Clefts in context: A QUD-perspective on *c'est* / *il y a* utterances in spoken French

Alessia Cassarà

Universität Stuttgart
alessia.cassara@ling.uni-stuttgart.de

Aria Adli

Universität zu Köln
aria.adli@uni-koeln.de

Lena Karssemberg

KU Leuven
lena.karssenberg@kuleuven.be

Abstract

In this paper we present the results of a pragmatic analysis of French full clefts and monoclausal c'est/il y a utterances (e.g. c'est la femme qui l'a tué 'it's the wife who

killed him' vs. c'est la femme 'it's the wife' respectively in answer to the question 'who killed him?'), when these structures are used as pragmatic strategies to focalize the subject in spoken French. Unlike full cleft sentences, monoclausal c'est and il y a utterances have received less attention in the literature, especially with regard to focus and its realization in spontaneous speech. Investigating the opposition between full clefts and monoclausal forms as well as the questions that these clefts answer allows us to arrive at a more precise understanding of the discourse functions of these structures and the pragmatic contexts in which they are felicitous. The corpus that is used (sgs, spontaneous spoken French) contains many question-answer pairs due to its interactive setup, thus enabling a clear analysis of the types of Question Under Discussion that the clefts answer. The data show that monoclausal utterances are more likely to answer highly active QUDs, whereas full clefts are more likely to answer less active QUDs. The level of activation is determined in terms of proximity and implicitness of the QUD (immediately-preceding the cleft, further away or implicit), and - when the question is uttered explicitly - modality (wh or yes/no) also plays a role.

Keywords: cleft, French, Question Under Discussion, Information Structure, focus.

1   Different types of clefts and information structure

Cleft sentences such as (1) and (2) have received an increasing amount of attention, especially in literature focusing on information structure. It is widely claimed that their main function is to indicate that the clefted element, such as *moi* 'me' in (1), is focal or that the whole sentence is focal as in (2). Constructions like (1) are called *c'est* clefts, and (2) *il y a* clefts.

(1)   A: *Et   qui   est-ce qui   l' a   découvert?*
              and who is   who it has found?
            'And who found him?'

         B: ***Alors,   c'est moi qui   ai   découvert la   personne***.
              so   it.is me who has discovered the person
            'So, I found the person.'

                                          (*c'est* cleft, taken from sgs French, 35/16-17)

(Out of the blue context)

(2)   A: *Maman!* **Il y   a le chat qui   déchire les rideaux!**
              mom   there is the cat who tears up the curtains
            'Mom! The cat is tearing up the curtains!'

                                          (il y a cleft, taken from Culioli 1990)

Numerous studies have shown that both *c'est* and *il y a* clefts are more frequently used as means to focalize a subject rather than another grammatical role (cf. the argument asymmetry in focus realization, e.g. Skopeteas & Fanselow 2010; Belletti 2015; Destruel 2016; Karssenberg 2018).[1] Yet, subject focus can also be expressed by other competing means, such as the monoclausal utterances in (3) (also called 'reduced clefts' in the literature, Jenkins 1973):

(3)   A: *Qui (est-ce que qui) a   parlé?*
              who (Q   who) has talked
              'Who talked?'
          B: ***C'est Jean** ~~qui a   parlé~~.*
              it.is   Jean   ~~who has talked~~
              'Jean ~~talked~~'

(Belletti 2009: 192)

Therefore, the precise conditions under which a cleft form rather than a similar varian is chosen by the speaker are not yet fully known. In this article, we conduct a corpus analysis of the discourse conditions under which cleft sentences and monoclausal utterances occur, and of their different pragmatic functions. A promising approach to analyze subject focus (or, more generally, focus) in its discourse context is the Question Under Discussion approach. The framework, proposed by von Stutterheim and Klein (1989), and van Kuppevelt (1995), then further developed by Roberts (1996) and Riester et al. (2018), amongst others, is a highly structured, meaning-based approach to identify focus and background of every utterance by suggesting a

---

[1]   Authors such as Skopeteas and Fanselow (2010) and Destruel (2016) have shown empirically the existence of an asymmetry with respect to focus realization in French: while subject focus is usually marked through a cleft sentence, object focus does not require any specific syntactic marking, and can be conveyed in situ (i.e. maintaining a SVO order). In a similar vein, Belletti (2015) claims that contrary to subject clefts, object clefts cannot function as a possible answer to a pure question of information, although this asymmetry between subject and object clefts seems to be contradicted by examples such as (i).

(i)   A: *Qu' est-ce   que   le fermier a   brossé?*
              what is   that   the farmer   has brushed
              'What is it that the farmer brushed?'
          B: *C'est le   cheval que le fermier a   brossé.*
              it.is   the horse that the farmer has brushed
              'The farmer brushed the horse.'

(taken from Destruel 2013: 101, our glosses)

discourse tree, in which every assertion is assigned to a corresponding QUD. In this paper, we apply this approach to the analysis of cleft sentences and monoclausal utterances in spoken dialogues. First, it allows a more solid analysis of the focus-background partition and the focus type (contrastive or non-contrastive). Second, it gives us a new insight into how focus expressed by these constructions is used to manage Common Ground (CG) as the discourse unfolds. This how is analyzed in terms of the distance between the construction under investigation and the corresponding QUD. Is the Is the QUD implicit or explicit, and if explicit, does the speaker utter the answering strategy immediately after the QUD or not? The results of this study will then lead to a follow-up analysis on the specific modality of the QUD. Are specific answering strategies more common with some question types rather than others?

In line with previous findings, we limit our investigation to subject clefts with a focus-background partition. Thus, clefts containing a constituent other than the subject in the clefted element (e.g. object clefts) or subject clefts that did not show a focus-background partition (e.g. all-focus clefts) have been excluded from the analysis. The factor that mainly distinguish cleft sentences from monoclausal utterances is the overt expression of the background proposition. In clefts with a focus-background partition, such a proposition is encoded in the relative clause. Monoclausal utterances – that, depending on one's theoretical persuasion, some may also call reduced or elliptical clefts (Jenkins 1973) – on the other hand, are typically concentrated on the focal part, not uttering the reconstructable background. We distinguish throughout this paper between full forms such as (1) and (2) and monoclausal ones such as (3), (4) and (5). Thereby, we can also identify the pragmatic motivations that determine speakers' preferences for a full cleft or a monoclausal utterance.

(4)   A: *Donc est-ce que tu   sais   qui   a   découvert le   corps*?

        so   Q   you know who has found   the body

        'So do you know who discovered the body?

    B: ***Oui, c'est sa femme***.

        yes it.is   his wife

        'Yes, his wife.'

                    (Monoclausal *c'est* utterance, taken from *sgs* French, 5/64-65)

(5)   A: *Mais dis, c' était qui   qui habitait au-dessus*?

        but   tell it.was who that lived   above

        'But tell me, who was it who lived above?'

B: ***Alors   au-dessus, donc, il y   a   un couple d' étudiants***.
     so   above   so   there is   a   couple of.students
   'So above, then, there is a couple of students.'

(Monoclausal *il y a* utterance, taken from *sgs* French, 39/277-278)

Given our QUD-based approach to the focus-background structure of clefts, we concentrate on assertive clefts that answer to a QUD. We leave interrogative clefts (see e.g. Druetta 2018: 32), which are themselves QUDs, for future research. We build our analysis on 347 sentences (both clefts and monoclausal utterances), extracted from spontaneous speech dialogues of the French sgs corpus (Adli 2011). This bottom-up corpus-based approach allows for a more accurate study of the use of these structures in real-life conversations. Thereby, we seek to gain a more valid empirical picture as opposed to a study that builds on constructed examples and the authors' introspection. We also take advantage of the fact that the sgs corpus contains a high number of explicit question-answer sequences. This will turn out to be crucial when we discuss the importance of the QUD as a factor influencing the choice between full clefts and monoclausal utterances. The structure of this paper is as follows. In section 2, we briefly state our definition of focus and the QUD approach. We then summarize previous analyses of *c'est* clefts, *il y a* clefts and monoclausal utterances. In section 3, we describe the pragmatic and conversational characteristics of the *sgs* corpus used for this study and explain how the data were extracted and annotated. Section 4 is devoted to the analysis of the data, first presenting an overview of the frequency of occurrence of each cleft type, after which we turn to the information structure properties of the clefts themselves, and finally to the properties of the preceding Question Under Discussion. The results issued from the Questions Under Discussion analysis motivated us to do a follow-up study on the possible role of the modality (yes/no or wh) of the Question under Discussion on the choice of the answering strategy, reported in section 5. Finally, in section 6, we discuss the findings that have emerged from the corpus studies and their implications for a theory of clefts in spontaneous speech.

