

# Information structure in Romance: from sentence to discourse

*This work is dedicated to the memory of Manuel Leonetti Jungl.*

# Index

|   |    |
|---|----|
| Foreword .....  | 3  |
| Part 1 - Information structure in Romance .....                             | 4  |
| 1 Introduction .....  | 4  |
| 1.1 Information structure in Romance .....                                  | 6  |
| 1.2 Information structure: methodology .....                                | 9  |
| 2 A QUD -based approach to information structure .....                      | 10 |
| 2.1 Introduction .....  | 10 |
| 2.2 QUD-reconstruction .....  | 11 |
| 2.2.1 Parallelism .....   | 13 |
| 2.3 Not-at-issue content and the information structure of adjuncts .....    | 14 |
| 2.3.1 The information structure of adjuncts .....                           | 16 |
| 2.3.2 Back to not-at-issue content .....                                    | 21 |
| 2.4 Some final remarks on the QUD-model .....                               | 23 |
| 3 Focus and focus fronting .....  | 23 |
| 3.1 Introduction .....  | 23 |
| 3.2 Focus fronting in Italian and Spanish .....                             | 24 |
| 4 Topic and (Clitic) Left Dislocation .....                                 | 29 |
| 4.1 Introduction .....  | 29 |
| 4.2 (Clitic) Left Dislocation .....   | 32 |
| 5 Contrast .....  | 37 |
| 5.1 Introduction .....  | 37 |
| 5.2 Contrastive relations .....   | 38 |
| 6 Future developments on IS modeling and the QUD-based model .....          | 44 |
| Part 2 Non-canonical questions .....  | 49 |
| 1 Introduction .....  | 49 |
| 2 <i>Comment</i> interrogatives in French .....                             | 50 |
| 2.1 The prosody of <i>comment</i> interrogatives .....                      | 52 |
| 3 French <i>comment ça</i> and Spanish <i>cómo que</i> interrogatives ..... | 55 |
| 4 Future developments on ‘reason’ interrogatives .....                      | 59 |
| References .....  | 62 |

# Foreword

The present work is not a detailed summary of all my past research; furthermore, the articles that I have chosen are not all summarized with the same level of details. There are various reasons for the choice of articles that I have made and for the way I discuss them here.

As for the older papers, which were published before my appointment at the Université Paris Diderot (now Université Paris Cité), I have decided to only summarize those that discuss phenomena whose analysis could be revised and bettered by my more recent discourse-oriented take on information structure (IS), based on the notion of ‘question under discussion’ (QUD) (Riester, Brunetti, and De Kuthy 2018). In Part 1 of this document, I therefore present a reworked version of my analysis of focus fronting (Brunetti 2009a) and of (clitic) left dislocation (Brunetti 2009c; 2009b) in Romance. I also dwell upon my recent work on adjuncts (Brunetti, De Kuthy, and Riester 2021) and on contrast (Brunetti 2024) within the same QUD-based model. The former article tackles a phenomenon that has not received much attention and that I think sheds some light on interesting aspects of the discourse-IS mapping. As for the article on contrast, it gives a – perhaps incomplete – idea of my more recent view on a phenomenon that has been the object of much of my past work; it therefore makes a link between my past and present research.

In Part 2, I summarize and re-discuss a set of phenomena – the non-canonical pragmatic functions of a certain type of interrogatives in Romance – that constitutes the object of great part of my most recent research (Brunetti et al. 2021; Brunetti, Tovená, and Yoo 2022; Fliessbach, Brunetti, and Yoo under review; Thiberge, Brunetti, and Fliessbach submitted) and that is therefore important to present here.

At the end of both Part 1 and Part 2, I present some of my ongoing work on the respective domains – information structure and non-canonical questions – and sketch my plans for future research. Note finally, that in the section of the general document to be submitted for the obtention of the *Habilitation à diriger des recherches*, where it is asked to present five main activities as a principal investigator, more publications of mine are summarized, which were not included or only briefly mentioned in the present summary.

Note on some editorial choices:

- I use by default the pronoun ‘she’ for the speaker and ‘he’ for the addressee.
- I follow the Leipzig’s glossing rules (<https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>), but add fully detailed glosses only when I find it necessary for the purposes of the point that I am making, and use English-word translations otherwise.
- The disfluencies of the examples extracted from spontaneous spoken conversations have sometimes been reduced or omitted for the sake of clarity.

# Part 1- Information structure in Romance

## 1 Introduction

Information Structure (IS) has a long research tradition and has been tackled from many different theoretical and empirical angles, generally strongly influenced by the languages under study: from the first works of the Prague School (by scholars such as Mathesius and Firbas), who focused on word order variation in Slavic languages, to Lambrecht (1994), who looked at word order and intonation in Romance and Germanic; from Vallduví's (1992) tripartite framework inspired by the way IS categories are expressed in Catalan syntax, to the Alternative Semantics framework (Rooth 1992; Büring 1997) and other frameworks developed in the English speaking community, which are inspired by English intonational phenomena (Halliday 1967; Jackendoff 1972; Steedman 2000, among many others). More recent studies – mostly on Germanic and Romance languages – address IS from a discourse perspective, and explain it by appealing to the discourse notion of 'Question Under Discussion' (QUD): from (Roberts 1996; 2012; Kuppevelt 1995; Ginzburg 1995; Büring 2003), from Roberts (1996; 2012), Kuppevelt (1995), Ginzburg (1995), Büring (2003) until more recent work such as Onea (2016), Riester, Brunetti and De Kuthy (2018) and others (for a review, see Velleman and Beaver 2016).

Despite such a long tradition, only in the last decades information structure has become an ineludible step of linguistic analysis, which must systematically be taken into account to explain linguistic phenomena, even beyond those specific syntactic or prosodic constructions that have very early been linked to it, such as dislocations, clefts, subject inversion on the syntactic side, F0 movements on the prosodic side. When a syntactic phenomenon is studied, an explanation that is partially or fully based on the informational properties of the sentence must be considered, whether or not this means to suppose a syntactic encoding of informational categories, as it is done in chomskian generative approaches. Informational notions have been proven to be crucial to account for phenomena at the sentence level that scholars were not able to fully explain in purely syntactic nor in semantic terms, among which ellipsis or fragments, and even old syntactic puzzles like islands (see Liu et al. 2022 for a review of the literature on IS explanations of island constraints).

Recognizing the importance of IS in linguistic analysis means to recognize the importance of context and language use to understand linguistic phenomena. In the generative, chomskian tradition, this has been seen as a potential problem, since linguistic analysis should account for the speaker's "competence", which is independent from language use. The problem has been tackled within this framework by proposing that pragmatic functions are incorporated in the syntactic tree via dedicated syntactic positions of focus and topic (among the first, see Brody 1990; Rizzi 1997; 2010; Belletti 2004). In semantics, a focus feature (F-feature) eventually visible to syntax has been proposed to account for the semantic effects of focus on the propositional content of the sentence (Rooth 1992). Focus and topic syntactic features are also present in lexicalist generative frameworks such as LFG, see for instance Butt (2014).

The problem with attempts at attributing a unique syntactic position to focus or topic is that it is rarely the case that such position corresponds to a *unique* informational notion. As we will see below, IS partitions (focus-background, topic-comment, contrast) are rather favored (sometimes strongly) than inherently determined by certain positions. Furthermore, in most languages, information structure is expressed by both morphosyntactic and intonational means: a morphosyntactic structure alone is generally insufficient to predict the informational partition of a sentence. At the same time, quite an important degree of indeterminacy and ambiguity is possible: information structure is not always clearly identified by form (be it morphosyntax or prosody) and context needs to be evoked for disambiguation, (see Riester, Brunetti, and De Kuthy 2018 for a discussion).<sup>1</sup>

Going back to the role of information structure in linguistic analysis, if its effects on syntax have been overlooked, its effects on the intonational properties of a sentence have been early acknowledged. That is perhaps due to the fact that, while not all languages use accents (it is argued for instance that French uses phrase boundaries, see Rossi 1980, Vaissière 1990, Delais-Roussarie et al. 2015), it is though rare to find a language that does not use prosody at all to mark focus, topic, or contrast.<sup>2</sup> I will return to prosody and intonation below, when presenting the characteristics of information structure in Romance.

Before moving on to discuss the way in which the interplay between information structure and syntax/prosody takes place in the Romance family, I would like to say a few more words on (the interpretive effects of) information structure. The terminology and also the definitions that are found in the literature on IS notions are numerous and often very different from each other. I believe that the classification below according to three dimensions proposed by Marandin (2006) gives credit to most of the accounts that are found in the literature, which may give preference and highlight one or the other of such dimensions:

(1) (adapted from Marandin 2006)

- a. The content of the utterance is divided into two parts, depending on whether it is *anchored in* the context or *added to* the context (and therefore, informative);
- b. A predicative relationship of ‘aboutness’ holds between a predicate and a predicator (logical subject, topic);
- c. The interpretation of a constituent triggers an inference that it belongs to a set of alternatives.

The first dimension divides the utterance into a ‘focus’ and a ‘background’ (see Dufter and Jacob 2009 for an overview of the literature on this partition). The focus corresponds to the information that is added to the context, and in contexts where the utterance answers a question, it is identified as the part of the utterance that provides its answer. In Riester, Brunetti, and De Kuthy (2018), along the lines of Roberts (1996; 2012) and Büring (2003) (and as we will see in details in Section 2), the same definition concerns assertions that are not preceded by a question, since the focus is identified as the part of the utterance that answers the *implicit* ‘question under discussion’ at that point in discourse. The notion of focus will be further discussed in Section 3.

The second dimension, which puts the notion of aboutness into play, accounts for the topic-comment partition of the utterance. A topic is the entity about which the information provided

---

<sup>1</sup> Other linguistic categories can be identified by more than one property, think for instance of the subject function, which can be recognized in many languages by its agreement with the verb as well as its position with respect to the verb. However, a grammatical function is a syntactic notion and no definition of the category beyond a syntactic one is available. In the case of information structure, a definition that goes beyond form can be given in terms of givenness, aboutness, or contrast, as it will become clearer below.

<sup>2</sup> A possible exception might be wolof, see Rialland and Robert (2004).

by the sentence is given. While the focus-background partition, as we will see, is well captured by the notion of ‘question under discussion’ (QUD), the notion of topic becomes less straightforward within a QUD-based approach to information structure and discourse. In Section 4, the analysis of (clitic) left dislocation in Romance will feed the debate that was started out by McNally (1998), about the necessity of a notion of sentence topic (an entity about which the information is entered) once a QUD framework of discourse is adopted.

The third dimension in (1) is the contrastive one, which orthogonally interacts with the first and the second. Understanding the nature of contrast and attempting at defining it by studying its linguistic marking (mostly) in Romance languages has been the main goal of my research on information structure. I will therefore come back to this notion repeatedly in the present work. In particular, in Section 5, I will show the limits of the definition in (1c) that is uniquely based on the notion of alternatives and, by taking inspiration from the work of Umbach (2004; 2005) and Repp (2016), will argue for the necessity to take the discourse dimension into account – both in terms of discourse relations among utterances and of discourse structure – in order to precisely characterize this phenomenon.

### 1.1 Information structure in Romance

Depending on the language and on the phenomenon at issue, information structure is involved at different degrees in the linguistic analysis and has an impact on different domains of grammar. As I said, the School of Prague was particularly interested in the effects of IS on word order, because their studies focused on Slavic languages, which display a particularly flexible and IS-dependent order of the constituents in the sentence. In English the role of word order – and syntax in general – for the identification of IS categories is marginal (though not completely absent), while intonation (accents, in particular) plays a crucial role. That is why the first important studies on focus and the first theoretical models of IS proposed within the English-speaking community aimed to explain the interpretation of certain accents (see Halliday 1967; Jackendoff 1972, and later Rooth 1992; Schwarzschild 1999; Steedman 2000).<sup>3</sup>

As far as Romance languages are concerned, studies like Vallduví and Engdahl (1996) have stressed out the importance of syntax in the expression of informational notions, compared to languages of other families such as the Germanic ones. For instance, in Vallduví and Engdahl’s (1996) examples below, while English allows for both (2b) and (3b) to express the contrastive topic *the knives*, Catalan (and the same could be said for other Romance languages) only allows for dislocation of the contrastive topic: the in-situ position in (4c) is degraded (small caps indicate the presence of an accent).

(2) Where can I find the cutlery? (Vallduví and Engdahl 1996, 473)

- a. The forks are in the CUPBOARD...
- b. but the knives I left in the DRAWER.

(3) Where can I find the cutlery?

- a. The forks are in the CUPBOARD...
- c. but I left the knives in the DRAWER.

(4) On són, els coberts?

‘Where’s the flatware?’

- a. Les forquilles són a l’armari, però...  
‘The forks are in the cupboard, but...’

---

<sup>3</sup> Though focus fronting too was an early object of study, see (Chomsky 1972).

- b. ...els ganivets<sub>1</sub> els<sub>1</sub> vaig ficar al CALAIX.  
the knives Cl.OBJ 1<sup>ST</sup>-PST put.INF in.the drawer
- c. ...#vaig ficar els ganivets al CALAIX.  
1<sup>ST</sup>-PST put.INF the knives in.the drawer  
‘...the knives, I left them in the drawer’

This constraint in Catalan supports Vallduví’s (1992) theoretical proposal to divide the background into ‘link’ and ‘tail’: indeed, the former would be represented in this language by a (Clitic) Left Dislocation (CILD), as shown in (4), while the latter by a (Clitic) Right Dislocation (CIRD).

