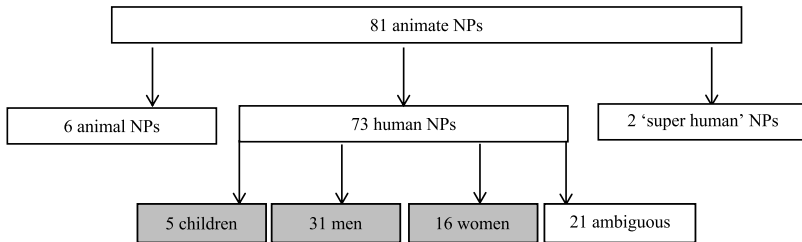
Twenty-one examples involved NPs of ambiguous or mixed gender. Since the masculine grammatical gender is employed in French when referring to an exclusively male group or a mixed-gender group, potentially mixed-gender adult NPs were not subject to gender analysis. For instance, it is impossible to determine the presence of women in NPs such as *les riches*:

¹⁷Metonyms that index animate beings were included in this category (i.e., mains; petits pieds; pieds nus; beaux cheveux blonds), as in: 'ces beaux cheveux blonds, ça tirera l'oeil' (Sand, *Le Marquis de Villemer*), and were classified as ambiguous with regard to gender, which is not explicitly expressed.

¹⁸The two 'super human' NPs include the soul (*l'âme*) and an angel (*un ange*).

Table 5. Critical References to Animate NPs (FRANTEXT)

References	Animals	Children	Men	Women
Critical references	1	2	6	12
Total references (N)	6	5	31	16
Total references (%)	17%	40%	19%	75%

**Figure 3.** Representation of Animate Sub-Categories (FRANTEXT)

(36) **les riches**, *ça* a toujours du temps à perdre. (Bazin, *Le Blé qui lève*)

Each example was examined within its surrounding context in order to determine whether it represented a criticism (Table 5). The difference in the percentage of critical references towards female versus male referents was found to be statistically significant.¹⁹

As Harris (1980: 154) observes, the distinction between proximity versus distance as conveyed through demonstrative determiners 'is at times pressed into service with a secondary function, e.g. to add a pejorative overtone'. That four of the 12 criticisms of women involve a NP that is modified by a demonstrative determiner (example 37) again calls into question the notion that dislocations characterized by co-referent clitic *ça* always and only impose generic readings.

(37) **Ces grandes dames**, *ça* ne va pas souvent à l'église (Stendhal, *Lucien Leuwen*)

4.2.3 Register

While detachment featuring *ça* may appear to reflect orality, it is surprising that only six examples from the FRANTEXT data sample present salient features of non-standard French (i.e., via contraction or particular non-standard orthography). A contraction is found in examples (34) and (38).

(38) **Les femmes**, *ça* n[e]'doute de rien (Leclercq, *Le Savetier et le financier*)

¹⁹The chi-squared test yielded a p-value of 0.0002 at the $\alpha = 0.05$ level of significance.

These examples can be understood as manifestations of the intermediary code, ROG (McLaughlin, 2011: 7), and the absence of the schwa [ə] can be interpreted as a feature of spoken French. A marked register is also signaled through lexical choices (example 39) and eye dialect, the representation of colloquial speech through non-standard orthography (examples 40 through 42).

- (39) Ces **guiantres** de noces, ça vous joue toujours ce tour-là (Leclercq, *Le Savetier et le financier*)²⁰
- (40) Non, **mossiou**.²¹ L'Italien, ça rit toujours, mais ça ne se radoucit point pour ça! (Sand, *La Daniella*)
- (41) Les miracles, ça se passe là-bas, loin, jamais **cheux** nous (Hamp, *Marée fraîche; Vin de Champagne*)
- (42) Un petit mémento de **tems en tems**, ça amuse (Pixérécourt, *Coelina, ou L'enfant du mystère*)

Conclusions on register or features of orality in graphic attestations of the type of dislocation under study cannot be drawn based on the limited examples highlighted above.²² Future syntactic research should explore and compare the register (as communicated on the lexical and syntactic level, e.g., the presence of ne-deletion) of the construction under investigation in both written and oral corpora.