2   Previous analyses

2.1   *Focus and the QUD approach*

A crucial element in our analysis of cleft sentences is the definition of focus. We follow Krifka's (2007) alternative-semantic approach, where focus is defined as signaling the presence of alternatives relevant for the interpretation of the respective expression. More specifically, a careful distinction between non-contrastive uses of focus (henceforth information focus, i-focus or [-contrast]) and contrastive ones (henceforth contrastive focus, c-focus or [+contrast]) is required. In order to do so, we adopt the differentiation advocated in Krifka (2007) between two macro-categories of focus types: information and contrastive focus. We consider that information focus expresses

the most important part of the utterance, or what is new in the utterance. Consider the example taken directly from the corpus under investigation:

(6)   A: *En fait, il y   avait qui dans l'immeuble   quand ça   s'est passé?*
            in fact there was   who in   the.building   when this se.is happened
            'So, who was there when this happened?'

        B: ***Alors, en fait, il y   avait la   grand-mère   du dessus qui était là***.
            so   in fact there was the grand-mother of above who was there
            'So, actually, there was the old lady from upstairs who was there.'

                    (Information focus, *Il y a* cleft, taken from *sgs* French, 9/261-262)

In (6b), the referent *la grand-mère du dessus* 'the old lady from upstairs' represents the part of the answer that is new to the hearer and that fulfils the wh-variable contained in the question. It is therefore a typical case of information focus. Moreover, we use the term contrastive focus for those cases where the focused subject is selected among a set of explicitly-mentioned alternatives, or when it replaces the variable already contained in the previous utterance. This implies that both the case of selection among a closed set of explicitly-mentioned alternatives (Chafe 1976; Kiss 1998) and the case of correction (Steube 2001; van Leusen 2004; Gussenhoven 2008; Repp 2010) are here conflated under the same category, as in (7):

(7)   A: *Est-ce que c'est toi qui   l'as   découvert, ou?*
            Q   it.is you who it.has found   or
            'Is it you who found it, or not?'

        B: ***Non, c'est sa femme qui   l'a   découvert***.
            no   it.is   his wife   who it.has found
            'No, his wife found it.'

                    (Contrastive focus, *c'est* cleft, taken from *sgs* French, 8/51-52)

In natural speech, the analysis of focus and background is anything but straightforward. At this point, the QUD-approach proves highly useful. Roberts (1996) provided both a general QUD-based approach to pragmatics and an account of focus. This inspired much further theoretical work on information structure including Büring (2003) on contrastive topics and Beaver and Clark (2008) on focus particles. The intuition behind it is the same as that behind the well-known question-answer test used to detect the focus structure of a sentence. An answer to a question is appropriate only if its focused constituent corresponds to the wh-phrase of the question. In other words, focus can be defined as that part of the answer that satisfies the wh-variable contained

in the question (Paul 1880). This method is particularly straightforward in the case of an explicit QUD, such as in a dialogical situation, like the one exemplified in (8):

(8)   A: **Who** ate the cake?
         B: [FOC Mary] did.

In spontaneous speech, however, utterances are often answers to *implicit* questions. In these cases the QUD has to be reconstructed starting from the answer itself and by looking at the preceding discourse context, as proposed by Riester et al. (2018). The authors suggest a series of constraints that regulate the reconstruction of the QUD and allow its derivation from the preceding discourse context. The constraints are derived from the focus literature of the past decades, in particular Rooth (1992), Schwarzschild (1999), and Büring (2008).

For instance, implicit Questions Under Discussions must be answerable by an assertion that they immediately dominate (called *Q-A-Congruence*), they must contain as much given material as possible (called *Maximize-Q-Anaphoricity constraint*) and can only consist of given or at least highly salient information (called *Q-Givenness constraint*).[2] Let us consider the following discourse section taken directly from Riester et al. (2018).

(9a)   And all I can say is that his condition was extremely bad during his last years
(9b)   He literally suffocated.

The assertion in (9a) can give rise to a set of alternative QUDs to which (9b) can function as the answer. Let us assume, for the sake of explanation, the following implicit QUDs:

Q1: What happened?
Q2: What about him?
Q3: Who literally suffocated?
Q4: Who owns a bicycle?

According to Q-A-Congruence, Q4 is ruled out because it cannot be answered by the assertion that it immediately dominates. Q1 is also excluded because of the Maximize-Q-Anaphoricity constraint, stating that the QUD must contain as much given or salient material as possible and that a question like Q1 does not contain any of the given material found in the assertion. Finally, the Q-Givenness constraint, according to which an implicit QUD cannot introduce new material (except for function words and wh-particles), excludes Q3, because the fact of being "literally suffocated" is new information and fails to provide a link with previous context. Conversely, the personal

---

[2]   The *salience* of a word is defined in Riester et al.'s (2018) guidelines as its active presence in the addressee's mind right before its actual occurrence in the text.

pronoun 'he' is connected to the previous sentence by virtue of reference continuity ("his condition" in(9a)). Q2 is thus the optimal candidate as the implicit QUD of the assertion contained in (9b). By virtue of these principles, (9b) is then interpreted with a topic-comment partition, in which the pronoun *he* is the topic, and the rest of the sentence is the comment part.

2.2   C'est clefts

In the literature on clefts, most attention has been given to clefts introduced by the expression *it is* and its cross-linguistic equivalents. This also holds for French: there is an abundant body of work on *c'est* clefts (e.g. Katz 2000b; Lambrecht 2001b; Dufter 2008; Carter-Thomas 2009; Dufter 2009a, b; Lahousse & Borremans 2014; Karssenberg & Lahousse 2018). *C'est* clefts can have a specificational interpretation, as illustrated in (10): the clefted element ("my brother") expresses the value for the variable implied in the preceding question, which is expressed by the cleft relative clause ("who did that").

(10)   *Tu   as   vu   ma voiture ?* **C'est mon frère   qui   a   fait   ça.**
        you have seen my car?   it.is my   brother who has done that
        'Did you see my car? It's my brother who did that.'

                    (taken from Katz 2000a: 264, our translation)

One key feature of cleft sentences is that they can be transformed into a sentence with subject-verb word order without a change in meaning (11).

(11)   *Mon frère   a   fait   ça.*
        my   brother has done that
        'My brother did that.'

        With respect to information structure, it has been argued that *c'est* clefts can convey a range of different articulations. Traditionally, the focus-background interpretation, illustrated by (10), is most well-known: the clefted element is focal (either with or without a contrastive interpretation), while the relative clause constitutes the background (see Smits 1989; Kiss 1998; Clech-Darbon et al. 1999; Katz 2000b; Lambrecht 2001a; Rialland et al. 2002; Doetjes et al. 2004; Dufter 2008; Mertens 2008; Destruel 2013; Lahousse & Borremans 2014). However, *c'est* clefts can also express all-focus as in the beginning-of-speech example (12), and it has been claimed that the clefted element can also be topical (13) (see Clech-Darbon et al. 1999; Dufter 2006; Huber 2006; Scappini 2006; Dufter 2008, 2009b). Indeed, Dufter (2008) shows in his diachronic corpus data that *c'est* clefts introducing an anaphoric (and hence discourse-given) expression and a discourse-new relative clause, such as (14), are increasingly frequent (see also Mertens 2012: 134). These studies show that there

is no one-to-one mapping between form (*c'est* cleft) and information structure articulation.

(12)   *C'est avec beaucoup de tristesse que l'on suit les événements qui se déroulent au Nigéria.*

'It is with great sorrow that we follow the events in Nigeria.'

(taken from Dufter 2009b: 105)

(13)   *Françoise, je te présente Philippe. C'est lui qui a écrit sa thèse sur Sade.*

'Françoise, this is Philippe. It's he who wrote his dissertation about Sade.'

(taken from Katz 2000a: 265)

(14)   *En quarante-deux, les bombardements ! Et c'est là que beaucoup de gens sont partis de Toulon.*

'In 1942, the bombings! And it's then that many people left Toulon.'

(taken from Blanche-Benveniste 2006: 282, our translation)

As for the motivations why speakers would use *c'est* clefts, various functions have been suggested. One prominent function is that of focus marking: the cleft format highlights part of the utterance (e.g. Lambrecht 2001b). Other motivations include increasing textual coherence at the level of the paragraph (Dufter 2009b; Jacob 2015). Moreover, in contrast to their canonical subject verb counterparts (e.g. *Il a écrit sa thèse…* 'He wrote his thesis...' in (13)), *c'est* clefts are argued to present part of the proposition as presupposed, and they carry an exhaustiveness implicature (see Dufter 2006; Destruel 2013; Karssenberg 2017b: 229)

2.3   Il y a clefts

Apart from the more well-known *c'est* clefts, an increasing number of studies has been devoted to *il y a* 'there is' clefts, prototypically illustrated by utterances such as (15).