Despite their similarities, Romance languages are sufficiently diversified to make their comparison extremely interesting, as it is often the case among typologically close languages. All Romance languages tend to be informationally transparent through word order, which allows for a clear bipartition of the sentence into focus and background, or topic and comment. Despite this apparently uniform crosslinguistic behavior, a fine-grained analysis shows some striking differences. As an example, Leonetti (2014) compares Romance languages with respect to the VSX order (where X corresponds to a verb complement or adjunct) and observes that such an order is accepted in Spanish, Portuguese, and Romanian, but rejected in Italian, Catalan, and French (Leonetti 2014, 38):

- |  |                            |
|--|----------------------------|
| (5) a. #Ha comprato Maria il giornale. | Italian                    |
| b. #Ha comprat la Maria el diari.      | Catalan                    |
| c. #A acheté Marie le journal.         | French                     |
| has bought Mary the newspaper          |                            |
| (6) a. Ha comprado María el periódico. | Spanish                    |
| has bought Mary the newspaper          |                            |
| b. Partiu o Paulo a janela.            | Portuguese                 |
| opened the Paul the window             |                            |
| c. Scrie Ion un articol.               | Romanian (from Soare 2009) |
| wrote John an article                  |                            |

The difference is explained by Leonetti by means of a generalization concerning the degree of transparency between syntax and IS within the Romance family. After showing that a VSX order in Romance must be interpreted as a “a single informational unit (typically, as wide focus, with athetic reading, and alternatively as background)”, he proposes that “the dividing line between VSX languages – Spanish, Portuguese, Romanian – and non-VSX languages – Catalan, Italian, French – can be based on the capacity of a language to admit complex constituents processed as single informational chunks, without internal partitions, especially in marked word orders such as subject inversion” (Leonetti 2014, 60). The theoretical conclusions that Leonetti draws from this comparison are the following:

The theoretical consequences of this view are all related with the place of IS inside the architecture of grammar. The major one is that IS, in line with a series of recent proposals, is viewed as an independent level of representation: it takes well-formed syntactic structures as input and evaluates them with respect to certain conditions. As a result, some structures that the syntax defines as grammatical may be rejected at the IS level and become (IS-)ungrammatical if they do not meet the requirements on informational partitions and focus structure. Notice that IS is not only responsible for issues of contextual adequacy, as usually assumed, but also for issues of idiomatic well-formedness. In languages like Catalan and Italian it must check whether a string obeys

the conditions on formal chunking that force dislocations and the marking of focus–background partitions, i.e. what makes these languages different from Spanish (and Romanian, and Greek). [...] The filtering role of IS thus accounts both for facts that were traditionally considered as syntactic and for facts that are usually seen as stylistic (a preference for right dislocation, or a tendency towards subject inversion with complex strings). (Leonetti 2014, 60)

Leonetti’s theoretical concern is to show that IS is independent from syntax and therefore to exclude an explanation of syntactic phenomena related to IS by means of the existence of dedicated syntactic projections for focus and topic categories (Focus Phrase, Topic Phrase, etc.). Leonetti’s concern is further compounded if one considers that intonation too plays an important role in Romance for the expression of information structure. Most Romance languages, in the same way as English does, use the nuclear accent to mark the focus and another accent to mark the topic. Contrast too is marked by prosody, though it is not always clear whether the marking is phonological or phonetic, since properties like higher pitch, greater intensity, longer duration than the corresponding non-contrastive parts have been identified in various languages (cf. Katz and Selkirk 2011). In French, prosodic phrasing rather than pitch accents seem to play a role in focus identification (Fery 2001).

Given that both syntax and prosody are responsible for the expression of IS in Romance, the question one may ask is how they interact with each other and what weight each of them has in the representation of IS. The answer to this question is not easy to give and I will only provide here a few examples that may help to reflect on the issue. The first example concerns French clefts. Clefts divide the sentence into two clearly distinguishable syntactic chunks, which typically correspond, in French and in other languages, to the focus (the clefted constituent) and the background (the coda) respectively. This is illustrated in (7): *le petit* is the focus because it answers the implicit question ‘Who fell down the stairs?’ and therefore represents the value of the variable *x* in *x est tombé dans l’escalier*. It also contrasts with the focal alternative *la grand-mère*.

(7) (Doetjes, Rebuschi, and Rialland 2004, 530)

C’est [le petit]<sub>F</sub> qui est tombé dans l’escalier, pas la grand-mère.

‘It’s the young one who fell down the stairs (, not the grand-mother).’

Nevertheless, French also has broad-focus clefts (cf. Doetjes, Rebuschi, and Rialland 2004; Dufter 2009; Karssenberg and Lahousse 2018). Though atypical, these clefts are frequent, especially when the clefted element is a PP adjunct, as in (8):

(8) (Doetjes, Rebuschi, and Rialland 2004, 535)

C’est [avec plaisir que je vous invite à participer à ce séminaire]<sub>F</sub>.

‘It is with pleasure that I invite you to this seminar’

The two sentences in (7) and (8) have different intonations, and the mismatch with the informational partition in (8) only occurs at the syntactic level, not at the prosodic one. Syntax being ambiguous, intonational marking is necessary here to disambiguate, and a mismatch between IS and intonation would give rise to an unacceptable sentence.

Consider now the Catalan example given above and repeated below.

(9) On són, els coberts ?

‘Where’s the flatware?’

a. Les forquilles són a l’armari, però...

‘The forks are in the cupboard, but...’



- b. ...els ganivets<sub>1</sub> els<sub>1</sub> vaig ficar al CALAIX.  
           the knives Cl.OBJ 1<sup>ST</sup>-PST put.INF in.the drawer  
           ‘...the knives, I left them in the drawer’
- c. ...#vaig ficar els ganivets al CALAIX.  
           1<sup>ST</sup>-PST put.INF the knives in.the drawer  
           ‘I left the knives in the drawer’

While the intonation in (9c) seems to be coherent with the informational partition of the sentence (the focal accent is on *al calaix*, which is the focus) the syntax is not acceptable, and a clitic right dislocation is expected instead. This example illustrates an opposite case than the one we have seen about French clefts, if we are sure that intonation does not play any role. But we may wonder whether the unacceptability of (9c) may not be related to the fact that *els ganivets* does not receive in such position a contrastive-topic accent, while it does in (9b). A prosodic analysis would therefore be necessary here to check whether intonation plays any role or not.

At any event, this couple of examples shows that a syntax-based account of IS phenomena would at best require a careful analysis of the interaction between syntax and intonation and evidence should be provided that intonational constraints are ultimately due to syntactic ones.

## 1.2 Information structure: methodology

Most of my research on IS has been dedicated to study the linguistic marking of information structure in Romance, but I have also been engaged in the search for the best methodology to study IS phenomena. I have soon understood that the study of data collected in natural settings, and in particular the study of spontaneous, unplanned oral speech (a genre in which the expression of the IS partition of the sentence strongly contributes to effective communication) is crucial to understand the exact function of IS markings, for various reasons: for the availability of the utterance’s linguistic context; for the possibility to take into account frequency (of a construction, or of some lexical choices); for the hints that the speaker’s lexical and grammatical choices may give. That is why I have contributed myself to the creation and annotation of spoken corpora, such as the Nocando corpus (see Brunetti et al. 2011).

Nevertheless, a problem that soon arises, while working on naturalistic data, is that the information structure of an utterance may not easily be identified in some contexts, independently from its form. If a researcher’s goal is to find the IS properties expressed by a certain linguistic construction (hence a certain form), then to start from its form to derive such properties looks like a circular procedure. It is therefore important to understand the informational properties of an utterance by looking at the context in which the utterance occurs. Even more so, if one considers that a certain linguistic form may be underspecified. Constructed examples have the advantage that they are provided with an ideal context, which makes the informational partition of the utterance completely transparent. In the literature, focus and aboutness topic are typically illustrated by examples where the target utterance is preceded by a question (see for instance (10) and (11)), and contrastive focus and topic by examples containing a sentence that constitutes the contrasting alternative, as in (12) or (13).

- (10) Q: Chi hai invitato alla festa?  
           ‘Who did you invite to the party?’  
       A: Ho invitato [Linda]<sub>F</sub>.  
           ‘I invited Linda’

- (11) Q: Com’è andata la cena? Che cosa ha mangiato Leo?

‘How did the dinner go? What did Leo eat?’

A: [Leo]<sub>T</sub> ha mangiato la pasta.

‘Leo ate pasta’

(12) A: Tommaso ha mangiato la pasta.

‘Tommaso ate pasta’

B: No, [Leo]<sub>CF</sub> l’ha mangiata.

no Leo CL.SG-has eaten

‘No, Leo ate it’

(13) Com’è andata la cena? Che cosa hanno mangiato i bambini?

‘How did the dinner go? What did the children eat?’

[Leo]<sub>CT</sub> ha mangiato la pasta e [Linda]<sub>CT</sub> ha preso il riso.

‘Leo ate pasta, and Linda took rice’

An utterance that is extracted from naturalistic data, however, may be preceded by a much less clear context. The awareness of these two facts – that naturalistic data are crucial to understand the IS properties of a linguistic construction and that context, through which the IS can be derived, can be more complex and therefore less clear in such kind of data – has led me to pursue the work that I will present in Section 2.

## 2 A QUD-based approach to information structure

### 2.1 Introduction

It is by now acknowledged in the literature that IS must be accounted for by appealing to the structuring of discourse; in particular, the importance of the discourse notion of question under discussion (QUD) is by now widely recognized. In order to find a way to annotate information structure in utterances of naturalistic data, Arndt Riester, Kordula De Kuthy and myself have proposed some criteria to annotate QUDs (that is, to find the implicit question preceding each utterance of a discourse), from the identification of which, through linking rules, the IS of each utterance is derived. Some guidelines about how to annotate QUDs are presented in Riester, Brunetti, and De Kuthy (2018) and further developed in several papers (De Kuthy, Reiter, and Riester 2018; Riester 2019; De Kuthy, Brunetti, and Berardi 2019; Brunetti, De Kuthy, and Riester 2021; Brunetti 2024). In the following section I will present such QUD-annotation procedure and the discourse structuring that derives from QUD annotation. In Sections 3 and 4 I will present, and reconsider through the lenses of such QUD-based discourse model, my previous analyses of two IS-related constructions in Romance: focus fronting and left dislocation.

The notion of QUD, which has been developed by many scholars since the ‘90s (see Kuppevelt 1995; Roberts 1996; 2012; Ginzburg 1995 and more recently; Onea 2016) is based on the idea that discourse is structured by means of (implicit or explicit) questions. In Riester, Brunetti, and De Kuthy 2018 we argue in particular that a question – usually an implicit one – precedes each utterance of a text. We think that the main contribution of our work to this line of studies is the fact that it provides an explicit procedure to formulate such implicit questions. In some QUD-based models (cf. e.g. Ginzburg 2012) much attention is given to dialogue, in which exchanges containing explicit questions occupy an important part. In our model, no special attention is given to dialogue, and QUD are mostly implicit. One study that also

discusses implicit questions at length is Onea (2016), who calls them ‘potential questions’. However, unlike other QUD frameworks (including ours) Onea takes potential questions to be questions that an utterance may *raise*, rather than questions that an utterance may *address*. They are called “potential” since only when the following utterance is considered, one knows which question is the intended one. In our model, an implicit question is the question that the target utterance answers, given the context in which it is uttered. Also, the way to derive the intended question goes in the opposite direction as compared to Onea’s procedure. In our case, the starting point is the target utterance, and the implicit questions this utterance may answer are reduced to one by applying the criteria to be discussed in the sections below.

## 2.2 QUD-reconstruction

We basically propose two methods of QUD reconstruction for an assertive discourse segment. The first one is guided by the principles listed in (14) (from Brunetti, De Kuthy, and Riester 2021, 17), the other method is presented in Section 2.2.1.

- (14) *Principles of QUD reconstruction*
- a. Q-A-CONGRUENCE: A QUD must be answerable by the assertion that it immediately dominates.
  - b. MAXIMIZE-Q-ANAPHORICITY: A QUD should be formulated using all the given semantic content of its answer.
  - c. Q-GIVENNESS: An implicit QUD can only consist of given content.

Principle (a) is trivial: the reconstructed question must be such that the target utterance (the utterance whose information structure we want to determine) answers it. This principle alone is too weak and many questions may fulfill it. Principle (b) further constrains the formulation of a question and prevents that the focus be too large (cf. Williams 1997; Schwarzschild 1999). In practice, it ensures that all parts of the assertion with anaphoric (given) content should be part of the QUD answered by that assertion. This is still not enough to exclude all possible questions but one, and a third principle is necessary, which ensures that no discourse-new content is present in the question. Indeed, if one assumes that the focus of the assertion is what answers the question and therefore corresponds to the *wh*-phrase (Rooth 1992), principle (c) makes sure that any part of the question that is not the *wh*-phrase will be given.