4.3 Discussion of findings

Some of the corpus data has violated assumed constraints on generic statements. We would expect dislocated NPs to be preceded by 'articles génériques' (Galmiche, 1983: 30–31), as is generally the case in the small Carruthers (2013) data sample, and the high number of demonstrative determiners in the FRANTEXT examples is therefore surprising. NPs are shown to vary in scope in the written examples, ranging from *all* instances of an object (i.e., *tous les journaux*) to restricted groups (i.e., *ces grandes dames; ces gens-là*) to particular individuals (i.e., *le riche banquier Leuwen*). The consultation of examples from FRANTEXT nuances the association between this variety of dislocation and generic statements by accounting for tense and determiners that restrict the scope of the dislocated NP. Further, the difference in critical evaluation according to gender was found to be highly significant in the FRANTEXT study. Future investigations should further probe the extent to which this type of dislocation is mobilized to criticize women, among others, as well as the role of demonstrative determiners in making pejorative statements about more restricted groups. While the use of contractions, eye dialect, and references to *la campagne* and *les paysans* in the data is not the primary focus of the present study, their attestation in the data

²⁰According to Nisard's (1872) *Étude sur le langage populaire ou patois de Paris et de sa banlieue*, *guiantre* is a marked form of *diantre*.

²¹Although this lexical item does not appear within the dislocation under study, it underscores the importance of examining the surrounding context when interpreting examples.

²²Carruthers (1999: 16) highlights the difficulty posed by quantitative analysis of a rare syntactic form due to low frequency.

justifies future research on this construction as a means of representing orality (and a particular register) in the written.

In comparing the findings from the Carruthers and FRANTEXT corpora, we observe similarities as well as notable differences. Examples from both corpora reveal the greater (or exclusive, in the oral corpus) presence of women compared to other animate human referents. Data from both oral and written sources also illustrate the predominance of generic particles, however, the FRANTEXT sample contains a higher frequency of examples involving a possessive determiner and a sizable proportion of demonstrative determiners. While no final conclusions can be made based on this observation, we could hypothesize that the writers manipulate the prototypical construction in ways that diverge from spontaneous speech practices and express greater specificity. Future research should endeavour to further analyse differences in dislocation across oral and written corpora. The presence of examples in tenses other than the present in both corpora and the relatively higher proportion of past tense verb forms (three of eight examples) in the Carruthers corpus should also motivate research on tense use in the construction under study.

4.4 Methodological considerations

Although the inclusion of a written corpus may appear unintuitive due to the longstanding association of dislocation with orality, this study evinces the insights that can be offered by written data and attests to the presence of orality within the written (McLaughlin, 2011). Indeed, researchers should be mindful so as not to overlook the fact that authors are themselves speakers who mobilize particular ways of perceiving and speaking about others. Even the authors of texts that are widely assumed to express no bias, such as dictionaries, should not be mistaken as producers of neutral reference works (Steuckardt et al., 2011; Linares, 2016). The exclusion or inclusion of particular terms and the ways in which these words are embedded within particular metalanguage reflect the ideologies and aims of dictionary creators (Steuckardt et al., 2011: 26).

From a functional approach to language, authors construct desired meanings through their selection of particular linguistic possibilities that extend beyond the lexicon. As Byrnes, Maxim, and Norris (2010: 46) write, 'Grammar provides the energy behind the semiotic system of language in that it presents networks of interlocking options, with particular wordings signaling construal of a particular sociocultural context'. The authors whose work contributed to the FRANTEXT data set are themselves speakers, aware of the signifying potential of the various linguistic choices available to them, such as the decision to avoid dislocation entirely, to construct a dislocation in which the clitic transparently corresponds to the NP, or to employ a dislocation with co-referent *ça*. In other words, the same choices available to speakers are available to writers who imagine and animate words through the narrators and characters of their literary productions. For this reason, while I acknowledge the motivation to attribute all written dislocations to orality, in so doing we risk overlooking other signifying functions of syntactic constructions. Researchers should not neglect

written corpora in future syntactic analyses of ‘oral’ constructions but should instead view them valuable sources of insight and comparison.

5. CONCLUSIONS AND FUTURE DIRECTIONS

This study, which represents an initial corpus-driven exploration of a rare syntactic form, dislocation in which a dislocated element occurs with co-referent clitic *ça*, has contributed three main findings to French syntactic research. Firstly, in light of the varying scope of dislocated NPs encountered in the corpus data, the research questions the notion that dislocation with clitic *ça* always and only imposes a generic reading. Secondly, the written data reveals a significant difference in the number of women versus men featured in examples expressing criticisms. Although it is premature to draw any definitive conclusions, further research should continue to examine how this type of dislocation may be deployed more frequently when making pejorative comments about particular groups. Thirdly, this investigation nuances the association between dislocation and orality (McLaughlin, 2011) and consequently encourages methodological practices that consider both written and spoken texts as complementary sources of dislocation data.