(15)   *Y   a   le   téléphone qui sonne.*

there is the phone   that rings

'The phone is ringing.'

(Lambrecht 1988: 136)

Like *c'est* clefts, *il y a* clefts can be transformed into a counterpart with subject-verb word order, as in (16).

(16)   *Le téléphone sonne.*

the phone   rings

'The phone is ringing.'

(Lambrecht 1988: 136)

Il y a clefts such as (15) are semantically predicational rather than specificational, i.e. they are argued to predicate a property (*it's ringing*) of a referent (*the phone*) (see Higgins 1979; Roy 2013: 8-9 about the predication-specification distinction, among others). Moreover, in terms of information structure, such examples display sentence focus (or all-focus): the whole sentence contains discourse-new information and is highlighted. However, corpus research has confirmed that next to these predicational, all-focus *il y a* clefts, there is also a less frequent category of semantically specificational *il y a* clefts expressing a focus-background articulation (Karssenberg 2017b; Verwimp & Lahousse 2017). This type of *il y a* cleft is illustrated in (17): the clefted element (*la citroën c1*) provides a value for the discourse-given variable "cars that cost…". The relative clause is backgrounded with respect to the focal clefted element.

(17)   A: *Je recherche des   modèles de voiture à acheter neuve moins de   10 000*

I   search   some models of cars   to buy   new   less   than   10.000

euros, où   aller?

euros   where to go?

'I'm looking for new car models that cost less than €10.000, where should I go?'

B: *bonjours. il y   a   la   citroen c1 qui est   a   moins de 10 000 euros.*

hello   there is   the   citroen c1 that is   at   less   of 10.000 euros.

'Hello.   The Citroën C1 costs less than €10.000.'

(taken from Karssenberg 2017b: 251, spelling mistakes from corpus example not corrected)

In contrast to *c'est* clefts, specificational focus-background *il y a* clefts such as (17) are less frequently associated to an exhaustiveness implicature (Lambrecht 2001b: 504; Karssenberg 2018: 195). For instance, while the Citroën C1 is one possible value for the variable "cars that cost less than €10,000", other values could also be added. It is possible to add the (*ne…*) *que* 'only' particle to the sentence in order to obtain an exhaustive reading (*il y a que* 'there is only'). However, in this case the exhaustiveness would be expressed by a lexical element and hence be part of the semantics of the sentence, rather than being a pragmatic implicature as in the case of *c'est* clefts (see Karssenberg 2018: 194-196, for more details about this distinction). Importantly, as is the case with *c'est* clefts, there is no one-to-one mapping of form (*il y a* cleft) and information structure articulation.

2.4   *Monoclausal c'est / il y a utterances*

Both clefts and monoclausal *c'est / il y a* utterances have been argued to function as answering strategies and are frequently found in spoken French in contexts of question-answer pairs. Monoclausal utterances such as (18) and (19) resemble the first part of a cleft: although a relative clause can easily be constructed (*qui a appelé* 'who called' in (18) and *qui habitait au-dessus* 'who lived upstairs' in (19)), this relative clause is not expressed.

(18)   A: *Qui   a   appelé?*

who has called

'Who called?'

B: ***C'est la   personne de l' appartement A, au   deuxième étage.***

it.is the person   of the.apartement   A at-the second   floor

'The person from the apartment A, on the second floor.'

(Monoclausal *c'est* utterance, taken from sgs French, 29/273-274)

(19)   A: *Et,   aux autres étages, qui   habite là?*

and in   other   floors   who lives   there

'And, who lives on the other floors?'

B: *Alors, aux autres étages**, il y   a un couple d'étudiants, une jeune fille seule.***

So   on other floors there is a   couple of.students   a   young girl alone

'So above, then, there's a couple of students, a young girl.'

(Monoclausal *il y a* utterance, taken from *sgs* French, 39/165-166)

There is no consensus about the syntactic relationship that holds between these forms. The first view is that monoclausal utterances such as (18) and (19) are clefts with unexpressed (or elided) relative clauses, as argued for by Belletti (2009), Hedberg (2000), Doetjes et al. (2004: 538), or Büring (1998). The opposite view is that clefts should in fact be seen as 'extensions' of sentences such as (18) and (19), meaning that there would be nothing 'reduced' about monoclausal utterances from a syntactic point of view (Birner et al. 2007: 319; Davidse et al. 2016; Karssenberg 2018: 172). In this paper, we remain agnostic on the exact syntactic nature of these structures. Rather, we are here concerned with their alternation with full clefts, and how their information-structural functions contribute to the universe of discourse.

Interestingly, Belletti (2009), building on acquisition studies on answering strategies in French, notices that the most natural answer to an immediately preceding question on the subject (e.g. *Who arrived?*) is not a 'full' cleft but rather a monoclausal *c'est* utterance (*It was Mary*). In a similar vein, Karssenberg (2017b) observes in her analysis of specificational *il y a* clefts such as (17) that in certain cases, the relative clause (almost) literally repeats the variable that is raised in the preceding question. In such cases, the relative clause could also be omitted, as in (20):

(20) [Same context as (17)]
    Bonjour. **Il   y   a   la   Citroen   c1**.
    Hello   there is the Citroen c1
    'Hello. There's the Citroën C1.'

Karssenberg argues that the writer nonetheless expresses the relative clause because it fulfills a specific pragmatic function: "(…) the speaker repeats the variable in order to emphasize the relevance of his/her utterance with respect to the QUD. Put somewhat informally, by repeating the question, the speaker indicates: 'See, I'm really answering your question'" (Karssenberg 2017a: 83; 2017b: 251).

Even if this intuition is true, however, this does not explain *why* speakers feel the need to explicitly relate back to the Question Under Discussion in some cases (full clefts), but not in others (monoclausal utterances). Important in this respect may be that in spontaneous speech, there is a tendency to elide material when this is highly activated in the mind of both speaker and hearer, while it is generally overtly expressed when its activation has somehow decreased or contains new information (see also Krifka 2007). In section 0 we will ask whether one can account for speakers' choice between a full cleft or a monoclausal form assuming that they are more likely to use a monoclausal form when the QUD is highly activated, and they are more likely to use a full cleft when the QUD is not highly activated.

3   The corpus approach

3.1   The French sgs corpus of spontaneous speech

The *sgs* corpus is a database of spoken language for French, Spanish, Catalan, and Persian. We use the French data collected by Adli in July/August 2005 in Paris. The strategy used for the collection of speech data was to engage the participants in a dialogue on a fictive scenario in which they took on the role of a police investigator in a suspected murder case. The interviewer talked to one of the well-trained local fieldwork assistants who took on the role of the doorman of the building that the dead body had been found in. This procedure allows participants to choose which aspects of the case they want to talk about. Moreover, it has the advantage of requiring the

interviewee to produce both declarative and interrogative utterances – the latter being rare in classic interview situations, in which the interviewer asks and the interviewee answers. The dialogical character containing many overt questions is particularly useful for the analysis of the QUDs. The interviewees consist of 101 French native speakers in the age range 19-49 with the mean value at 29 and the median at 27 years. 56% are female.

3.2   *Finding and categorizing cleft and monoclausal sentences in the corpus*

Any cleft type containing a subject in the clefted element has been automatically extracted using keywords, such as *ce/c'...qui* or *i(l) y... qui*. Each cleft sentence has been extracted together with the (either explicit or implicit) QUD it constitutes an answer to. To prevent any error of the automatic extraction, the corpus has been manually screened as well. As already mentioned in the introduction of this paper, prior to undertaking the annotation process, we excluded any clefts not exhibiting a focus-background (e.g. all-focus clefts or clefts with a topic-comment partition). Additionally, syntactic constructions that look like clefts on the surface, but do not meet the requirements to be clefts, e.g. due to the presence of a restrictive relative clause or the impossibility of declefting (i.e. "cleft lookalikes", Karssenberg 2017b), have been taken out as well.

Subsequently, the items have been annotated for the following factors: (i) answer type (distinguishing between *c'est* cleft, *il y a* cleft, pseudo cleft, monoclausal *c'est* sentence, and monoclausal *il y a* sentence), (ii) focus type in terms of [±contrast] (see examples (5) and (6), (iii) proximity of the QUD (distinguishing between immediately preceding QUD (tagged as "immediate QUD", non-immediate QUD, and implicit QUD). Moreover, the follow-up analysis required the annotation of, (iv), the modality of the QUD (distinguishing between *wh*-QUD, and yes/no-QUD).