For an illustration of how the three principles work, consider (15), from an excerpt of a speech by Obama.<sup>4</sup>

- (15) English, Obama’s speech (Brunetti, De Kuthy, and Riester 2021, 17)

A<sub>1</sub>: While studying here, my father met my mother.  
Q’: #*What did he do after studying here?*                      \*Q-A-CONGRUENCE  
Q’’: #*What else?*    \*MAXIMIZE-Q-ANAPHORICITY  
Q’’’: ***What about Obama’s mother?***  
Q<sup>iv</sup>: #*Where was Obama’s mother born?*                      \*Q-GIVENNESS  
A<sub>2</sub>: She was born in a town on the other side of the world.

The assertion whose QUD we want to reconstruct is A<sub>2</sub>. Question Q’ might naturally arise from A<sub>1</sub>, but it cannot be answered by A<sub>2</sub>, so it is excluded by Q-A-CONGRUENCE. Q’’ to Q<sup>iv</sup> are all questions that A<sub>2</sub> is able to answer. However, MAXIMIZE-Q-ANAPHORICITY rules out Q’’,

---

<sup>4</sup> By convention, implicit QUDs are given in italics and they are always worded in English.

because this question is not specifically about Obama’s mother, who is given in the context and referred to in  $A_2$  by means of the pronoun *she*.  $Q'''$  and  $Q^{iv}$  both contain a reference to Obama’s mother, but Q-GIVENNESS rules out  $Q^{iv}$ , because some of its content is not yet given in the preceding context (*was born* is not anaphorically retrievable from  $A_1$ ). The only question that adheres to all three principles is therefore  $Q'''$  (in boldface), which simply asks for some additional information about Obama’s mother.

The formulation of the QUD is therefore a way to reconstruct the target utterance’s given content (the background), while the focus is the new piece of information that answers the QUD. For instance in (15), given the local linguistic context represented by  $A_1$ , the only given content of  $A_2$  is *she*. Its QUD must therefore only contain a reference to the speaker’s mother (in addition to the wh-phrase and to anything that may be necessary to formulate a grammatical question). Consequently, *was born in a town on the other side of the world* is labeled as focus and *she* as background (see (16)). If, instead, question  $Q''$  had been chosen, *she* in  $A_2$  would have been wrongly included in the focus part; if, on the other hand,  $Q^{iv}$  had been chosen, then some linguistic material, specifically *was born*, would have been wrongfully excluded from the focus part. Following Rooth’s (1992) conventions, the annotation only marks the focus by means of F-indexed square brackets, while the background is what is left outside the focus within the ‘focus domain’ (the sum of focus and its background), marked by brackets and a squiggle sign  $\sim$  (again, following Rooth’s notation).<sup>5</sup>

- (16) English, Obama’s speech  
 $Q_2$ : *What about Obama’s mother?*  
 $>A_2$ : [She [was born in a town on the other side of the world]<sub>F</sub>.]<sub>~</sub>

Through the formulation of QUDs, a discourse structure is derived, under the form of a tree whose terminal nodes are the answers to the QUDs (the actual utterances of the text). The symbols  $>$ ,  $>>$ ,  $>>>$ , etc. in the examples correspond to the level of embedding of questions and answers in the discourse tree.

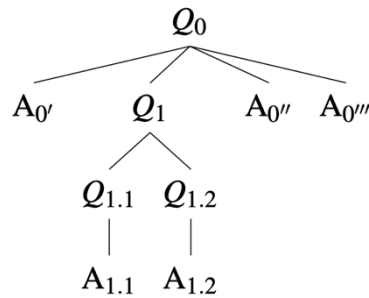


Figure 1: An example of QUD tree, taken from (Brunetti, De Kuthy, and Riester 2021, 20).

In Figure 1,  $Q_0$  is answered by  $A_0'$ ,  $A_0''$  and  $A_0'''$ ; however, the last two answers follow the answers to the two sub-questions of  $Q_1$ .<sup>6</sup>

<sup>5</sup> Rooth’s Alternative Semantics’ framework will be recalled more in details in Section 3.2.

<sup>6</sup> Note that the content dependency of a question  $Q$  on a preceding assertion  $A$  is not represented vertically but horizontally. That way the tree is rendered compact, the actual assertions of the text remain the terminal nodes of the QUD tree and can be read off from left to right. In other words, although  $A_0'$  and  $Q_1$  are represented at the same level,  $Q_1$  is anaphoric to the content of  $A_0'$  and informationally subordinated to it. See Riester (2019) for a discussion and alternative representations.

### 2.2.1 Parallelism

The second method of QUD reconstruction is based on the principle called ‘Parallelism’. The principle is given in (17) (Brunetti, De Kuthy, and Riester 2021, 18), and accounts for pairs or lists of utterances – generally coordinated ones – that answer the same QUD:

- (17) PARALLELISM: A QUD that is directly answered by two or more answers is formulated on the basis of the semantic content that is shared by the answers.

Unlike the principles in (14), PARALLELISM is forward-looking, in that the QUD can be made of new content if that content is shared by an utterance *following* the target one. In other words, this principle overrides Q-GIVENNESS. Example (18) is taken from a French spoken sociolinguistic interview of the CFPP2000 corpus (Branca-Rosoff et al. 2012). The speaker is answering a question from the interviewer about what districts of Paris she likes.<sup>7</sup>

- (18) French, CFPP2000 (7<sup>ème</sup>) (Brunetti 2024, 6)  
 A<sub>23</sub> : alors euh j’aime beaucoup euh tout c’qui est euh: aux environs de bah  
 j’aime beaucoup le septième euh où qu’ce soit  
 ‘so ehm I love very much all that is close to... well I love the 7th district  
 very much, no matter where’  
 Q<sub>24</sub> : *Where specifically, within the 7th district?*  
 >A<sub>24</sub>’ : [que ça soit [vers la Tour-Maubourg]<sub>F</sub>]~  
 ‘be it towards the Tour-Maubourg’  
 >A<sub>24</sub>” : [que ce soit [ici]<sub>F</sub>]~  
 ‘be it here’

This is an example of ‘simple’ PARALLELISM. The segments A<sub>24</sub>’ and A<sub>24</sub>” share some semantic content (represented by the almost identical parts *que ça soit* and *que ce soit*) and differ with respect to the answer they give to Q<sub>24</sub>. Example (19), from the same corpus, is an instance of ‘complex’ PARALLELISM. Spk1 is the interviewer (a linguist) and Spk2 is the interviewee, a person living in the 13<sup>th</sup> district of Paris.

- (19) French, CFPP2000 (13<sup>ème</sup>) (Brunetti 2024, 6)  
 >Q<sub>1</sub> : Spk1: comment est-c’que euh toi ou tes parents vous êtes arrivés dans  
 l’quartier (...)  
 ‘how did you or your parents arrive in the district (...)?’  
 >>Q<sub>1.1</sub> : *How did you arrive in the district?*  
 >>>A<sub>1.1</sub> : Spk2 : alors, donc [[moi]<sub>CT</sub> j’suis arrivé à Paris [j’étais tout petit ]<sub>F</sub>]~  
 (...)  
 ‘well, as for myself, I arrived in Paris as a little child’  
 >>Q<sub>1.2</sub> : *How did your parents arrive in the district?*  
 >>>A<sub>1.2</sub> : et [[mes parents]<sub>CT</sub> sont venus à Paris [pour le boulot]<sub>F</sub>]~  
 ‘and my parents came to Paris for work’

<sup>7</sup> Some of the pauses, fillers, repetitions, etc. that are present in the transcription have been removed for the sake of clarity.

In this example, the QUD is not implicit but is explicitly uttered by the interviewer.<sup>8</sup> The characteristic of complex PARALLELISM is that the question is only partially answered by each utterance (A<sub>1.1</sub> and A<sub>1.2</sub>); each utterance however fully answers a *subordinate* QUD (Q<sub>1.1</sub> and Q<sub>1.2</sub>). This is basically the QUD-based analysis that Buring (2003), following Roberts (1996; 2012), gives to the effects of (contrastive) topic accents in English and German: they signal a discourse configuration where a superordinate QUD is followed by sub-QUDs whose answers are partial answers to the super-question. Such a discourse structure not only identifies the focus variable but also the topic variable, which is indexed with CT (for “Contrastive Topic”) in (19).<sup>9</sup>

The difference between simple and complex PARALLELISM is basically that, with the former, the topic is fixed and the only evoked alternative set is the one introduced by the focus, while with complex PARALLELISM, the topic varies too.<sup>10</sup> I will return to PARALLELISM in Section 5, where I will discuss the notion of contrast within the QUD-based model described above.

### 2.3 Not-at-issue content and the information structure of adjuncts

In Riester, Brunetti, and De Kuthy’s (2018) annotation procedure, some parts of the utterance are excluded from the focus domain: they are neither focus nor background, that is, they are not part of the answer to the QUD. These expressions are discourse markers and coordinating or subordinating conjunctions, the meanings of which stay outside the propositional content and are non-truthconditional (they are “conventional implicatures”, see Grice 1975, 45):

- (20) German, (Riester, Brunetti, and De Kuthy 2018, 408)
- A<sub>3</sub>: Wir haben ja nun alle von Konflikten gehört,  
‘Now, we have all heard of conflicts’
- Q<sub>4</sub>: *What about conflicts?*
- > A<sub>4</sub>: **aber** [[es gibt immer]<sub>F</sub> Konflikte]~  
‘but [[there are always]<sub>F</sub> conflicts]~’

Other than such expressions, it has soon become evident, during the annotation process, that Potts’s (2004) supplements, which include non-restrictive relative clauses (NRRCs), parentheticals, appositions, must also be left outside the focus domain. As a matter of fact, these expressions too, according to Potts’ analysis, trigger conventional implicatures. Speaker-oriented expressions, such as ‘thank God!’ are considered as external to the focus domain as well.

- (21) French blog (<https://www.bloghoptoys.fr/dossiers/page/42>)
- >> Q<sub>1</sub>: *How do you know Nathalie?*
- >>> A<sub>1</sub>: [Nathalie] **maman de deux enfants nés en 2007 et fin 2011**, [nous [a été présentée par une maman blogueuse]<sub>F</sub>.]~  
‘Nathalie, **mum of two children born in 2007 and end of 2011**, was introduced to us by a blogger mom’

More interestingly, during the QUD-annotation of different kinds of texts, we came to the conclusion that verb-dependent adjuncts – subordinate clauses, Prepositional Phrases, adverbials – also might have to be excluded from the focus domain in some contexts (though

<sup>8</sup> QUDs can be questions that are actually produced by a speaker. Unlike implicit questions, explicit questions can be made of new content. I will return to non-assertive speech acts and how to deal with them in the QUD-model in Section 6.

<sup>9</sup> I will return to Buring’s analysis of contrastive topics in Section 4.1, when I will discuss the notion of topic.

<sup>10</sup> For the same discourse structures, Umbach (2005) uses the terms “simple contrast” and “double contrast”.

not always). An example of a temporal subordinate clause that is excluded from the focus domain is given in (22), from the French spoken Rhapsodie corpus (Lacheret, Kahane, and Pietrandrea 2014), and in particular from the sub-corpus from Mertens (1987). The excerpt comes from a radio interview of a journalist to a French writer, Françoise Giroud:

- (22) French, Rhapsodie (Brunetti, De Kuthy, and Riester 2021, 5)  
 ‘It is an experience that I’ve never forgotten’  
 >Q<sub>8</sub>: *What about the harshness of this experience?*  
 >>A<sub>8</sub>’: [ce qui est dur ce n’est pas] **surtout quand on est très jeune** [ce n’est vraiment pas [d’être pauvre (...)]<sub>F</sub>]~  
 ‘What is hard is not, especially when one is very young, it is not really to be poor’  
 >>A<sub>8</sub>’’: (...) [c’est [de se dire je n’en sortirai jamais]<sub>F</sub>]~  
 ‘it is to say “I will never get out of it”’

In Riester et al (2018), we propose that such an expression conveys ‘not-at-issue’ content, which we define as follows:

- (23) *Non-at-issue content (relative to Q)*  
 An expression *X* whose denotation is discourse-new and which is contained in an assertion *A* is non-at-issue with respect to the current QUD *Q* iff *X* is optional with respect to *Q*, where optional means that under deletion of *X*, *A* is still an answer to *Q*.

This definition well accounts for the exclusion of adjuncts, since optionality is a characteristic of adjuncts.<sup>11</sup> The definition, however, does not take into account a property that an important part of the literature considers as crucial to define not-at-issue meanings, namely “projection”. In the last fifteen years, a number of studies by Simons, Beaver, Roberts, and Tonhauser have tried to give a precise definition of the notion of not-at-issueness. Simon et al (2010)’s characterization relies on Roberts’s (1996; 2012) QUD framework of discourse structure and is defined in terms of relevance to the QUD, where an assertion is defined as relevant to the QUD iff it contextually entails a partial or complete answer to the QUD. Another crucial property that they ascribe to non-at-issue content is that it projects (Simons et al. 2010; Tonhauser et al. 2013). These authors propose a biunivocal relationship between projective and not-at-issue content: all projective content is not at issue and all non-at-issue content projects. In their words, projection “is intimately related to the structuring of information in discourse. It is a consequence of the fact that in the totality of information conveyed by an utterance, some is central to the speaker’s conversational goals, and some is peripheral. The peripheral projects.” (Simons et al. 2010, 325).