Future corpus-based investigations focused on this construction should consider a broad range of spoken as well as written data, including but not limited to online forums or blogs that may generate discussions of perceived sociocultural differences,²³ social media platforms, literary texts, and the press. Elicitation techniques can also be employed to collect examples of this construction. The elicited data could then be complemented by oral and written questionnaires aimed at assessing semantic overtones that speakers attribute to dislocation featuring co-referent *ça* versus other syntactic alternatives (see Carruthers, 1999). Beeching (1997) reminds us that ‘the relationship between form and function is not static’, and it is necessary to look to the surrounding discursive context—and, when possible, to speakers (or writers) themselves—in an attempt to identify the function of a particular construction embedded within a particular set of circumstances.

The data from this study compel us to interrogate the notion of dislocation as a ‘linguistic variant’. Coveney (2003: 112) asserts that ‘les variantes doivent représenter “différentes façons de dire la même chose”’, but Carruthers (1999: 14) argues that ‘it is very difficult to argue that two syntactic constructions are ever truly semantically and pragmatically equivalent’, citing the case of dislocation. She holds that ‘it is not possible to argue that such constructions “mean the same thing”’ (p. 16). Surveys would help us to better understand the extent to which constructions ‘mean the same thing’ when *ça* substitutes a ‘traditional’ subject (e.g. *il*, *elle*) and would assist us in assessing the complex ways in which syntax contributes to nuanced meaning.

²³Research on generalizations or stereotypes about speakers or individuals from different genders, cultures (e.g., Les Français, *ça* ... Les Américains, *ça*...), or socioeconomic classes represents a potentially fruitful avenue for future research.

ACKNOWLEDGEMENTS. I am grateful to Mairi McLaughlin for her feedback while I was carrying out this study. I would also like to thank the anonymous reviewers for their very careful reading of earlier versions of this manuscript and many insightful comments.

REFERENCES

- Ashby, W. (1988). The syntax, pragmatics, and sociolinguistics of left- and right-dislocation in French. *Lingua*, 75: 203–229.
- Ayres-Bennett, W. and Carruthers, J., with R. Temple (2001). *Problems and Perspectives: Studies in the Modern French Language*. Harlow: Longman.
- Barnes, B. (1985). *The Pragmatics of Left Detachment in Spoken Standard French*. Amsterdam: John Benjamins.
- Beeching, K. (1997). French for specific purposes: the case for spoken corpora. *Applied Linguistics*, 18(3): 374–394.
- Beeching, K. (1999). Update on spoken French. *Bilingualism, Language and Cognition*, 2(2): 81–83.
- Blanche-Benveniste, C. (2006). Detachment constructions. In: K. Brown (ed.), *The Encyclopedia of Language and Linguistics* (2nd edn) Vol. 3. Oxford: Pergamon, 477–485.
- Blanche-Benveniste, C., Delofoeu, J., Stefanini, J. and van den Eynde, K. (1984). *Pronom et syntaxe. L'approche pronominale et son application au français*. Paris: SELAF.
- Blanche-Benveniste, C., Deulofeu, J., Stéphanini, J., and van den Eynde, K. (1987). *Pronom et syntaxe. L'approche pronominale et son application en français*. Paris: SELAF.
- Blasco, M. (1997). Pour une approche syntaxique des dislocations. *Journal of French Language Studies*, 7: 1–21.
- Blasco-Dulbecco, M. (1999). *Les dislocations en français. Étude syntaxique*. Paris: Éditions Champion.
- Blasco-Dulbecco, M., and Caddeo, S. (2001). Apposition et dislocation: la séquence pronom + lexique + clitique. *Recherches sur le français parlé*, 16: 125–149.
- Blasco, M., and Cappeau, P. (1993). Les relations clitique/lexique. Étude de deux cas particuliers. *Recherches sur le français parlé*, 12: 35–57.
- Burston, J. L. and Burston-Monville, M. (1981). The use of demonstrative and personal pronouns as anaphoric subjects of the verb être. *Linguisticae Investigationes*, 5: 231–257.
- Byrnes, H., Maxim, H. H. and Norris, J. M. (Eds.). (2010). Monograph issue: Realizing advanced foreign language writing development in collegiate education: Curricular design, pedagogy, assessment. *The Modern Language Journal*, 94.
- Cadiot, P. (1988). De quoi ça parle? À propos de la référence de ça, pronom-sujet. *Le Français Moderne*, 65: 174–192.
- Carlier, A. (1996). 'Les gosses, ça se lève tôt le matin': l'interprétation générique du syntagme nominal disloqué au moyen de ce ou ça. *Journal of French Language Studies*, 6/2: 133–162.
- Carruthers, J. (1999). A problem in sociolinguistic methodology: investigating a rare syntactic form. *Journal of French Language Studies*, 9/1: 1–24.
- Carruthers, J. (2013). *French Oral Narrative Corpus*. Queen's University Belfast: Oxford Text Archive.
- Cinque, G. (1977). The movement nature of left dislocation. *Linguistic Inquiry*, 8/2: 397–412.
- Coveney, A. (2003). Le redoublement du sujet en français parlé: une approche variationniste. *Études Romanes*, 54: 111–139.
- De Cat, C. (2005). French subject clitics are not agreement markers. *Lingua*, 115: 1195–1219.
- De Cat, C. (2007). *French Dislocation: Interpretation, Syntax, Acquisition*. Oxford: Oxford University Press.
- Delais-Roussarie, E., Doetjes, J. and Sleeman, P. (2004). Dislocation. In: F. Corblin, and H. de Swart (eds), *Handbook of French Semantics*. Stanford, CA: CSLI Publications, pp. 501–528.
- Deulofeu, J. (1977). La syntaxe et les constructions binaires. *Recherches sur le français parlé*, 1: 30–62.
- Fradin, B. (1988). Approche des constructions à détachement. La reprise interne. *Langue française*, 78: 26–56.
- Gadet, F. (1991). Le parlé coulé dans l'écrit: le traitement du détachement par les grammairiens du XXe Siècle. *Langue française*, 89: 110–124.
- Galmiche, M. (1983). L'Utilisation des articles génériques comme mode de donation de la vérité. *LINX*, 9/2: 29–87.