We present here some examples directly extracted from the *sgs* corpus, where it is possible to observe the three levels of immediacy of the QUD: immediate QUD (21), not-immediate QUD (22) and implicit QUD (23). In the first condition (immediate), the QUD immediately precedes the answer, and the background proposition (someone found a body) is, hence, highly activated. Consider the example in (21):

(21)   A: *Et   qui   est-ce qui l' a   découvert?*
              and who Q   that it.has discovered
              'And who found it?'

        B: ***Alors c'est   moi qui ai   découvert la   personne.***
              So   it.is   me who have discovered the person
              'So, I found the person.'

                                                                (*sgs* French, 35/16-17)

Conversely, in the not-immediate condition, the QUD ('Who takes care of the children') is not answered immediately. Rather, speaker B provides other pieces of information (e.g. 'They're not there right now') before answering the actual question. In this case, we consider that the level of activation of the background proposition has somewhat decreased, because of the presence of intermediate material (22).

(22)   A: *Et   les enfants, qui   s' en occupe?*

       and the children who s'en take care

       'And the children, who takes care of them?'

    B: *Donc, en ce   moment, ils   sont pas là.*

      so   in this moment they are not there

      'So, at the moment they are not there.'

    B: *Ils   sont en vacances.*

      they are   in holidays

      'They are on holidays.'

    B: *Sinon   **c'est une nourrice   qui les   garde le   soir.**[3]*

      otherwise   it.is   a   babysitter   that them watch the evening

      'In general, a babysitter takes care of them during the evening.'

                     (*C'est* cleft, taken from *sgs* French, 14/133-136)

A particular challenge has been the reconstruction of the implicit QUD. In line with the QUD approach, we have built on Riester et al.'s (2018) guidelines, whose constraints are illustrated in section 2.1. Let us try to reconstruct the implicit QUD answered by the assertion *C'est les voisins du haut qui ont organisé cette soirée* 'It is the neighbors from above who organized this party' contained in (23), considering both the material in the assertion and the context preceding it:

(23)   A: *Est-ce qu'il y   a   eu des   problèmes de bruit, etcetera?*

      Q   there has had some problems   of noise, etcetera

      'Were there noise problems, etcetera?'

---

[3]   It has to be pointed out that, in this example, the NP *le soir* (the party) constitutes a new piece of information and can be considered not-at-issue material with respect to the current QUD. These cases are discussed in Cassarà (2021), where it is argued that sometimes the use of a full cleft is motivated by the fact that the speaker attaches a new element to the background proposition contained in the relative clause of the cleft.

B: *Ben, la nuit   dernière, il y   a   jamais eu autant   de bruit dans*
   well the night last   there has never   had so much of noise in
   l'immeuble.
   the.building
   'Well, there has never been so much noise in the building like yesterday night.'

B: *Mais il y   avait une soirée.*
   but   there was a   night
   'But there was a party.'

A: *m.*
   'm.'

B: *Donc justement, **c'est les voisins   du   haut   qui ont   organisé***
   so   rightly   it.is the neighbours of-the above that have organized
   cette soirée.
   this   night
   'So, the neighbours from upstairs have organized the party.'

(*C'est* cleft, taken from *sgs* French, 59/ 334-338)

To avoid being biased by morphosyntactic cues, we do not take into account the fact that the assertion is a cleft sentence, and we only build on the semantic content of the dialogue. The Q-A-Congruence constraint rules out an implicit QUD such as 'what about him?' because it would not be semantically congruent with the answer. The referent *les voisins du haut* 'the neighbors upstairs' is discourse-new, and, therefore, cannot be part of the background proposition. In other words, a QUD such as 'What about the upstairs neighbors?' is excluded as well. The referent *soirée* 'party', however, has already been mentioned in the preceding utterance. Its co-reference with its antecedent is also testified by the use of the demonstrative *cette* 'this'. Applying the Q-Anaphoricity constraint, we assume that the implicit QUD must somehow contain this referent. The verb *organiser* 'to organize' has not been previously mentioned either. However, the word *soirée* 'party' activates a series of possible QUDs in the mind of the hearer, such as 'Who came to the party?', 'How many people?', 'How was the music?', 'How was the food?' or 'Who organized this party?'. Among the different possibilities, it seems plausible that the last question (Who organized this party?) is the most felicitous in context. Speaker A wants to know about possible noises in the building, speaker B informs speaker A of a party, and assumes that speaker A wants to know who the people responsible for the party/noise of that evening are. Thus, it is possible to assume that *organiser* 'To organize', although not properly given, is highly salient, and that its salience is activated by the referent soirée 'party'. For this reason,

it can be considered part of the implicit QUD. The discourse marker *justement* 'Rightly' is additional evidence of the link between the current assertion and the previous context, proving that 'Who organized the party?' is a sub-question of the main QUD *Est-ce qu'il y a eu des problémes de bruit?* 'Were there usually problems with noisy parties?'. Speaker A is thus signaling that s/he is providing additional information with respect to the noise heard that evening.

4   Empirical results

We will begin in section 0 with a general overview of the frequency of occurrence of the different cleft and monoclausal types and analyze, one by one, the three different factors focus type (section 0) of the cleft sentence, contextual immediacy/distance of the QUD before the assertive cleft (section 0), and the modality of that QUD (section **Error! Reference source not found.**). More precisely, section 0 presents the quantitative distribution of different focus types in *c'est* and *il y a* clefts, consistently distinguishing between cleft types and monoclausal utterances. Section 0 takes up the hypothesis previously formulated in section **Error! Reference source not found.**, according to which monoclausal forms tend to follow highly activated QUDs whereas full clefts tend to follow less activated QUDs. We separate our frequency analyses by immediate, non-immediate and implicit QUDs. Finally, we come back in section **Error! Reference source not found.** to our theoretical assumption on the internal QUD structure that differs between yes-no and *wh*-questions. The frequency pattern presented there will show modality to play a role in the speaker's choice between a full cleft or a monoclausal form.

4.1   Descriptive overview of the frequencies

The corpus exhibits 388 occurrences of subject cleft sentences out of approximately 34,000 utterances.[4] Unsurprisingly, *c'est* clefts (*n*=107, 27.58%) and *il y a* clefts (*n*=98, 25.26%) are by far the most frequent cleft types in the data, covering more than half of the occurrences. The corpus query also revealed a considerable number of occurrences for monoclausal utterances introduced by *c'est* (*n*=77, 19.85%) and *il y a* (*n*=60, 15.46%), covering together more than a third of the occurrences. In what follows, we will concentrate on these four types. The remaining cleft types are pseudoclefts (*n*=30, 7.73%) (e.g. *Moi, ce qui m'intèresse bien c'est son mari* 'Me, what interests me the most is her husband'), *(s')il y a… c'est* clefts (*n*=8, 2.06%) (e.g. *S'il*

---

[4]   This number builds on the corpus segmentation rule in Adli (2011: annex 1): Each utterance consists of a single root clause with its embedded clauses. Coordinated root clauses count as several utterances. Incomplete sentences are represented as an utterance, too. However, the total number of utterances does not included turns that only consist of an interactional marker (e.g. *d'accord* 'Ok') or of an isolated polarity expression (e.g. *ouais* 'Yep', *oui* 'Yes', *non* 'No').

y a bien quelqu'un qui a pu faire une chose pareille, c'est elle 'If there is someone that could have done something like this, it's her), inferential clefts (*n*=7, 1.8%) (e.g. *C'est qu'ils sont certainement passés par là* 'It's that they have certainly passed from here', and reverse pseudoclefts (*n*=1, 0.26%) (e.g. *C'est ça ce qui m'énerve vraiment* 'This is what really bothers me') have all resulted too scarce to conduct an in-depth study and have been excluded from the following empirical analyses.