The ability to project is tested by means of the *family-of-sentences* (FOS) diagnostics, defined in Chierchia and McConnell-Ginet (2000) – and developed further in Tonhauser et al. (2013) – which consist in adding such operators to the sentence; if the content conveyed by a part of the sentence is not affected by such operators, then it projects. Projection is a property that was first noticed for presuppositions, but since Potts (2004), it has been found to apply to conventional implicatures too (see also Simons et al. 2010; Tonhauser et al. 2013; Faller 2014). In (24), for instance, the inference triggered by the apposition *maman de deux enfants nés en 2007 et fin 2011* (‘mum of two children born in 2007 and end of 2011’) is neither negated, nor

<sup>11</sup> In fact, it depends what one means by ‘optionality’. As we will see below, adjuncts are always optional in syntax, but may be pragmatically obligatory in certain contexts.



questioned when the main clause is, nor its truth is affected when it occurs inside the scope of the antecedent of a conditional or of a modal.

- (24) French blog (<https://www.bloghoptoys.fr/dossiers/page/42>)  
Nathalie, maman de deux enfants nés en 2007 et fin 2011, nous a été présentée par une maman blogueuse.  
'Nathalie, mum of two children born in 2007 and end of 2011, was introduced to us by a blogger mom'  
*Negation.* Nathalie, maman de deux enfants nés en 2007 et fin 2011, ne nous a pas été présentée par une maman blogueuse.  
'Nathalie, mum of two children born in 2007 and end of 2011, was not introduced to us by a blogger mom'  
=> '*Nathalie is the mum of two children... 'is true*  
*Interrogative.* Est-ce que Nathalie, maman de deux enfants nés en 2007 et fin 2011, vous a été présentée par une maman blogueuse?  
'Was Nathalie, mum of two children born in 2007 and end of 2011, introduced to you by a blogger mom?'  
=> '*Nathalie is the mum of two children... 'is true*  
*Conditional.* Si Nathalie, maman de deux enfants nés en 2007 et fin 2011, vous a été présentée par une maman blogueuse, alors...  
'If Nathalie, mum of two children born in 2007 and end of 2011, was introduced to you by a blogger mom, then...'  
=> '*Nathalie is the mum of two children... 'is true*

The definition of not-at-issue content as a meaning that projects is supported by the behavior of discourse markers, supplements or speaker-oriented elements. On the contrary, it is not at all evident that the same definition holds for adjuncts. We must therefore separate the issue about the IS of adjuncts into two parts: we must first find out whether some adjuncts stay outside the focus domain, and if yes, whether they project. In Brunetti, De Kuthy, and Riester (2021) we explicitly tackle these issues by analyzing the information-structural properties of adjuncts within our QUD-based model of discourse.

### 2.3.1 The information structure of adjuncts

The role of adjuncts in the information-structural partitioning of the sentence has been rarely addressed in the literature. An exception is the work by Fabricius-Hansen and colleagues, see e.g. the collective work on co-eventive adjuncts represented by non-finite predicate- or clause-like adverbial phrases (Fabricius-Hansen and Haug 2012). Some work has also tackled the prosody of adjuncts, which can reveal some of their informational properties (see Selkirk 1996; Samek-Lodovici 2005; Truckenbrodt 2007). In Brunetti, De Kuthy, and Riester (2021) we explore the informational properties of adjuncts that modify either the verb or the whole sentence (we call them 'V/S-adjuncts'), and which can be either clausal or non-clausal expressions. The data come from oral corpora and written texts in English, French, and German.

Adjuncts are "syntactically parasitic in the sense that they are added within fully fledged syntactic constituents that would be perfectly well-formed without them" (Schlenker 2023, 40). In the paper we show that the informational and discourse properties of V/S-adjuncts goes hand in hand with their special syntactic status. Indeed, we argue that in some contexts, V/S-adjuncts convey new information without being part of the sentence focus (cf. (23)); for such cases, we propose that the adjunct forms an independent segment in the QUD discourse tree. The adjunct's syntactic autonomy matches its discourse autonomy. Such correspondence however does not occur all the time: adjuncts can also represent focus, background, or (contrastive)



topics. In such cases, while syntactically autonomous and still optional, they are not autonomous at an informational and discourse level. Examples of adjuncts as focus, background or contrastive topic respectively are given below, together with their QUD annotation. Ex. (25) comes from the French spoken CFPP2000 corpus. The causal PP adjunct *pour le boulot* ‘because of their job’ is the answer to the implicit sub-QUD introduced by the wh-word ‘how’ (*How did you arrive in the district?*) and therefore constitutes the sentence focus.

- (25) French, CFPP2000  
 >Q<sub>1</sub>: Spk1: euh comment est-c’que euh toi ou tes parents vous êtes arrivés dans l’quartier (...)  
 ‘How did you or your parents end up in this district [of Paris] (...)?’  
 >>Q<sub>1.1</sub>: *How did you arrive in this district [of Paris]?*  
 >>>A<sub>1.1</sub>: Spk2: alors, donc [[moi]<sub>CT</sub> j’suis arrivé à Paris [j’étais tout petit]<sub>F</sub>]~ (...)  
 ‘well, as for me, I arrived in Paris as a little child’  
 >>Q<sub>1.2</sub>: *How did your parents arrive in the district?*  
 >>>A<sub>1.2</sub>: et [[mes parents]<sub>CT</sub> sont venus à Paris [pour le boulot]<sub>F</sub>]~  
 ‘and my parents came to Paris for work’

The French blog example in (26) illustrates a case of adjunct that is part of the background (the given content). The denotation of *dans ces ateliers* ‘in these workshops’ is given in A<sub>1</sub>, and it is then included in Q<sub>2</sub> (it is therefore part of the background).

- (26) French blog (<https://www.bloghoptoys.fr/dossiers/page/42>)  
 A<sub>0</sub>: Elle (...) a créé (...) Kestumdis,  
 ‘She has created Kestumdis’  
 Q<sub>1</sub>: *What does Kestumdis do?*  
 >A<sub>1</sub>: qui propose des ateliers familiaux de Langue des Signes avec les bébés.  
 ‘which offers family workshops on sign language with babies.’  
 >Q<sub>2</sub>: *What does she do in these workshops?*  
 >>A<sub>2</sub>: [Dans ces ateliers, elle [accueille aussi les enfants différents]<sub>F</sub>]~ .  
 ‘In these workshops, she also hosts children that are different’

Finally in (27), an example of contrastive topic adjuncts is given, from the same corpus.

- (27) French, CFPP2000  
 > Q<sub>14</sub>: What happened, when you moved?  
 >> Q<sub>14.1</sub>: What happened during the first week?  
 (...) en fait j’crois que [[la première semaine]<sub>CT</sub> [mes parents m’emmenaient à l’école]<sub>F</sub>]~  
 (...) ‘Well, I think the first week my parents brought me to school’  
 >> Q<sub>14.2</sub>: What happened after you moved?  
 et [[après, dès qu’on a emménagé]<sub>CT</sub> euh bah [ça allait]<sub>F</sub>]~  
 ‘and then, when we moved, ehm well, it was ok’

As the QUD structure shows, the two temporal expressions *la première semaine* ‘the first week’ and *après, dès qu’on a emménagé* ‘later, once we moved’ are the two contrastive topics of a complex PARALLELISM construction, since they constitute the two times at which the speaker reached her school in the two different ways described, respectively (from *ça allait* ‘it was ok’

one gets the inference that, when she moved, the distance of the speaker’s home from school was not a problem anymore, so she did not need anymore to be accompanied – supposedly by car – by her parents).

In other contexts, as I said, adjuncts provide some peripheral or parenthetical information that is *not* part of the focus domain (neither as focus, nor as topic or background), namely that is not part of the answer to the current QUD. In Brunetti et al (2021) we call such adjuncts ‘I(nformation)S(tructure)-peripheral’ ones. Our claim is that IS-peripheral adjuncts answer a *different* QUD than the one that is answered by the sentence hosting them. This new QUD, however, is related to the current QUD in that it anaphorically depends on the host. Let us see some examples. IS-peripheral adjuncts can occupy a sentence final, initial, or middle position. An example final position is given below.

(28) French, CFPP2000 corpus (Brunetti, De Kuthy, and Riester 2021, 14)

Q<sub>1</sub>: Spk1: ouais donc vous allez beaucoup à René Le Gall [parc] ou au petit [parc]...?

‘yeah, so do you often go to René Le Gall [playground] or to the little one...?’

> Q<sub>2</sub>: *Which playground do you go to?*

>> Q<sub>2.1</sub>: *Which playground do you go to, when you are not very motivated?*

>>> A<sub>2.1</sub>: Spk2: ouais bah [[quand on est pas très motivés]<sub>CT</sub> [on peut descendre juste en bas]<sub>F</sub>]~ (...)

‘yeah, well, when we are not very motivated we may go just down the road (...)’

>> Q<sub>3</sub>: *For how long?*

>>> A<sub>3</sub>: [[pour une petite heure]<sub>F</sub>]~  
‘for about an hour’

Q<sub>2.1</sub> is completely answered by A<sub>2.1</sub>, and the temporal adjunct *pour une petite heure* ‘for about an hour’ answers the separate sub-question Q<sub>3</sub> *For how long (do you stay at the playground that is down the street)?* The tree representation of (28) is given in Figure 2. The independent segment constituted by the adjunct is the circled one.

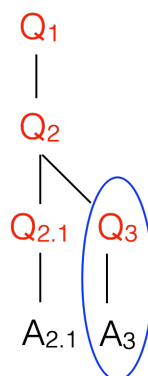


Figure 2: QUD tree corresponding to (28).

It is less obvious that sentence-initial adjuncts can constitute an independent segment:

(29) German, *Neue Westfälische*

‘In a solo accident on the L 756 a 19-year-old girl sustained serious injuries on Sunday afternoon.’

Q<sub>2</sub>: *What happened to her?*

- >A<sub>2</sub>': [Die Frau [war mit einem Fiat (...) unterwegs]<sub>F</sub>]~  
 'The woman was driving a Fiat (...).'  
 >Q<sub>3</sub>: *Why did she come off the road?*  
 >>A<sub>3</sub>: [[**Aus bislang noch unbekannter Ursache**]<sub>F</sub>]~  
 'For reasons still unknown,'  
 >A<sub>2</sub>'': [[kam]<sub>F</sub> sie (...) [nach rechts von der Fahrbahn ab und überschlug sich]<sub>F</sub>]~  
 'she came off the road to the right (...) and rolled over.'

The sub-question that the adjunct answers is in a *cataphoric* relation with the upcoming host sentence, which is A<sub>2</sub>''. This is also illustrated in Figure 3.

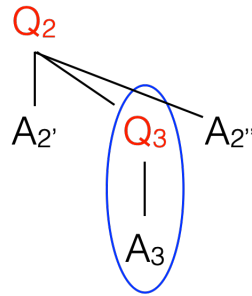


Figure 3: QUD tree corresponding to (29).

Even more problematic is the structural position of sentence-internal adjuncts. In (30), the temporal clause *quand on est très jeune* is inserted in the middle of the segment A<sub>8</sub>'. We propose to represent the segment that is split into two chunks by repeating its label (here, A<sub>8</sub>') twice, see also Figure 4:<sup>12</sup>

- (30) French, Rhapsodie (Mertens corpus) (Brunetti, De Kuthy, and Riester 2021, 29)  
 'It's an experience that I've never forgotten'  
 > Q<sub>8</sub>: *What is hard, in this experience?*  
 >> Q<sub>9</sub>: *Is it to be poor that is hard?*  
 >>>A<sub>9</sub>...: (...) [ce qui est dur ce n'est pas  
 'What is hard is not...'  
 >>>Q<sub>10</sub>: *When is it not to be poor that is hard?*  
 >>>>A<sub>10</sub>: [surtout quand on [est très jeune]<sub>F</sub>]~  
 'especially when one is very young'  
 >>>...A<sub>9</sub>: ce n'est vraiment pas [d'être pauvre (...)]<sub>F</sub>~  
 '...it is not really to be poor'

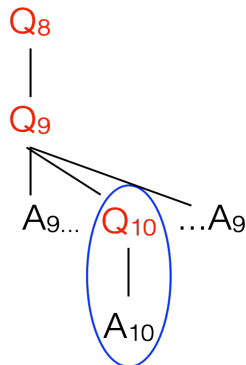


Figure 4: QUD tree corresponding to (30).

<sup>12</sup> See Riester (2019) for a different account.

Internal adjuncts, similarly to sentence-initial ones, answer a question that is reconstructed on the basis of the host sentence *before* the host sentence has ended. In this case, we may call such a relation *intraprophic*.

The observation that certain informational elements, despite being new, do not answer the current QUD but a separate question has already been made in the literature. Onea (2016) makes this observation with respect to supplements like those in (31). He says that the questions are “over-answered” by speaker B, and that the utterances provide “too much information” (Onea 2016, 216).

- (31) (Onea 2016, 216)
- a. A: Did any stock rise yesterday?  
B: Yes, **Alcatel and Telefonos Mexicanos**. (from Zeevat 2007)
  - b. A: Whom did Mary kiss?  
B: She kissed JOHN, **who you met yesterday**.
  - c. A: Whom did John introduce to Jane?  
B: John, **cheerfully**, introduced MAX to Jane.