- Harris, M.** (1980). The marking of definiteness in Romance. In: J. Fisiak (ed.), *Historical Morphology*. The Hague: Mouton, pp. 141–156.
- Jeanjean, C.** (1985). Le ‘redoublement du sujet’. Étude grammaticale d’un corpus de français parlé. *Actes du XVIIe congrès international de Linguistique et Philologie Romanes*, Aix-en-Provence, 4: 233–246.
- Jones, M.** (1996). *Foundations of French Syntax*. Cambridge: Cambridge University Press.
- Lambrecht, K.** (1981). *Topic, Antitopic and Verb Agreement in Non-Standard French*. Amsterdam: John Benjamins.
- Lambrecht, K.** (2001). Dislocation. In: M. Haspelmath, E. König, W. Oesterreicher, and W. Raible (eds), *Language Typology and Language Universals: An International Handbook*. (Handbücher zur Sprach- und Kommunikationswissenschaft, 20). Vol. 2. Berlin/New York: Walter de Gruyter, pp. 1050–1078.
- Larsson, E.** (1979). *La dislocation en français: étude de syntaxe générative*. Lund: CZK Gleerup.
- Linares, E.** (2016). Resuscitating a bygone political vision through language: Antoine Furetière and the resurrection of the corps politique. *Early Modern French Studies*, 38(2): 148–163.
- McLaughlin, M.** (2011). When written is spoken: dislocation and the oral code. *Journal of French Language Studies*, 21(2): 209–229.
- Nisard, C.** (1872). *Étude sur le langage populaire ou le patois de Paris et de sa banlieue, précédée d’un coup d’œil sur le commerce de la France au Moyen Âge, les chemins qu’il suivait, et l’influence qu’il a dû avoir sur le langage*. Paris: A. Franck.
- Riegel, M., Pellat, J.-C. and Rioul, R.** (1994). *Grammaire méthodique du français*. Paris: Presses Universitaires de France.
- Rowlett, P.** (2007). *The Syntax of French*. Cambridge: Cambridge University Press.
- Shuh, R.** (1999). Du Salon à la Chambre: Wit and Politics in Stendhal’s Lucien Leuwen. *The Romantic Review*, 90(4).
- Steuckardt, A., Leclercq, O., Niklas-Salminen, A. and Thorel, M.** (Eds.). (2011). *Les dictionnaires et l’emprunt: XVI^e-XXI^e siècle*. Aix-en-Provence: Publications de l’Université de Provence.
- Tristram, A.** (2012). L’accord sujet-verbe en français contemporain: une étude de variation grammaticale et sociale. In CORELA– Cognition, Représentation, Langage. Retrieved from <https://journals.openedition.org/corela/2614>

Cite this article: Linares E (2019). ‘Ces grandes dames, ça ne va pas souvent à l’église’: A corpus-based investigation of dislocation featuring co-referent clitic *ça*. *Journal of French Language Studies* 29, 423–441. <https://doi.org/10.1017/S0959269519000152>