4.2   *Information focus vs. contrastive focus*

While, in many languages, cleft sentences have been often claimed to mainly express *contrastive* focus (e.g. Drubig 2003 for English clefts), we know that the common functional range of clefts is broader in French: Being the prototypical structural expression of narrow focus, clefts cover both contrastive and non-contrastive (information) focus, i.e. [±contrast] focus. Our results show the following (see **Error! Reference source not found.**): Out of 342 critical items (*c'est* and *il y a* clefts; monoclausal *c'est* and *il y a* sentences), 5 have been tagged as doubt cases and excluded from the analysis. Of the remaining 337 sentences: (i) The majority are cases of information focus, i.e. [-contrast] (see example (6) at p3). This is unsurprising given that in most situative contexts [+contrast] focus, being a more marked pragmatic interpretation, is scarcer. (ii) When [+contrast] is expressed (see example (7) at p3), speakers mostly utter a full cleft, and not a monoclausal utterance: Only 15% of [+contrast] focus is expressed in monoclausal forms. Contrastive focus requires additional semantic operations. Regarding the corrective variant of [+contrast] focus, the speaker needs to signal to the hearer to "go back" to information previously proposed to be added to the Common Ground, in order to assign a new value to the corresponding variable. In line with Zimmermann (2008), we assume that the full expression of the relative clause in this case facilitates processing and makes the communicative intention less prone to misunderstanding if the background into which the corrected value needs to be integrated is uttered overtly by means of a full cleft structure. (iii) [+contrast] focus is rarely expressed in *il y a* clefts. Given that *il y a* clefts mostly express a non-exhaustive focus interpretation (see Karssenberg 2018) and that [+contrast] is a sub-type of exhaustive focus (see Kiss 1998),[5] it is unsurprising that *c'est* clefts, by virtue of their pragmatic association with exhaustivity, are more frequent in the case of [+contrast].

---

[5]   A different view is, however, suggested by Cruschina (2021), who argues that exhaustive focus would in fact be a sub-type of contrastive focus, rather than the contrary.

Figure 1. Cleft type and IS-functions
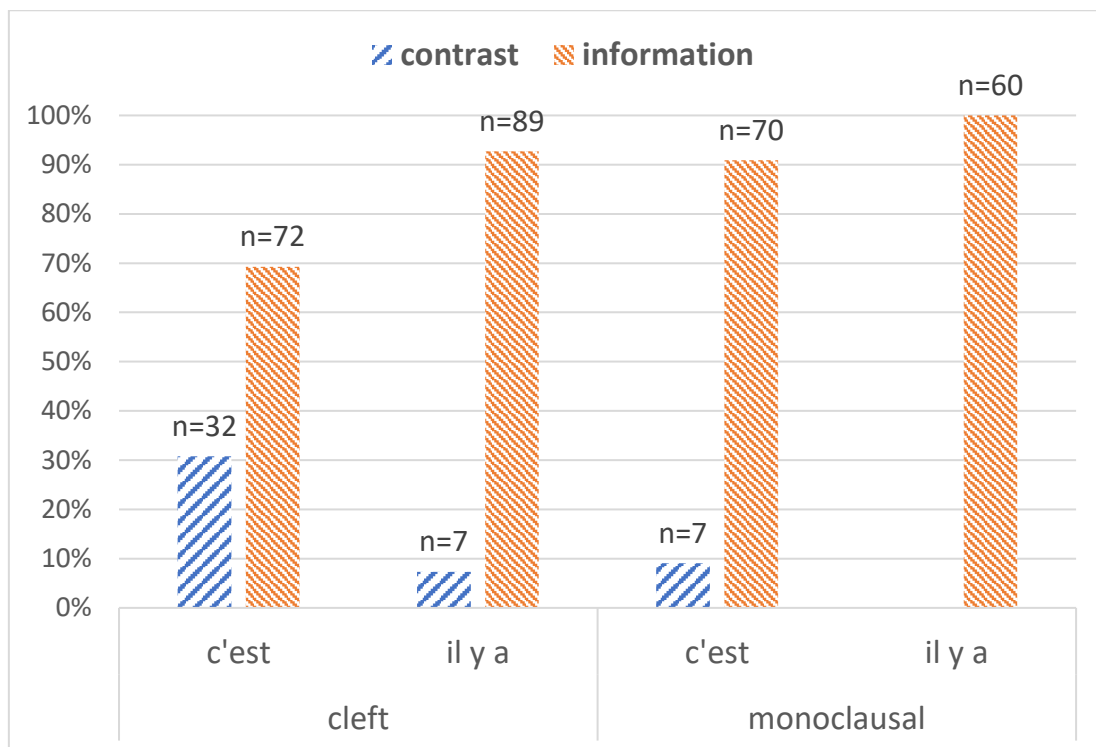
### 4.3   Distance and explicitness of the Question under Discussion

Recall that we hypothesized in section 2.4 that speakers are more likely to use a monoclausal utterance when the QUD is highly activated, while they are more likely to use a cleft when the QUD is not highly activated. The data (see **Error! Reference source not found.**) indeed confirm this hypothesis. Out of 342 critical items:

(i) Monoclausal *c'est* and *il y a* utterances appear much more frequently as answers to immediate QUDs than full *c'est* and *il y a* clefts, which most often answer either implicit or non-immediate QUDs (see Cassarà 2021, for full details and a comparative analysis with Spanish). This effect is particularly salient with monoclausal *c'est* utterances, 79% of them immediately following the QUD. The immediacy of the QUD means that the background information, i.e. the non-*wh* part in the case of *wh*-questions and all tokens in the case of polar questions, is highly activated. As already predicted by Belletti (2009), it is unsurprising that speakers tend not to repeat the background information in the immediately following assertion, thereby avoiding redundancy. In cleft sentences, background information is expressed by the relative clause. At the same time, the monoclausal construction is highly correlated with the degree of activation of the background information. According to us, this observation suggests that the hypothesis according to which monoclausal constructions are clefts with an elided relative clause is worth to be further scrutinized in future research.

(ii) We observe an interesting difference between full *c'est* and *il y a* clefts: Full *c'est* clefts are preferred when the QUD is implicit (50%), while full *il y a* are preferred with explicit, yet non-immediate QUDs (47%).

(iii) When the QUD is not immediately and explicitly uttered, a full biclausal cleft is preferred. In implicit QUDs, the background proposition is never clearly formulated, but rather evoked by other contextual cues, and its degree of activation is therefore not as high as it would be in the case of an explicit and immediate QUD. Similarly, in not-immediate QUDs, the degree of activation of the background proposition decreases due to the presence of intervening utterances. Therefore, speakers seem to prefer the use of full clefts, where the whole background proposition is explicitly uttered.

Figure 2. Cleft type and immediacy of the QUD
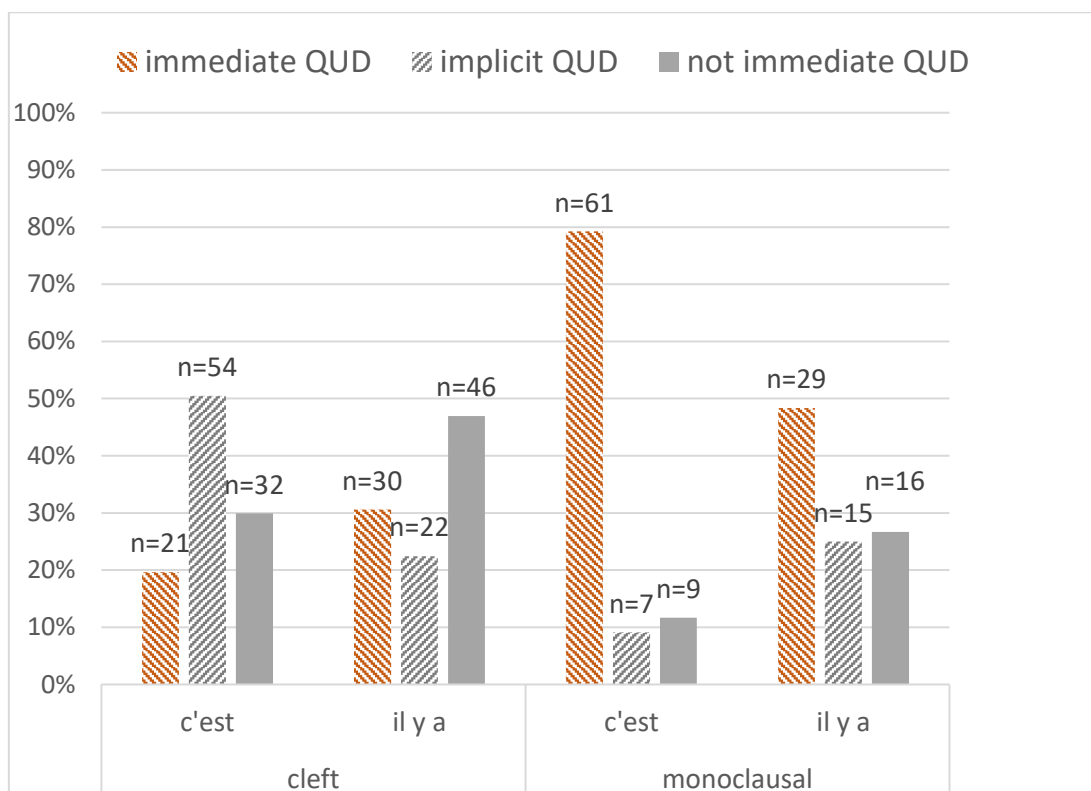
5   Modality of the QUD

As it can be observed by various examples reported in this paper, both clefts and monoclausal sentences can be preceded by questions of different types, i.e. QUD exhibiting a different modality. *Wh*-questions and yes/no questions are the most common types. At first sight, this may seem surprising, because assertive cleft-sentences are always the logical answer to a *wh*-question. However, it is well known

that in spontaneous speech many explicit yes-no questions contain *wh*-variable that is not overtly uttered (Busemann 1982). Consider the example in (24):

(24)   Q1: Did you see anybody entering the building?
        Q2: (Implicit QUD: Who did you see entering the building?)
        A1: Yes. I saw two men entering.