Onea analyzes the redundant parts (in boldface) within his theory of ‘potential questions’. He assumes that the extra information is not at issue and claims that “once we add the at-issue content to the discourse, this will license a new potential question which will accept the not-at-issue content of the candidate assertion as congruent answer” (Ibid.).

Ramm (2011: 146) makes similar comments on a constructed example where the redundant element is a manner adjunct. His example is given in (32). Ramm says that, while *yesterday evening* is clearly focused, since it answers the question about the time of the hearer’s arrival, the adjunct *with some friends* “encodes information that is new, i.e. not part of the background, but does not contribute to answering the relevant question and thus cannot be part of the focus in the strict sense either.” (Ramm 2011, 146)

- (32) (Ramm 2011, 146)
- When did you arrive?
  - I arrived yesterday evening **with some friends**.

Ramm does not discuss this example further, but notably suggests that “the adjunct, in a way, answers a question that has not been asked” (Ibid.).

Our proposal also receives some support from prosody. In their analysis of English and German co-eventive adjuncts, Fabricius-Hansen and Haug (2012b), follow Truckenbrodt (2007) and argue that adjuncts “are prominent by default and do not trigger deaccenting of an adjacent verb (predicate), in contrast to what may happen when the verb occurs next to a prominent complement.” An example of this difference is given in (33):

- (33) (Truckenbrodt 2007, 446)
- a. - What does he do?  
- He [teaches linguistics]<sub>F</sub>  
- Er soll Linguistik [unterrichten]<sub>F</sub>.  
he MODAL linguistics teach  
'He is said to teach linguistics'
  - b. - [What does he do?]  
- He [teaches in Ghana]<sub>F</sub>.  
- Er soll [in Ghana unterrichten]<sub>F</sub>.

he MODAL in Ghana teach  
'He is said to teach in Ghana'

Fabricius-Hansen and Haug point out that “deaccenting the lexical predicate adjacent to a prominent adjunct indicates narrow focus on the latter” (Fabricius-Hansen and Haug, 2012b: 22). In other words, when the adjunct is the only focus, then the verb and the adjunct constitute a prosodic unit, but when they are both focused, they can be treated prosodically as two independent units. Since, according to our proposal, the adjunct in (b) constitutes an independent discourse unit that answers its own QUD (see the QUD annotation below), it is then expected that the adjunct bears its own focal accent.<sup>13</sup>

- (34) Q<sub>1</sub>: *What does he do?*  
 > A<sub>1</sub>...: Er soll  
 > Q<sub>2</sub>: *Where?*  
 >> A<sub>2</sub>: [[in Ghana]<sub>F</sub>]~  
 > ...A<sub>1</sub>: [[unterrichten]<sub>F</sub>]~  
 'He is said to teach in Ghana'

Summarizing, if an adjunct has a specific informational role inside an utterance, namely it is its focus or background or contrastive topic, then its role is obviously not compatible with the adjunct being an independent discourse unit. In these cases, the adjunct cannot be omitted without making the host utterance unacceptable *in that context*. Its omission does not have an effect on the syntax of the host, which remains well-formed, but the adjunct is obligatory from a pragmatic point of view. When the adjunct is IS-peripheral, on the contrary, it is so to speak useless within the utterance from an information-structural point of view; however, as an independent unit, it acquires a role in the structuring of discourse.

### 2.3.2 Back to not-at-issue content

As I mentioned in the previous section, Onea (2016) proposes a similar analysis for certain supplements, namely appositions and Non-Restrictive Relative Clauses (NRRCs): he claims that they do not answer the same QUD as their host sentence, but a sub-question that is raised by the host. In Brunetti et al (2021), we indeed account for supplements in the same way as IS-peripheral adjuncts, as shown in (35) for the apposition *maman de deux enfants nés en 2007 et fin 2011*.

- (35) French blog (<https://www.bloghoptoys.fr/dossiers/page/42> )  
 >>Q<sub>1</sub>: *How do you know Nathalie?*  
 >>>A<sub>1</sub>...: [Nathalie  
 >>>Q<sub>2</sub>: *What about Nathalie?*  
 >>>>A<sub>2</sub>: [**maman de deux enfants nés en 2007 et fin 2011**]<sub>F</sub>]~,  
 >>>...A<sub>1</sub>: nous [a été présentée par une maman blogueuse]<sub>F</sub>]~  
 'Nathalie, mum of two children born in 2007 and end of 2011, has been introduced to us by a blogger mom'

<sup>13</sup> As Fabricius-Hansen and Haug (2012b: 26) put it: “forming intonation phrases of their own, detached adjuncts have no impact on the information structure of their host [and] behave like ‘orphans’.” Truckenbrodt (2007) though shows that the picture is more complex than that, since the prosodic shape of a phrase is also determined by other phonological factors, such as rhythmic constraints or constituent stress clash. According to the impressionistic judgment of our data, indeed IS-peripheral adjuncts do not seem to have an uncontroversial detached prosody.

The similarity between IS-peripheral adjuncts and supplements is not surprising. Appositions and NRRCs are, syntactically, adjuncts too: the only difference is that they depend on a noun instead of a verb or clause.<sup>14</sup> Furthermore, both supplements and IS-peripheral adjuncts represent new information but do not constitute the main point of the utterance (see discussion at the beginning of Section 2.3). Nevertheless, IS-peripheral but can have all kinds of information-structural properties, depending on the context, while supplements do not seem to ever be able to take part in the information structure of their host.<sup>15</sup> Another crucial difference is, as I mentioned above, their ability to project. We have seen in (24), by applying the family-of-sentences diagnostics, that supplements project. Different results are obtained with adjuncts. Consider the results of the same diagnostics to the English sentence below (from a spoken interview of a journalist to Snowden):<sup>16</sup>

- (36) English (Brunetti, De Kuthy, and Riester 2021, 13)
- “The greatest fear I have”, and I quote you, “regarding the disclosures is nothing will change.” That was one of your greatest concerns at the time
- a. *Negation*. and now, there is still no discussion about the situation with the NSA in America or elsewhere.
- b. *Interrogative*. What about today? Is there a discussion about the situation with the NSA, in America or elsewhere?

The meaning of the adjunct is the location where the event takes place, so if the event does not take place, then the location is one where the event does *not* take place. If the event realization is put into question, then the same holds for the event’s location. In (36a), the speaker says that there is no discussion in America, hence, the question of its location does not even arise. In (36b), the question is whether there is any discussion. Again, only if the discussion takes place, it can have a location.

In Brunetti et al (2021) we propose that the difference between the behavior of V/S-adjuncts and supplements is linked to a difference already mentioned earlier, namely that the former depend on a verbal head, while the latter – in particular appositions and NRRCs – depend on a nominal head. Indeed, such a difference has drastic effects on the semantic import of these elements. By adopting Venhuizen et al’s (2014) account, we argue that it is this last difference that explains why supplements project and IS-peripheral adjuncts do not. Venhuizen et al. observe that the supplement’s nominal anchor always triggers an existential presupposition. They therefore propose that supplements project because they piggyback on their anchor, since they provide an elaboration on the description of the anchor’s referent. Since IS-peripheral adjuncts attach to a verbal anchor, they cannot take advantage of any existing projective power, and are by consequence unable to project.

<sup>14</sup> This assumption is perhaps less obvious for NRRCs. We assume here the standard view that they are attached to the Determiner Phrase, as in Del Gobbo (2003); Koev (2013); Poschmann (2018).

<sup>15</sup> Such a difference must perhaps be nuanced. While it seems rather obvious that supplements cannot be narrow foci, the constructed example in [i] suggests that an apposition may function as a contrastive topic. The context is one where Mary has two different roles, one as professor and one as secretary, and behaves differently depending on the role she performs. Since the referent is the same, the contrast lies on the property described in the apposition (Mary’s role).

(i) Mary, the professor, would never do that, but Mary, the secretary, would absolutely go for it.

Concerning the possibility of being background elements, if we assume that supplements are conventional implicatures and therefore convey new information (as a definitory property of conventional implicatures) then they are never part of the background.

<sup>16</sup> The examples of IS-peripheral adjuncts seen above are too complex and make the results of the diagnostics more difficult to interpret. That’s why I use here this English example (still from Brunetti et al 2021).

What are the consequences of this analysis, with respect to the notion of not-at-issueness? If V/S-adjuncts do not project, not even when they are IS-peripheral, then either we must exclude them from the realm of not-at-issue content, or a different less strict definition of not-at-issue content must be given, which does not include projection as an obligatory property. We leave this problem open here. At any rate, within Riester et al.'s (2018) representation of discourse by means of QUD trees, projective and non-projective content cannot be distinguished, since both supplements and IS-peripheral adjuncts are treated as independent discourse units, which answer independent sub-questions with respect to their host, and their difference has rather to be captured in the semantic analysis.<sup>17</sup>

## 2.4 Some final remarks on the QUD-model

The QUD-based model of information structure presented above has revealed itself to be a promising framework within which to explain IS phenomena. We have seen in the preceding sub-sections that by its application to the analysis of the information structure of adjuncts, a novel informational analysis of these syntactic objects has emerged. The discussion has also led us to a rethinking of the definitory properties of not-at-issue content.

In the following sections I will review and reconsider, in the light of the QUD-based model, the analysis of the informational properties of two Romance constructions that have been the object of much of my work: focus fronting and (clitic) (left) dislocations. In Section 5 I will also reconsider the notion of contrast – which has always been central in my study of information structure – in the light of the QUD-model and more in general within the idea that IS phenomena must be analyzed and explained by taking the properties of discourse and discourse structure into account.

In Section 6 I will finally discuss a weakness of the QUD-model's explanatory power, namely the fact that it provides tools to derive the IS of assertions, but says little or nothing at all about how to derive the IS of non-assertive speech acts. I will briefly sketch how the model could be adapted to fill this gap.

# 3 Focus and focus fronting

## 3.1 Introduction

As discussed in Section 1, information structure concerns a set of pragmatic phenomena that may be expressed in different ways in the languages of the world, and which can therefore interact with different modules of grammar. According to Vallduví (1992), Neeleman and Van De Koot (2008), Leonetti (2014) and others, IS constitutes an independent module of grammar, which evaluates the suitability in a certain context of the morpho-syntactic and prosodic configurations that are made available by a language. In this perspective, there is no need that IS categories be encoded in the syntactic tree nor that syntactic features corresponding to IS categories make a link with “core” syntax, as assumed in cartographic approaches (such as Rizzi 1997; 2010; Belletti 2004, among many others). My point is that while a certain syntactic position may strongly favor the representation of a certain IS configuration, it is not bound to it. Furthermore, even in cases in which syntax always forces a certain IS interpretation, we do

---

<sup>17</sup> Either in a multidimensional analysis à la Potts (2004), in which non-at-issue content provided by CIs is not part of the semantic content of the main assertion, or like in Venhuizen et al.'s (2014) analysis, where the semantic anchor has to be chosen in a way that the projective behaviour of these expressions can be explained.

not need to appeal to a dedicated position to explain why this is the case. A good illustration of this perspective is focus fronting, a construction that I've been studying for my PhD thesis and in later work.

### 3.2 Focus fronting in Italian and Spanish

In Italian or Spanish, the focus of a sentence typically occupies a postverbal position, even if the focus is the subject, which is not obligatorily preverbal in these languages:

- (37) Italian  
 a. Chi ha mangiato la marmellata?  
     'Who ate the jam?'  
 b. L'ha mangiata GIANNI.  
     it.CL-has eaten Gianni  
     'GIANNI ate it'

When a focused constituent is fronted, both in these languages and in other European languages, the focus is usually associated with some special interpretation, such as contrast, mirativity, exhaustivity, depending on the language and on the context (see Cruschina 2021 and the literature quoted there).

In Brunetti (2009a), I argue for Italian and Spanish that the special interpretations that the fronted focus gets are not due to an intrinsic property of the fronted syntactic position, but to the fact that what follows the fronted constituent is unambiguously interpreted as background. I assume an Alternative Semantics account for focus as in Rooth (1992). Within such a framework, which is proposed to explain the effects of accent on sentence interpretation in English, focus is defined as an alternative-evoking phenomenon. More precisely, Rooth (1992) proposes that a sentence containing a focal accent in English, such as *MARY likes Sue* (where capital letters indicate the word carrying the accent), has an ordinary semantic value plus an additional 'focus semantic value' that corresponds to a set of propositions of the type *x likes Sue*, where the value of the variable in each proposition is one of the possible alternatives to the focus in the relevant context. For instance, in a sentence like (38a) where *a book* is the focus, the set of propositions is of the kind displayed in (38b).

- (38) a. I gave a BOOK to John.  
       b. *Focus semantic value:*  
         {I gave a book to John, I gave a t-shirt to John, I gave a cookie to John, etc.}

Obviously, not all potential alternatives are relevant in the context in which the sentence occurs. Rooth proposes that the focus value has to be restricted by a pragmatic process. Rooth's idea is that focus interpretation introduces a free variable whose antecedent is a discourse object which is either a subset or an element of the focus semantic value. The antecedent is fixed by the context and determines what pragmatic function the focus has: contrast, correction, answer to a question, etc. Consider (38a). The antecedent of the focus variable must be an element or a subset of the focus value *I gave x to John*. Therefore, the antecedent can be a sentence of the type in (39a) – which (38a) contrasts – or it can be a *wh*-question like (39b) – which (38a) answers.

- (39) a. You gave John a magazine.  
       b. What did you give to John?