   In (24), it is possible to observe that the sentence uttered by speaker A constitutes an answer to two questions: the polarity item *yes* answers the explicit QUD, while the sentence "I saw two men entering" answers the implicit *wh*-question contained in Q2.[6] Since *wh*-questions and yes/no questions are of a different nature, we expect these differences to be reflected in the kind of answers they trigger. By virtue of containing implicit material, we consider yes/no questions more costly to accommodate for the speaker compared to *wh*-questions, where the *wh*-variable is explicitly uttered.
Given that the results in section 0 revealed the importance of the QUD for the choice of the cleft variant, we decided to further disentangle the various facets of the QUD in a follow-up analysis, turning our attention to the interrogative modality of the QUD.
        As already argued by Zimmermann (2008), speakers tend to utter more material when they are facing more costly (or unexpected) communicative operations, such as contrast. Therefore, we would expect yes/no QUDs, which - as we have seen - contain an implicit *wh*-QUD, to trigger more frequently full clefts, while we expect *wh*-QUDs to be followed more frequently by monoclausal utterances. For obvious reasons, we have only included explicit, both immediate and non-immediate, QUDs into this follow-up analysis. Example (25) illustrates a cleft sentence preceded by a *wh*-QUD, and (26) illustrates a monoclausal sentence preceded by a yes/no QUD.

(25)   A: *En fait, il y   avait qui   dans l'   immeuble quand ça s' est passé?*
            in fact there was who in   the.building   when it *se*.is happened
            'Actually, who was in the building when all this happened?'

        B: ***Alors en fait, il y   avait la   grand-mère du   dessus   qui était là.***
            so   in fact there was   the grand-mother of-the upstairs who was there
            'So, actually the old lady from upstairs was there.'

                                    (*Il y a* cleft, taken from *sgs* French, 9/261-262)

---

[6]   Although containing implicit material, yes/no questions are somewhat different from the implicit QUDs discussed in section 4.3. The implicit element is *wh*-variable. However, the background proposition is still uttered and, hence, highly activated.

(26)   A: *Et   sur son palier, il y   a   des   voisins?*

      and on his floor   there is some neighbours?

      'And in his floor, are there any neighbours?'

    B: ***Ouais, il y   a   un jeune couple.***

    yes   there is a   young couple

    'Yes, there is a young couple.'

(Monoclausal *il y a*, taken from *sgs* French, 37/75-76)

Error! Reference source not found. shows the distribution of the cleft and monoclausal variants, with separate bars for *wh*-QUDs and yes-no QUDs. This part of the analysis builds on 245 items, which represent the total number of answers to explicit QUDs.

Figure 3. Cleft type and modality of the explicit QUD
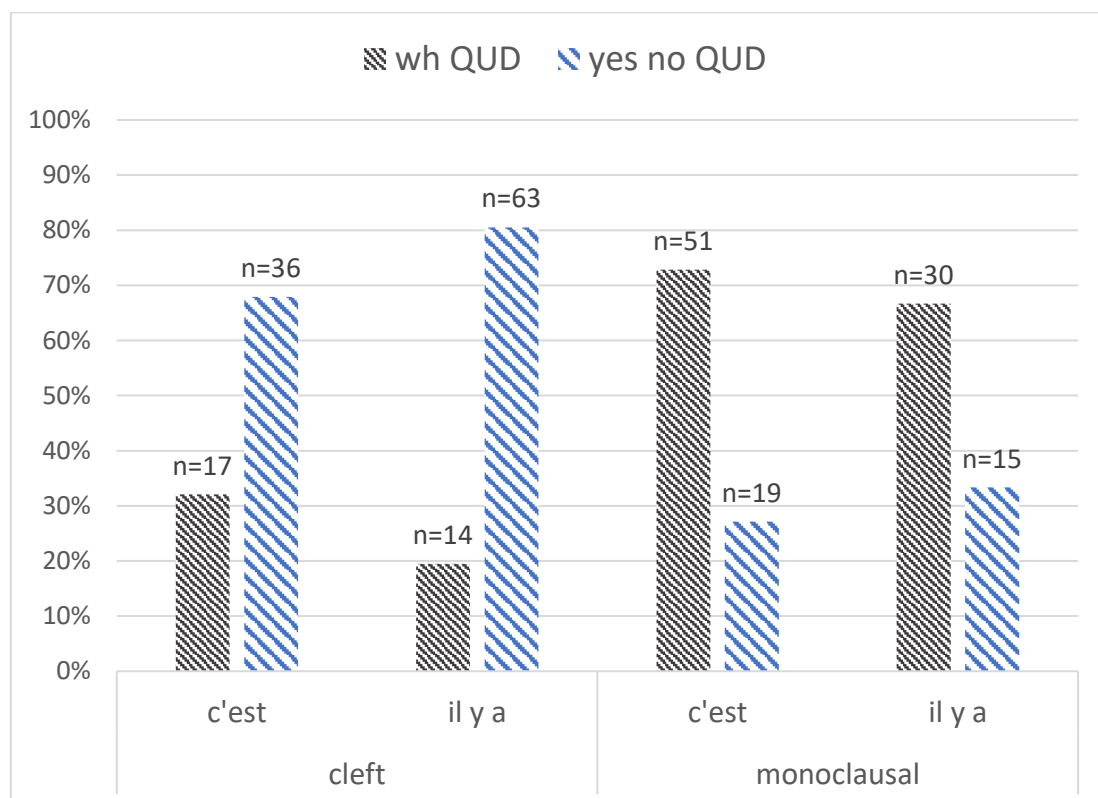
We do not observe a notable difference between the *c'est* and *il y a* type of monoclausal utterances. However, we observe a difference with full clefts: While the proportion of *wh* QUDs and yes/no QUDs is 1:2 with full *c'est* clefts, it is approximately 1:4 with *il y a* clefts.

Furthermore, the QUD-modality reveals a clear difference between full clefts and monoclausal utterances. Full clefts, whether they be of the *c'est* or the *il y a* type, are mostly triggered by yes/no-questions, whereas the opposite is true for monoclausal utterances, which most often occur as answers to *wh*-questions. We have seen in section 0 that yes-no QUDs of assertive clefts must entail an implicit *wh*-QUD. As can be seen in (27), the explicit QUD uttered by A is answered by the polarity item contained in the first part of B's answer (*ouais* 'yes'). The *il y a* cleft, on the other hand, constitutes an answer to the implicit QUD 'who came yesterday?'.

Conversely, when the QUD is an explicit *wh*-question, a clear correlation can be observed with monoclausal utterances (both of the *c'est* and *il y a* type), as in the case of (28).

(27)   A: *Tu dis qu' il y a des gens qui sont venus hier?*
         you say that there is some people that are come yesterday?
         'You say that there are people that came yesterday?'

                                        (Implicit QUD = Who came yesterday?)

        B: ***Ouais, il y a quatre gars qui sont passés hier.***
         yes there is four guys who are passed yesterday
         'Yes, four guys came yesterday.'

                                        (*Il y a* cleft, taken from *sgs* French, 83/231-232)

(28)   A: *Donc, tu sais qui a retrouvé le corps?*
         so you know who has found the body
         'So, do you know who found the body?'

        B: ***Oui, c'est sa femme.***
         yes it.is his wife
         'Yes, his wife.'

                                (Monoclausal c'est, taken from *sgs* French, 05/64-65)

The question is why the more complex QUD structure of yes/no-QUDs triggers more frequently a cleft sentence in its full form. Interestingly, yes/no QUDs behave like implicit only QUDs analyzed in section 0: Both trigger mostly full cleft variants. The determining factor is implicitness, i.e. the fact that the cleft answers to an implicit *wh* QUD. While speakers tend to avoid too many redundancies when unnecessary, the reverse also holds, namely avoiding too much elided and implicit material. Speakers tend to avoid that an implicit *wh*-QUD is followed by a cleft variant in which the background material is also elided, as the results in section 4.3 show. Similarly, they tend to avoid that the implicit *wh*-element contained in yes/no questions is followed by another elliptical structure (e.g. a monoclausal utterance). Thus, there seems to be

a general tendency to avoid both redundancy and implicitness to achieve successful communication.

6   Towards a discourse model of cleft sentences

The results reported in this paper have revealed new insights on the behavior of clefts and monoclausal variants when they compete to express the same pragmatic function of subject focus. Firstly, the dialogues of the *sgs* corpus have confirmed that *c'est* clefts and *il y a* clefts are the most frequent cleft types in spontaneous speech, and that elliptical structures, such as monoclausal variants are far from being uncommon. Second, it has been shown that three factors determine the choice of the different variants: i) the presence or absence of contrast, ii) the immediacy of the QUD (or degree of activation of the background proposition) of the backgrounded material, and iii) the modality of the QUD.

From a pragmatic perspective, the findings regarding the immediacy and the implicitness of the QUD are in line with the "Principle of the Least Effort" proposed by Zipf (1949), according to which speakers prefer to express less when they have enough contextual cues at their disposal, while they express more when these cues are not as clear. Thus, the use of a full cleft in the case of an immediately-preceding QUD would provide more information than required by the context, given that the background proposition contained in the QUD is highly activated and does not need to be repeated by the relative clause of a full cleft. The results reported in this paper confirm that in those cases, speakers opt for a more economical structure, i.e. a monoclausal utterance, and that the use of full clefts is preferred in contexts where the activation of the background proposition decreases. For instance, when the QUD is not immediate, the presence of intervening utterances may diminish the saliency of the background material by diverting the attention of the speaker towards a somewhat related but still different sub-QUD. Additionally, it seems that, when speakers have to choose between a more or less economic variant, they tend to do so by respecting cooperative pragmatic principles, such as the ones advocated by Grice (1975). For instance, the Maxim of Manner describes how people achieve effective conversational goals by avoiding unnecessary prolixity and redundancy. In this paper, we have shown that such principles can be applied to different answering strategies when they express narrow focus.

To investigate these strategies, we adopted a variationist approach, according to which a variable is defined as a general or abstract feature that is subject to variation, while the actual instantiations of the variable in speech are known as the variants. In this study, we chose a pragmatic function as our linguistic variable (subject focus), and identified possible competing variants to express this function, in line with the principle of accountability (Labov 1982: 30). Such an approach has the advantage to capture the phenomenon in its entirety, thus avoiding the exclusion of certain syntactic structures that would otherwise not being compared. According to this perspective,

there are reasons to consider both full and monoclausal forms as variants of the same variable, based on the fact that they effectively function as alternates for expressing subject focus in spontaneous speech – a phenomenon called "neutralization in discourse" (Sankoff 1988: 153). Similarly, looking at the same pragmatic contexts from a cross-linguistic perspective might reveal that different languages use syntactically-different -but nevertheless pragmatically-equivalent strategies- to convey the same communicative function. For instance, Cassarà (2022) has shown that, *ceteris paribus*, Spanish speakers have the tendency to answer a question on the subject (e.g. *Quién abrió la puerta?* 'Who opened the door?') with a fragment (*Juan* 'Juan'), rather than a monoclausal utterance (e.g. *Ha sido Juan* 'It was Juan'), as is the case in French. Intuitively, this phenomenon could also be observed in languages like English or German. It would be interesting, therefore, to adopt this variationist comparative perspective on both related and unrelated languages to point out the differences, and to try to understand why they occur.

When one draws on the variationist perspective as we have just suggested, the question of change in progress arises, in particular whether *c'est* and *il y a* clefts are becoming increasingly interchangeable. The same change hypothesis arises with monoclausal and biclausal forms. Yet, we leave these issues to future research. An age-stratified sample that allows to compare the usage of these variants by different age groups could provide evidence in favor or against the assumption of such change in progress. Apart from the question of inter-individual variation between age groups, the question about intra-individual variation of clefts also needs further inquiry. Dufter's (2008) or Destruel's (2016) study on the difference between written and spoken French is an indicator that register is also a dimension of variation. Future research in which we compare different situative settings within the spoken and written modality would allow us to tackle this issue as well.

Abbreviations

IS   Information Structure
FOC   Focus
Q   question particle/marker
QUD   Question Under Discussion

References

Adli, Aria. 2011. Gradient Acceptability and Frequency Effects in Information Structure: a quantitative study on Spanish, Catalan, and Persian. Habilitationsschrift, Universität Freiburg.

Beaver, David & Clark, Brady Z. 2008. Sense and Sensitivity: How focus determines meaning. Oxford: Wiley-Blackwell.

Belletti, Adriana. 2009. Answering strategies: New information subjects and the nature of clefts. Structures and Strategies. 242–265. London: Routledge.

Belletti, Adriana. 2015. The Focus map of clefts: Extraposition and Predication. In Shlonsky, U. (ed.), Beyond Functional Sequence, The cartography of Syntactic Structures V. 10. 42–59. Oxford & New York: Oxford University Press.

Birner, Betty J., Kaplan, Jeffrey P. & Ward, Gregory. 2007. Functional compositionality and the interaction of discourse constraints. Language 83(2): 317–343.

Blanche-Benveniste, Claire. 2006. Les clivées françaises de type: C'est comme ça que, c'est pour ça que, cest là que tout a commencé. Moderna Språk 100(2): 273–287.

Büring, Daniel. 1998. Identity, modality, and the candidate behind the wall. In Strolovitch, D. & Lawson, A. (eds.), Proceedings of SALT 8. 36–54. Ithaca, NY: Cornell University.

Büring, Daniel. 2003. On D-trees, beans, and B-accents. Linguistics and Philosophy 23: 511–545.

Büring, Daniel. 2008. What's new (and what's given) in the theory of focus? Proceedings of the 34th Annual Meeting of the Berkeley Linguistics Society, Berkeley. 403–424.

Busemann, Stephan. 1982. Problems involving the automatic generation of utterances in German. Memo ANS-8. Research Unit for Information Science and AI, University of Hamburg.

Carter-Thomas, Shirley. 2009. The French c'est-cleft: Function and frequency. In Banks, D., Eason, S. & Ormrod, J. (eds.), La linguistique systémique fonctionnelle et la langue française. 127–157. Paris: L'Harmattan. <halshs-00276942>.

Cassarà, Alessia. 2021. Subject Focus in Spanish and French: a comparative corpus analysis. PhD dissertation, Universität zu Köln.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects and topics. In Li, C. N. (ed.), Subject and topic. 25–55. New York: Academic Press.

Clech-Darbon, Anne, Rebuschi, Georges & Rialland, Annie. 1999. Are there cleft sentences in French? In Tuller, L. & Rebuschi, G. (eds.), The grammar of focus. 83–118. Amsterdam: John Benjamins.

Cruschina, Silvio. 2021. The greater the contrast, the greater the potential: On the effects of focus in syntax. Glossa: a journal of general linguistics 6(1): 1–30.

Culioli, Antoine. 1990. Pour une linguistique de l'énonciation. Paris: Ophrys.

Davidse, Kristin, Kimps, Ditte & Njende, Ngum Meyuhnsi. 2016. Making the case for defining clefts broadly as secondary specification structures: Prosody and information

structure. Paper presented at 'The second Round Table on Communicative Dynamism', Namur.

Destruel, Emilie. 2013. The French c'est-cleft: Empirical studies of its meaning and use. PhD dissertation, University of Texas at Austin.

Destruel, Emilie. 2016. Focus marking asymmetries in Colloquial and Standard French: A stochastic optimality-theoretic account. Journal of French Language Studies 26(3): 299–326.

Doetjes, Jenny, Rebuschi, Georges & Rialland, Annie. 2004. Cleft Sentences. In Corblin, F. & De Swart, H. (eds.), Handbook of French semantics. 529–552. Stanford: CSLI Publications.

Drubig, Bernhard H. 2003. Towards a typology of focus and focus constructions. Linguistics 41: 1–50.

Druetta, Ruggero. 2018. Syntaxe de l'interrogation en français et clivage écrit-oral: une description impossible? In Béguelin, M.-J., Coveney, A. & Guryev, A. (eds.), L'interrogative en français. 19–50. Bern & Berlin: Peter Lang.