In Brunetti (2009a) I argue that a fronted focus is defined in the same way; the only difference with respect to a postverbal focus is due to ‘focus projection’ rules of the focal accent. In some languages like Italian or Spanish, the nuclear accent marks the right boundary of the focus constituent but does not specify the left boundary, which can be the edge of any constituent containing the nuclear accent up, along the main branching of the syntactic tree, in these languages to the left (see Cinque 1993; Zubizarreta 1998). The extension of the focus to the left is therefore ambiguous, as shown in (40): B can be interpreted as having any of the foci within brackets:

- (40) Italian  
 Q: *A chi ha dato un bacio, Paolo?*  
     ‘Who did Paolo kiss?’  
 Q’: *Che cosa ha fatto, Paolo?*  
     ‘What did Paolo do?’  
 Q’’: *Che cosa è successo?*  
     ‘What happened?’  
 A: [Paolo [ha dato un bacio [a FRANCESCA]<sub>F</sub>]<sub>F</sub>]<sub>F</sub>  
     Paolo has given a kiss to Francesca  
     ‘Paolo gave a kiss to Francesca.’

When a sentence has a sentence final narrow focus, such as *a Francesca* in (40), in the absence of context it can be misinterpreted as having a wider focus. On the contrary, when a narrow focus (bearing the nuclear accent) is at the beginning of the sentence and precedes the background, given that the accent “projects” to the left, the post-focal part cannot be misinterpreted as focused. In other words, since the right edge of the focus is marked, the linguistic material following it is unambiguously interpreted as background. The syntactic configuration forces a narrow focus interpretation, and no context is needed to retrieve the focus-background partition of the sentence.<sup>18</sup>

As a consequence, focus fronting is useful when a context making the focus-background partition of the utterance transparent is not available and must be inferred. There are various cases in which this may occur:

- (41) a. when the focus answers a question that is not explicitly given in the immediately preceding discourse;  
 b. when the focus is used to correct a preceding utterance;  
 c. when the information provided by the focus goes against the speaker’s expectations (“mirative focus”, cf. Cruschina 2021).

I will illustrate each case below. An example of point (a) is (42), an Italian example from the spoken C-ORAL ROM corpus of spontaneous conversations, dialogues, and monologues in different communicative settings (Cresti and Moneglia 2005). The speaker (a goldsmith’s employee) is about to leave the goldsmith’s atelier to go and buy some material. The fronted focus in the last utterance answers the employee’s question corresponding to *fammi vedere*

<sup>18</sup> In a similar perspective, Neeleman and van de Koop (2007) propose that focus fronting facilitates a transparent mapping between syntax and information structure. By displacing the focused element to the left, focus and background are represented by two continuous constituents, as the difference between (a) and (b) shows:

- (i) a. [[XP] [ t<sub>XP</sub> ]]  
       focus background  
     b. [ [XP] ]  
       background focus

*quanti ne servono* ‘let me see how many we need’, at the beginning of the exchange. Thanks to the fronting, which makes it unambiguous that the focus of the sentence is *quaranta bianchi* ‘forty white ones’, the listener interprets the sentence as the answer to that previous question.

(42) Italian, C-ORAL ROM (Brunetti 2009a, 59)

WOM: Okay, se lui ce li ha sfusi... Fammi vedere quanti ne servono.

‘Ok, if he has them loose... Let me see how many we need.’

AND: Sì.

‘Yes.’

WOM: Ma non credo che ce li ha sfusi.

‘But I doubt he has them loose.’

AND: Questo grigio, dove l’hai preso?

‘This grey one, where did you take it?’

WOM: Dieci, dieci... mah, un mi ricordo, comunque sento. (...)

‘Ten, ten... well, I don’t remember; in any case I’ll ask.’

Io sento dove vado, intanto dal Celerini, e poi dipende dove devo andare.

‘I’ll ask wherever I go: first at Celerini’s, and then, it depends where I have to go.’

Dieci, venti, trenta, quaranta... [quaranta BIANCHI]<sub>F</sub> mi servono, all’incirca.

‘Ten, twenty, thirty, forty... forty white ones I need, roughly.’

The annotation of QUDs according to Riester, Brunetti, and De Kuthy’s (2018) model illustrates the discourse structure of this exchange. Q<sub>3</sub> is asked at the beginning of this chunk of conversation, but not answered immediately. While speaking with speaker AND about something else, which makes her answer the implicit questions Q<sub>1</sub>, Q<sub>4</sub>, Q<sub>5</sub>, and Q<sub>6</sub>, speaker WOM is trying to answer Q<sub>3</sub> by counting the number of objects she needs. Only at the end of the exchange she eventually provides the actual answer for Q<sub>3</sub>, and in doing so, she brings the discourse structure back to a higher level (>). The QUD-structure clearly shows that the question about the number of objects that she had to buy was pending, and that the question is separated by its answer by a chunk of discourse triggered by Q<sub>4</sub>.

(43) WOM: Okay, se lui ce li ha sfusi....

‘Ok, if he has them loose...

*Q<sub>1</sub>: Does he have them unpacked?*

Q<sub>2</sub>: Fammi vedere quanti ne servono

‘Let me see how many we need.’ (= *Do you allow me to see how many we need?*)

>A<sub>2</sub>: AND: Sì.

‘Yes.’

>Q<sub>3</sub>: *How many do we need?*

>A<sub>1</sub>: WOM: Ma non credo che ce li ha sfusi.

‘But I doubt he has them loose.’

Q<sub>4</sub>: AND: Questo grigio, dove l’hai preso?

‘This grey one, where did you take it?’

>A<sub>4</sub>: WOM: Dieci, dieci... mah, un mi ricordo,

‘Ten, ten... ehm, I don’t remember’

>Q<sub>5</sub>: *What will WOM do then?*

>>A<sub>5</sub>: WOM: comunque sento. [...]

‘In any case I’ll ask.’

>>Q<sub>6</sub>: *Where?*

>>>A<sub>6</sub>: WOM: Io sento dove vado,  
                   'I'll ask wherever I go'  
 >>>Q<sub>6.1</sub>: *Where first?*  
 >>>>A<sub>6.1</sub>: WOM: intanto dal Celerini,  
                   'First at Celerini's'  
 >>>Q<sub>6.2</sub>: *Where then?*  
                   WOM: e poi dipende dove devo andare.  
                   'and then, it depends on where I have to go.'  
 >A<sub>3</sub>: Dieci, venti, trenta, quaranta... [quaranta BIANCHI]<sub>F</sub> mi servono, all'incirca.  
           'Ten, twenty, thirty, forty... Forty white ones I need, roughly.'

Consider now contrast (point (b) in (41)), and more precisely correction.<sup>19</sup> When a correction is made, unlike what happens in question-answer exchanges, where the question is clearly identifiable by the interrogative form, the context utterance that the speaker wants to correct is not identified before the correction is made. Furthermore, correction is typically made by replacing one part of the utterance that one wants to correct, but there is often no marking of the part to be corrected, before the correction is made. When a correction is made by replacing part of a previous utterance, the correcting utterance contains a narrow focus. Focus fronting, by unambiguously marking a constituent of the sentence as narrow focus, thus allows the listener to identify where the correction takes place. Consider the Italian QUD-annotated example in (44), still from the C-Oral-ROM corpus.

(44) Italian, C-ORAL ROM

allora / lo sapete / ho visto un manifesto / a Firenze / indove c'era / una un ristorante indiano / in via Pisana

'So you know, I saw a poster in Florence showing an Indian restaurant, in Pisana street...'

(...)

Q<sub>1</sub> : *What does this restaurant propose?*

>A<sub>1</sub> : e in questo xxx c'è la cena

          'and in this xxx there is dinner'

>Q<sub>1.1</sub> *What does this restaurant propose on Wednesday?*

>>A<sub>1.1</sub> : e poi / dopo xxx / mercoledì sera / c'è / lo spettacolo di danza del ventre

(...)

          'and then after xxx on Wednesday night there is a show of belly dancing

(...)

>Q<sub>1.2</sub> : *What does this restaurant propose on Friday?*

>>A<sub>1.2</sub> : il venerdì sera / c'è invece sempre la cena /

          'on Friday night instead there's still dinner'

>>A<sub>1.2</sub> : e dopo c'è / la lettura della mano // c'è una / che fa / la lettura della mano (...)

          'and then there's hand reading there's someone who does hand reading

(...)

>Q<sub>1.3</sub> : *What does this restaurant propose on Saturday?*

>>A<sub>1.3</sub> : e poi invece / i' sabato sera / c'è sempre la cena

          'and then instead on Saturday night there's still dinner'

>>A<sub>1.3</sub> : e dopo la cena / fanno l'oroscopo

          'and after dinner they make the horoscope'

<sup>19</sup> Correction is not the only possible kind of contrast (cf. Repp 2016). I will discuss the different kinds of contrast in details in Section 5 below.

(...) *About 25 exchanges follow where speakers talk about astral charts*

>>Q<sub>n</sub> : *Do they have horoscope reading on Saturday ?*

>>>A<sub>n</sub> : no / scusami /

‘No, sorry’

>>A<sub>1.3</sub>''' : [[la CARTOMANTE]<sub>F</sub> c'è / i' sabato sera ]~

‘The fortune teller there is, on Saturday night’

The fronted focus allows the addressee to understand that the utterance *dopo la cena fanno l'oroscopo*, ‘after dinner they do horoscope reading’, which is uttered several exchanges before, is the alternative that is corrected by the fronted-focus sentence. The QUD structure shows that the speaker has resumed question Q<sub>1.3</sub> and has answered it once more by means of the assertion A<sub>1.3</sub>'''.<sup>20</sup>

Let us finally discuss point (c) of (41), that is, mirative focus. An Italian example is given in (45), which is another C-Oral ROM corpus used in Brunetti (2009a), and which is again annotated according to the QUD-model by Riester et al. (2018). The speaker is discussing De Niro's acting. His point is that De Niro's way of acting is strongly affected by his personality. Despite common shared knowledge (that an actor's skill consists in being able to impersonate different characters), the speaker argues that De Niro is a good actor precisely because he imposes *his own* personality on each character he plays. The fronted focus goes against the expectations that De Niro's acting skills, like those of any other actor, are to be able to transform into different characters. The demonstrative *questo* ‘this’ refers to what the speaker has just said right before, namely that De Niro always impersonates the same character in his movies. I propose that the QUD structure must include the shared inference about what an actor's skill consists of. The inference is added under the form of an implicit answer to Q<sub>2</sub>, hence as a parallel answer to Q<sub>2</sub>, together with the fronted-focus assertion.<sup>21</sup>

(45) Italian, C-Oral ROM

>Q<sub>1</sub>: *What about De Niro's acting?*

>>A<sub>1</sub>': De Niro (...) sa imporre solo se stesso, come un certo tipo di personaggio, e basta;

‘De Niro (...) is only able to impose himself, as a certain type of character, and that's all;

>>A<sub>1</sub>''': cioè, in tutti i film che ho visto si comporta e agisce sempre allo stesso modo (...);

‘that is, in all movies I've seen, he always behaves and acts in the same way (...);

>>A<sub>1</sub>'''': cioè, lui, sul set, può essere solo in quel modo (...)

‘that is, he, on set, can only be in that way (...)’

<sup>20</sup> In Section 5 I will come back to contrast and will dwell more on the various QUD-based discourse configurations that represent different contrastive types.

<sup>21</sup> A similar analysis is given by Prince (1999) of fronted foci in Yinglish, a variety of English spoken by Jewish communities in the US. Prince argues that a fronted focus is possible if the information in the background is already known to or at least plausibly inferable by the listener. For instance, in (i), the fronted focus is acceptable because it is considered to be a well-known fact in the Yiddish community that sons ask parents to buy them things.

(i) A: *Hello, Mrs. Goldberg. How's everything? How's your son?*

B: *Oy, don't ask. [A SPORTSCAR]<sub>F</sub> he wants – that's all I was missing.*

I think that this is not enough, however, to account for (iB). Indeed, the sentence expresses a negatively charged surprise. According to my analysis, the fronting indeed indicates that the answer violates some sort of shared rule of behavior, namely that sons must ask parents to buy them cheaper things than a sportscar.

- >Q<sub>2</sub>: *Where does De Niro's acting skill lie?*  
 >>A<sub>2</sub>: [*His acting skill lies [in being able to play very different characters]<sub>F</sub>~*.  
 >>A<sub>2</sub>'': ed [[in questo]<sub>F</sub> sta la sua bravura]~ .  
 'and in THIS his skill lies.'

Other examples of this kind will be discussed in Section 5.

To conclude, the different “uses” that focus fronting may have – answer an implicit question, replace the focus with some contrasting alternative, express surprise/deception for some unexpected information – are possible because the focus is unambiguously narrow and as such, it can be recognized in the absence of the appropriate context or when the context is ambiguous.

## 4 Topic and (Clitic) Left Dislocation

### 4.1 Introduction

The discussion on focus fronting is a good illustration of the debate that can arise, when studying IS-related phenomena, about the degree of encoding in the grammar – by syntax, by prosody, or by both – of informational notions such as focus and topic, but also of interpretive effects such as contrast, mirativity, exhaustivity, etc., which can be associated with focus or topic. An analogous debate arises when another set of IS-related constructions that is frequent in Romance is studied, namely (clitic) left and right dislocations, which typically are associated with topicality. This set of constructions has been the object of study of great part of my past research. Before presenting some of my work on this phenomenon, a few introductory words on the notion of topic are necessary.

The term ‘topic’ has been used in different ways in the literature. When by topic it is meant *sentence* topic, then the notion of ‘aboutness’ or of ‘address’ comes into play: a topic can be defined as “an entity or entity-type discourse referent” (McNally 1998, 147) that the sentence is about. As McNally points out (following Reinhart 1981; Vallduví 1992; Portner and Yabushita 1998; and see also Jacobs 2001), the idea of sentence topic as an entity can be viewed as deriving “from a procedure for up dating the common ground: The topic of S [=sentence, *LB*] is a ‘file card’ onto which the information contained in S is entered or a referent with which the information in S is associated (...).” (McNally 1998, 149). As an illustration, see the example from De Cat (2007) below. *La clé* denotes what Spk2 is asked to talk about, and therefore the information provided by Spk2’s sentence is about the key, or is entered onto the file-card ‘the key’.

- (46) French, De Cat (2007, 20)  
 Spk1 : Et la clé, où elle est ?  
 ‘And the key, where is it?’  
 Spk2 : La clé, je pense qu’ elle est restée dehors.  
 the key I think that CL.SG.F AUX.be.3SG.PRS remained outside  
 ‘I think that the key has remained outside.’

In the same paper, McNally views this notion of topic as problematic. As a matter of fact, topic marking can target expressions (such as quantifiers) that do not denote entities, as in the

example below, where the topic is marked in English by the so-called ‘B accent’, opposed to the focal ‘A’ accent (see Jackendoff 1972):<sup>22</sup>

- (47) English, McNally (1998:6)  
 - How many people expressed interest in your house?  
 - Well, [lots]B of people [called]A, and [three]B [looked at it]A, but [nobody]B [made an offer]A.

Following Roberts’s (1996) seminal work on the relation between information structure and the structuring of discourse into questions under discussion (see Section 2.1), McNally argues that a better way to define sentence topic is by assuming that topics are questions that structure discourse, rather than entities. The two linguistic phenomena that lead McNally to opt for a QUD-based notion of topic are the B accent in English and the *wa* morpheme in Japanese.

I agree with McNally that these topic markings do not support an account where topic is viewed as an entity about which the information is entered. I will not dwell much on the issue of the *wa* morpheme in Japanese. I will only say that, since Kuno (1973), it has been recognized that *wa* can be a marker of contrast (evoke an alternative set). As an example of proposal within the vast literature on *wa*, which seeks to reconcile its apparent double function (thematic and contrastive), Vermeulen (2013) proposes that *wa* is either a topic or a focus alternative-evoking marker (so it is not restricted to *topical* contrast), and that the aboutness function of the *wa*-phrase, when present, rather depends on its sentence initial position.

I want to discuss more at length the function of the English B-accent (or similar topical accents in other languages). In his seminal work, Büring (1997; 2003) provides an analysis of the English and the German topic accents by extending to topic the analysis made by Rooth (1992) on focus within the Alternative Semantics framework. As I recalled in Section 3.2, Rooth proposes that a sentence containing a focus feature – represented by a focal accent, as in *MARY likes Sue* – has an additional ‘focus semantic value’ that corresponds to a set of propositions of the type *x likes Sue*, where the value of the variable in each proposition is one of the possible alternatives to the focus in the relevant context. Along these lines, Büring proposes that a topic accent evokes a set of alternatives, but unlike the focus ones, they are alternative *questions*. For instance in (48), the topic feature (marked by a B accent on *female*) evokes a set of alternatives of the type: *What did the x pop stars wear?*.

- (48) English, Büring (1997: 69)  
 A: What did the pop stars wear?  
 B: The FEMALE pop stars wore caftans.

Since a question can be defined as a set of propositions corresponding to its potential answers (Hamblin 1973), it turns out that *two* sets of alternatives are at stake (*the x pop stars wear y*), as illustrated in (49), where the sets are given in curly brackets.

- (49) Büring (1997:68)  
 { {the female pop stars wore caftans, the female pop stars wore dresses, the female pop stars wore overalls,...},  
 {the male pop stars wore caftans, the male pop stars wore dresses, the male pop stars wore overalls,...},  
 {the female or male pop stars wore caftans, the female or male pop stars wore dresses, the female or male pop stars wore overalls,...},

---

<sup>22</sup> But see Endriss (2009) for an account of quantifier topics based on the notion of aboutness.

{the Italian pop stars wore caftans, the Italian pop stars wore dresses, the Italian pop stars wore overalls,...},...}

The difference between the focus and the topic variable lies on their effects on discourse and can be accounted for by appealing to a QUD-based discourse model. As we have seen in Section 2.2.1, the QUD structure of cases like (48) proposed by Riester et al (2018), which is based on Büring (2003), is such that the utterance containing the topic accent answers a sub-question, namely it *partially* answers the current QUD ( $Q_1$  in the illustration below).

- (50)  $Q_1$ : What did the pop stars wear?  
 $>Q_{1.1}$  : *What did the female pop stars wear ?*  
 $>>A_{1.1}$ : [The FEMALE pop stars wore [CAFTANS]<sub>F</sub>]~

The value of the focus variable is given by the part of the sentence that provides an answer to the sub-QUD ( $Q_{1.1}$ ), while the value of the topic variable is provided by a part of the sub-QUD itself (*female* in  $Q_{1.1}$ ), which is also part of the background of its answer (see  $A_{1.1}$ ). According to this analysis, the topic accent introduces a member of the set mentioned in the higher QUD (the set of pop stars). Notice that this member was not mentioned before, but constitutes new information: indeed, the sub-QUD is not formulated according to Q-GIVENNESS, but according to the principle of (complex) PARALLELISM, which allows for new content to be present in the sub-QUDs.

As a matter of fact, things are more complicated than that, since it has been observed that the topic accent might mark some more general phenomenon than the one described above. The same intonation that is present in the answer to the sub-question in (50) can be found in French in exchanges like (51) and (52), from Beyssade (2017), where the French topic accent (an initial rise according to Marandin et al. 2002) does not indicate a partial topic. In these examples, the accent is required to make the exchange coherent and seems to signal a more general “complex discourse strategy” (Marandin et al. 2002, 1).

- (51) (Beyssade 2017, 179)  
 Spk1: Qui a répondu à l’invitation ?  
 ‘Who answered the invitation?’  
 Spk2: La POSTE est en grève.  
 ‘The post office is on strike’  
 Spk1: Ok, on annule.  
 ‘Ok, then we cancel it’
- (52) (Beyssade 2017, 180)  
 Spk1: Jean-Marie est arrivé ?  
 ‘Did Jean-Marie arrive ?’  
 Spk2: Oui. BERNADETTE n’est pas contente.  
 ‘Yes. Bernadette is not happy (about it)’

More precisely, Beyssade (2017) proposes that the topic intonation is used by the speaker to indicate that, despite appearances, she has more to say about the question asked. In other words, the accent signals thematic continuity. I would rather suggest to take the opposite perspective to explain its function. Rather than indicating thematic continuity (which, incidentally, usually is not made prominent by linguistic material or intonation), I suggest that the special intonation on *poste* and *Bernadette* signals discourse *discontinuity*, namely the fact that a different QUD is answered by the utterance than the one that would be derived by the application of the

principles seen in Section 2.2. Specifically, the intonation indicates what part of the QUD is new content (against Q-GIVENNESS).

I agree with Beyssade that, in (51) and (52), there are no plausible alternatives to *la poste* or to *Bernadette* respectively, so Büring's proposal of the topic accent as an alternative-set-evoking device does not seem to hold here. Nevertheless, these cases still share with Büring's examples the fact that the QUD preceding the utterance contains some new information, as seen when discussing (48). In (53), 'the post-office' in  $Q_2$  is new information, and that is why *la poste* in the answer ( $A_2$ ) is intonationally prominent.  $Q_1$ 's answer is inferred via Grice's Maxim of Relevance; if one accepts to include it in the QUD-tree, it will be as an implicit assertion (in grey below).<sup>23</sup> In (54),  $A_2$  means 'Jean-Marie arrived', but the following utterance answers a question about Bernadette, who therefore represents new content.

- (53)  $Q_1$ : Qui a répondu à l'invitation ?  
           'Who responded to the invitation?'  
        $>Q_2$  : *What about the post-office (through which the responses to the invitation are sent) ?*  
        $>>A_2$  : [La poste [est en grève]<sub>F</sub>]~ .  
               'The post office is on strike'  
        $>A_1$  : *Nobody.*
- (54)  $Q_1$ : Jean-Marie est arrivé ?  
           'Did Jean-Marie arrive ?'  
        $>A_1$  : Oui.  
               'Yes'  
        $>Q_2$  : *How does Bernadette feel about Jean-Marie's arrival ?*  
        $>>A_2$  : [Bernadette [n'est pas contente]<sub>F</sub>]~.  
               'Bernadette is not happy (about it)'

In conclusion, going back to the discussion based on McNally (1998) about the notion of topic and the utility of an entity-based definition of sentence topic, we have seen that the two phenomena McNally mentions are not apt to support a notion of topic as an entity the sentence is about. These markings rather seem to signal a *change* of topic, as in (53) or (54), and the topic is not necessarily an entity, even though in the examples shown above it is.

A construction that seems more promising for an entity-based analysis of topic is (clitic) left dislocation (CILD). Indeed, this construction divides the sentence into two parts, one of which generally is a referential expression. I will discuss CILD in the section below, where we will see that in fact, its association with an entity-like notion of topic is less straightforward than one might expect.

## 4.2 (Clitic) Left Dislocation

(Clitic) left dislocation exists in several languages that belong to different typological families. It is a frequent construction in Romance languages (particularly in their spoken variety),<sup>24</sup> and it has received much attention in the last decades, either from a syntactic, a pragmatic, or a prosodic point of view (for Romance languages see, among many others, Ashby 1988; Lacheret and François 2003; E. Delais-Roussarie, Doetjes, and Sleeman 2004; Tellier and Valois 2006; De Cat 2007; Avanzi 2012 for French; Vallduví 1992; Villalba 2000; López 2009 for Catalan

<sup>23</sup> Cf. ex. (45) at the end of Section 3.2, and see more on this choice in Section 5.

<sup>24</sup> See Villalba (2000) for examples from a large variety of standard and non-standard Romance languages.



and Spanish; Benincà, Salvi, and Frison 1988; Cimmino 2016 for Italian; and see my own work on Italian and Spanish in Brunetti 2009c; 2009b; and on Catalan in Brunetti, Mayol, and Villalba 2020). CILD consists of the detachment to the left periphery of the sentence of a constituent (typically an argument or an adjunct of the verb) that is “resumed” within the core sentence by a coreferential clitic pronoun, as in the Italian constructed example in (55):<sup>25</sup>

- (55) Il libro sulle interrogative, l’ho dato alla studentessa.  
 the book on.the interrogatives CL.SG-have.1SG given to.the student  
 ‘I gave the book on interrogatives to the student’

CILD has been described as marking the sentence topic, either an *aboutness* topic (Reinhart 1981; Lambrecht 1994; Vallduví 1992; Jacobs 2001) or a *contrastive* one (Arregi 2003; López 2009). In Brunetti (2009c), I propose that CILD in Italian represents a ‘link’ in the sense of Vallduví (1992) (that is, it points to a “file card” onto which the information of the sentence is entered). Therefore, in this work I give an entity-based account of topic. However, I also try to show that this function has the pragmatic effect of evoking alternatives in contexts where the CILDed expression refers to the current discourse topic (that is, when it is a continuous topic, cf. Givón 1983) and therefore it is not expected to be expressed by a full NP.

To better understand my claim, consider the constructed example in (56). The repetition of *a Dante* in the answer has the effect of evoking other people to whom the speaker might give a present, among a contextually relevant set. I explain this effect by appealing to Grice’s Maxim of Manner: the lexical realization of the continuing topic triggers the inference that the speaker intends to compare that topic with some contextually relevant alternative.<sup>26</sup>

- (56) (adapted from Brunetti 2009c, 764)  
 Two friends are talking about a common friend, Dante.  
 A: Che cosa gli regalerai per Natale?  
 ‘What will you give him for Christmas?’  
 B: A Dante (gli) regalerò un LIBRO.  
 to Dante CL.DAT.SG.M give.1SG.FUT a book  
 ‘I’ll give a BOOK to Dante’

Within a QUD-based discourse model, the presence of the CILD triggers the inference that by uttering (56B), the speaker intends to answer a larger question, such as: *What will you give to your friends for Christmas?* of which (56B) is a partial answer. Most probably, in fact, a yet more general function can be attributed to CILD, as we have seen above for the topic accent, namely the function of signaling that a different QUD is at stake than the explicit question, no matter whether an alternative set is involved or not. This analysis, if correct, would make the function of CILD parallel to that of topic accents in English or French. Yet this analysis does not take into account intonation, therefore we cannot be sure that the marking of a change of QUD is due to the presence of the CILD construction or to the intonational pattern that goes with it.