Dufter, Andreas. 2006. Zwischen Kompositionalität und Konventionalisierung: Satzspaltung mit 'c'est' im Französischen der Gegenwart. Romanistisches Jahrbuch 57: 31–59.

Dufter, Andreas. 2008. On explaining the rise of c'est-clefts in French. In Detges, U. & Waltereit, R. (eds.), The Paradox of Grammatical Change: Perspectives from Romance. 31–56. Amsterdam: John Benjamins.

Dufter, Andreas. 2009a. Beyond focus marking: Fine-tuning the evolution of cleft types from Latin to Modern French. Paper presented at 'Paper presented at the DGfS-Jahrestagung 31. AG 9: Focus marking strategies and focus interpretation', Osnabrück.

Dufter, Andreas. 2009b. Clefting and discourse organization. Comparing Germanic and Romance. In Dufter, A. & Jacob, D. (eds.), Focus and Background in Romance Languages. 83–121. Amsterdam & Philadelphia: Benjamins.

Grice, Herbert Paul. 1975. Logic and Conversation. In Cole, P. & Morgan, J. L. (eds.), Syntax and Semantics. Vol. 3: Speech Acts. 41–58. New York: Academic Press.

Gussenhoven, Carlos. 2008. Types of Focus in English. In Lee, C., Gordon, M. & Büring, D. (eds.), Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation. 83–100. Dordrecht: Springer Netherlands.

Hedberg, Nancy. 2000. The Referential Status of Clefts. Language 76(4): 891–920.

Higgins, Francis Roger. 1979. The pseudo-cleft construction in English. New York: Garland.

Huber, Stefan. 2006. The complex functions of it-clefts. In Molnár, V. & Winkler, S. (eds.), The architecture of focus. 549–578. Berlin/New York: Mouton de Gruyter.

Jacob, Daniel. 2015. Anaphorische Spaltsätze im Französischen: Grammatik - Text - Rhetorik. In Adam, S., Jacob, D. & Schecker, M. (eds.), Informationsstrukturen in Kontrast: Strukturen, Kompositionen und Strategien. 101–122. Frankfurt am Main: Peter Lang.

Jenkins, Lyle. 1973. Cleft reduction. In Rohrer, C. & Ruwet, N. (eds.), Etudes de syntaxe. 182–191. Tübingen: Max Niemeyer.

Karssenberg, Lena. 2017a. Specificerende il y a-clefts (er is-clefts). Handelingen der Koninklijke Zuid-Nederlandse Maatschappij voor Taal- en Letterkunde en Geschiedenis (KZM): 77–91.

Karssenberg, Lena. 2017b. Ya les oiseaux qui chantent. A corpus analysis of French il y a clefts. PhD dissertation, KU Leuven.

Karssenberg, Lena. 2018. Non-prototypical clefts in French: A corpus analysis of il y a clefts. Berlin: De Gruyter Mouton.

Karssenberg, Lena & Lahousse, Karen. 2018. The information structure of French il y a clefts and c'est clefts: A corpus-based analysis. Linguistics 56(3): 513–548.

Katz, Stacey. 2000a. Categories of C'est-Cleft Constructions. Canadian Journal of Linguistics / Revue Canadienne de Linguistique 45(3/4): 253–273.

Katz, Stacy. 2000b. A functional approach to the teaching of the French c'est-cleft. The French Review 74(2): 248–262.

Kiss, Katalin É. 1998. Identificational Focus versus Information Focus. Language 74(2): 245–273.

Krifka, Manfred. 2007. Basic Notions of Information Structure. In Féry, C., Fanselow, G. & Krifka, M. (eds.), Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS) 6. 13–56. Potsdam: Universitätsverlag Potsdam.

Labov, William. 1982. Building on empirical foundations. In Lehmann, W. P. & Malkiel, Y. (eds.), Perspectives on Historical Linguistics. 17–92. Amsterdam & Philadelphia: John Benjamins.

Lahousse, Karen & Borremans, Marijke. 2014. The distribution of functional-pragmatic types of clefts in adverbial clauses. Linguistics 52(3): 793–836.

Lambrecht, Knud. 1988. Presentational cleft constructions in spoken French. In Haiman, J. & Thompson, S. A. (eds.), Clause combining in grammar and discourse. 135–179. Amsterdam: John Benjamins.

Lambrecht, Knud. 2001a. Dislocation. In Haspelmath, M., König, E., Oesterreicher, W. & Raible, W. (eds.), Sprachtypologie und sprachliche Universalien. Ein internationales Handbuch. 1050–1079. Berlin: Walter de Gruyter.

Lambrecht, Knud. 2001b. A framework for the analysis of cleft constructions. Linguistics 39(3): 463–516.

Mertens, Piet. 2008. Syntaxe, prosodie et structure informationnelle: une approche prédictive pour l'analyse de l'intonation dans le discours. Travaux de linguistique 56: 87–124.

Mertens, Piet. 2012. La prosodie des clivées. In Caddéo, S., Roubaud, M.-N., Rouquier, M. & Sabio, F. (eds.), Penser les langues avec Claire Blanche-Benveniste. 127–139. Aix-en-Provence: Presses Universitaires de Provence.

Paul, Hermann. 1880. Prinzipien der Sprachgeschichte. Halle: Niemeyer.

Repp, Sophie. 2010. Defining 'contrast' as an information-structural notion in grammar. Lingua 120(6): 1333–1345.

Rialland, Annie, Rebuschi, Georges & Doetjes, Jenny. 2002. What is focused in C'est XP qui/que cleft sentences in French? In Bel, B. & Marlien, I. (eds.), Speech Prosody. 595–598. Aix-en-Provence: ISCA Archive.

Riester, Arndt, Brunetti, Lisa & De Kuthy, Kordula. 2018. Annotation guidelines for Questions under Discussion and information structure. In Adamou, E., Haude, K. & Vanhove, M. (eds.), Information Structure in Lesser-Described Languages: Studies in Syntax and Prosody. 403–443. Amsterdam: Benjamins.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In Hak Yoon, J. & Kathol, A. (eds.), OSU Working Papers in Linguistics 49: Papers in Semantics. 91–136. Columbus: Ohio State University.

Rooth, Mats. 1992. A theory of focus interpretation. Natural Language Semantics 1: 75–116.

Roy, Isabelle. 2013. Nonverbal Predication: Copular Sentences at the Syntax-Semantics Interface. Oxford: Oxford University Press.

Sankoff, David. 1988. Sociolinguistics and syntactic variation. In Newmeyer, F. J. (ed.), Linguistics: the Cambridge Survey. Vol IV: The Socio-cultural Context. 140–161. Cambridge: Cambridge University Press.

Scappini, Sophie-Anne. 2006. Etude du dispositif d'extraction en 'c'est...qu', différenciation entre une relative en 'c'est...qu' et une proposition clivée. PhD dissertation, Université d'Aix-en-Provence.

Schwarzschild, Roger. 1999. GIVENness, AVOIDF and Other Constraints on the Placement of Accent. Natural Language Semantics 7(2): 141–177.

Skopeteas, Stavros & Fanselow, Gisbert. 2010. Focus types and argument asymmetries: A cross-linguistic study in language production. Comparative and Contrastive Studies of Information Structure. 169–197. John Benjamins.

Smits, Reinier Johannes Charles. 1989. Eurogrammar. The relative and cleft constructions in the Germanic and Romance languages. Dordrecht: Foris Publications.

Steube, Anita. 2001. Correction by contrastive focus. ZAS Papers in Linguistics 27(2-3): 211–230.

von Stutterheim, Christiane & Klein, Wolfgang. 1989. Referential Movement in Descriptive and Narrative Discourse. In Dietrich, R. & Graumann, C. F. (eds.), North-Holland Linguistic Series: Linguistic Variations. 39–76. Elsevier.

van Kuppevelt, Jan. 1995. Discourse structure, topicality and questioning. Journal of Linguistics 31(1): 109–147.

van Leusen, Noor. 2004. Incompatibility in context: A diagnosis of correction. Journal of Semantics - J SEMANT 21: 415–415.

Verwimp, Lyan & Lahousse, Karen. 2017. Definite il y a--clefts in spoken French. Journal of French Language Studies 27(3): 263–290.

Zimmermann, Malte. 2008. Contrastive Focus and Emphasis. Acta Linguistica Hungarica 55: 347–360.

Zipf, George Kingsley. 1949. Human Behavior and the Principle of Least Effort. Cambridge: Addison-Wesley Press.