In Brunetti (2009b) I propose that CILD in Spanish and in Italian is a way to front a constituent, so that the constituent is recognized as a topic. I argue that it is the sentence initial position that makes the constituent a topic, not the CILD-ed position. I also claim that fronting occurs so that an expression different from the subject and that does not have prototypical topic properties can be promoted to a position where it is recognized as a topic. This happens when

<sup>25</sup> Clitic *Right* Dislocation (CIRD), where the constituent is detached to the right, has similar though not identical syntactic properties and pragmatic functions. I will not discuss them in the present work.

<sup>26</sup> This idea is already in Vallduví 1990, see Villalba (2000, 77).

that expression must be the topic for some pragmatic needs, specifically when contrast must be expressed. In fact, when one looks under the light of the QUD model at the examples provided in Brunetti (2009b) (extracted from the spoken Nocando corpus), once more the idea seems to be confirmed that a CILD signals a different QUD than the one expected in that context. Consider one of such examples, given in (57), and its QUD annotation in (58). The speaker is narrating the story of a frog that creates disarray in a restaurant.

- (57) Spanish, Nocando corpus (Brunetti 2009b)
- a. La rana los                      está mirando, feliz,  
the frog CL.ACC.PL.M    is    watching happy
  - b. dándose cuenta de que eso lo              ha hecho ella.  
realizing                  of that that CL.ACC.SG.M    has done her  
'The frog is watching them, happy, realizing that HE was the cause of (all) that.'
- (58) *Q<sub>1</sub> What about the frog?*  
>A<sub>1</sub>' : La rana los está mirando,  
          'the frog is watching them,  
>*Q<sub>2</sub> : What are the frog's feelings when he is looking at them?*  
>>A<sub>2</sub> : feliz,  
          'happy'  
>A<sub>1</sub>' : dándose cuenta de que eso lo ha hecho ella  
          'realizing that HE was the cause of (all) that'

The two utterances provide information about the frog, but clearly this annotation cannot be the end of the story, since the question in *Q<sub>1</sub>* cannot account for the presence of a CILD in the subordinate clause.<sup>27</sup> The subordinated CILD *de que eso lo ha hecho ella* 'that HE did that' must then be separated from its matrix clause, so that it answers its own QUD.<sup>28</sup> Note that such a QUD is not, as one might expect given Riester et al's principles of QUD annotation, *What does the frog realize?*, but an unexpected question on the subject, which is compatible with the CILD construction (the focus bears the sentence final accent and the direct object is part of the background):

- (59) *Q<sub>1</sub> What about the frog?*  
(...)  
>A<sub>1</sub>' : [[dándose cuenta]<sub>F</sub>]~  
          'realizing'  
>*Q<sub>3</sub> : Who did (all) that (mess) ?*  
>>A<sub>3</sub> : [de que eso lo ha hecho [ELLA]<sub>F</sub>]~  
          that HE was the cause of (all) that.'

<sup>27</sup> In Riester et al (2018) we claim that within our model, linguistic form can be dispensed with and that the QUD is only determined by linguistic context. We add, however, that there are exceptions and that "the annotator is encouraged to narrow down the question as soon as language-specific rules have been established with a reasonable degree of certainty" (Riester et al 2018 :416). This precisely corresponds to a case like (58), where the CILD does not leave much space for doubt that the QUD that the subordinate clause answers is not the one predicted by the context.

<sup>28</sup> This is a violation of another rule stated in Riester et al (2018), namely that only adjuncts, not complements, can constitute independent segments. Note however that the matrix verb is *realize*. According to Hooper and Thompson (1973), with the class of verbs to which *realize* belongs, which "assert the manner in which the subject came to know that the complement proposition is true" (Hooper and Thompson 1973, 480), the complement clause may replace the matrix clause as the complex sentence's *main assertion*. If that is correct, then the separation in (59) of the subordinate clause from the matrix verb is fully justified.

Another Italian example is given in (60), with its QUD structure in (61). The implicit argument in the two sentences is the boy who is the owner of the frog, in the same story as the example above. The boy is forced to leave the restaurant because of the disarray that the frog has created, without the boy's knowledge of it.

- (60) Italian, Nocando corpus  
 a. Della cena non gli interessava, però  
     of.the dinner not CL.DAT interested but  
 b. la rana, poverina, l'aveva quasi persa  
     the frog poor.thing CL.ACC.SG-had almost lost  
     ‘As for the dinner, he didn't care, but the poor frog, he had almost lost it!’
- (61) *Q<sub>1</sub>: What about the boy?*  
 >A<sub>1</sub>: E il bambino, poverino era (...) il più dispiaciuto,  
     ‘And the boy, poor thing, was the sorriest’  
 >*Q<sub>2</sub>: About what was the boy sorry?*  
 >*Q<sub>1.1</sub> Was the boy sorry about the dinner?*  
 >>A<sub>1.1</sub>: [[Della cena]<sub>CT</sub> [non gli interessava]<sub>F</sub>]~,  
     ‘about the dinner, he didn't care’  
 >*Q<sub>1.2</sub> Was the boy sorry about the frog?*  
 >> A<sub>1.2</sub>: però [[la rana, poverina]<sub>CT</sub>, [l'aveva quasi persa]<sub>F</sub>]~ [=he was sorry  
     about the frog]  
     ‘but the poor frog, he had almost lost it!’

The QUD structure is derived from complex PARALLELISM (cf. Section 2.2.1) for an OPPOSE type of contrast. As we will see in Section 5, contrast can be of different kinds and the OPPOSE type requires polar QUDs, namely QUDs whose answer has focus on the verb, and the two contrasting assertions must have opposite polarity. In (61), the first assertion explicitly states that the boy was *not* sorry for having to stop the dinner, while the second implicitly means that the boy *was* sorry for almost losing the frog.<sup>29</sup> By removing the complement to the left, the CILD construction allows for the verb to occupy a sentence final position and therefore to be focused. The CILD also makes the addressee understand that the dinner and the frog are new content that is part of the QUDs that A<sub>1.1</sub> and A<sub>1.2</sub> answer respectively, unlike what predicted by the preceding context.

In conclusion, what we gather from these examples is that a CILD helps formulating a QUD – and therefore identifying a certain informational partition of the sentence – when such a QUD does not correspond to what is expected given the principles of QUD formulation stated in (14). In other words, the CILD is used to mark an unexpected QUD.<sup>30</sup>

The analysis of CILD as a QUD shifter does not seem to be supported by two examples of polarity focus given in Garassino and Jacob (2018). Garassino and Jacob provide corpus data

<sup>29</sup> I will return on polar QUDs in Section 5. Polar QUDs are not used in Riester et al's (2018) model; they are introduced in Brunetti (2024) in order to account for certain kinds of contrast.

<sup>30</sup> In Brunetti, Mayol, and Villalba (2020) we look at the relationship in Catalan between the position of a direct object – dislocated to the left, to the right, or in situ – and the strength of the anaphoric link with its antecedent. Results of two acceptability-judgment tasks show that a decrease of ratings from the strongest to the weakest bridging type is present with CIRD; however, CILD is not as much affected by the bridging strength. This result is in line with the idea just presented that CILD's main function is to express either a change of topic (that is, a change of QUD) or contrast, namely a specific case of changing topic where the topic is a member of a set of alternatives (still a change of QUD, but occurring in a special discourse configuration as the one described by Büring's work and Riester et al's complex PARALLELISM).

from various languages showing that the CILD provides “a positive answer to the polar QUDs raised in the previous context” (Garassino and Jacob 2018, 18). Consider (62):

(62) Italian (Garassino and Jacob 2018, 16)

Dovremmo, credo, dare maggiore informazione sugli strumenti a disposizione, valutare le cose positive che sono state fatte - anche se noi non siamo contente fino in fondo, perché ancora resta molto da fare - ma non dobbiamo abbatterci per il fatto che non abbiamo strumenti a disposizione:

‘We must, I think, provide greater information on the tools available, emphasize the positive things that have been done - even if we are not completely happy because there is still a lot to do - but we must not become dejected that we do not have any tools at our disposal:’

gli strumenti li                      abbiamo  
the tools      CL.ACC.PL   have.1PL  
‘we *have* the tools’

The QUD structure that I propose for this example is given below. The CILD here does not seem to be justified by any unexpected change of QUD. The polar QUD that the CILD answers is formulated according to Riester et al’s principles. the CILD seems to simply mark the fact that the speaker is making a statement of opposite polarity with respect to the preceding utterance. According to Garassino and Jacob, the CILD also signals topic continuity, namely that the dislocated expression denotes an entity that is already “under discussion”, already mentioned before in the discourse.

(63) *Q<sub>1</sub>: What tools should we develop?*

(...)

>*Q<sub>2</sub>: What should we do, once we have such tools?*

>>*A<sub>1</sub>*: Dovremmo, credo, dare maggiore informazione sugli strumenti a disposizione,

‘We must, I think, provide greater information on the tools available’

>>*A<sub>1</sub>*’: valutare le cose positive che sono state fatte

‘appreciate the positive things that have been done’

>>*Q<sub>2</sub>: Despite what ?*

>>>*A<sub>2</sub>*: anche se noi non siamo contente fino in fondo,

‘even if we are not completely happy’

>>>*Q<sub>3</sub>: Why ?*

>>>>*A<sub>3</sub>*: perché ancora resta molto da fare –

‘because there is still a lot to do –’

>>>>*Q<sub>4</sub>: Given that, should we become dejected that we do not have any tools?*

>>>>>*A<sub>4</sub>*: ma non dobbiamo abbatterci per il fatto che non abbiamo strumenti a disposizione:

‘but we must not become dejected that we do not have any tools at our disposal’

>>>>>*Q<sub>5</sub>: Do we have any tools?*

>>>>>> *A<sub>5</sub>*: [gli strumenti li [abbiamo]<sub>F</sub>]~  
‘we *do* have the tools’

In (64), still from Garassino and Jacobs, the dislocated elements are two: *noi* and *questo sforzo*.

(64) Italian (Garassino and Jacob 2018, 17)

Per riuscire è indispensabile tener fede a un principio che è alla base del nostro stare nell'Unione europea. (...) È quello secondo il quale nello sviluppo della costruzione europea occorre sempre fare uno sforzo per comprendere le ragioni degli altri, farsene in qualche modo carico.

‘To be successful in this, we have to hold faith with a principle that lies at the basis of our existence as the European Union (...). It is the principle whereby, in furthering European integration, it is essential always to make an effort to understand the arguments of the other side and in some way to take them on board.’

Noi questo sforzo lo abbiamo sempre fatto e continueremo a far-lo  
 we this effort CL.ACC.SG have.1PL always made and continue.FUT.1PL to make-CL.ACC.SG  
 ‘We always made this effort and will continue to make it’

The QUD structure that I propose is the following:

(65) ‘To be successful in this, we have to hold faith with a principle that lies at the basis of our existence as the European Union (...)’

*Q<sub>1</sub> : What principle ?*

>A<sub>1</sub>’ : È quello secondo il quale nello sviluppo della costruzione europea occorre sempre fare uno sforzo per comprendere le ragioni degli altri,

‘It is the principle whereby, in furthering European integration, *it is essential always to make an effort to understand the arguments of the other side*’

>A<sub>1</sub>’’ : farsene in qualche modo carico.

‘and in some way to take them on board’

>Q<sub>2</sub> : *Do EU countries make this effort?*

>>Q<sub>2.1</sub> : *Does Italy (=we) make this effort ?*

>>>>A<sub>2.1</sub>: [[Noi]<sub>CT</sub> questo sforzo lo [abbiamo sempre fatto]<sub>F</sub>]~

‘We always made this effort’

While the dislocated *noi* ‘we’ indeed signals a sub-QUD that contains new semantic content, since the question is about one of the EU members, Italy, which is not mentioned before, ‘this effort’ is expected to be part of the QUD, since it is mentioned in A<sub>1</sub>’. The dislocation therefore only seems to mark the polarity focus by removing non-focused material from the part of the sentence where the focal accent falls.

## 5 Contrast

### 5.1 Introduction

As we have seen throughout the preceding discussion, the notion of contrast is often relevant when focus and topic phenomena are discussed, since both focus and topic markings may trigger a contrastive interpretation. We have also seen that focus and topic, which have always been described as phenomena occurring at the sentence level, are in fact strongly related to discourse structuring. Contrast is no exception: it cannot be accounted for without appealing to discourse and discourse structuring. In Brunetti (2024) I assume that contrast is a phenomenon that arises between two discourse segments and, following Umbach (2004) and Repp (2016) (among others), that it must be tackled from the angle of (contrastive) discourse relations.

Furthermore, I make the different types of contrastive relations to derive, at least in part, from the QUD structure in which contrast occurs.

By definition, contrast requires that two or more members of an alternative set (that is, elements that share similarities and dissimilarities) be compared, and we have seen that both focus and topic may evoke an alternative set. However, the presence of alternatives is not sufficient to get a contrastive interpretation: certain contextual or semantic constraints must apply. For Rooth (1992), for instance, contrast arises in situations where the linguistic context explicitly provides a phrase/sentence whose semantic value corresponds to one of the alternatives evoked by the focus. Various scholars take it as crucial that the contrasting alternatives be clearly identifiable in the context; others claim that they should be limited in number; others that they should be both (see Molnár 2002 for a survey). Neeleman and Vermeulen (2013) propose a constraint for contrastive focus, namely that its meaning must encode at least one false alternative.<sup>31</sup> For contrastive topic, they propose that the additional interpretive effect is that the speaker is unwilling to utter (at least) one alternative.<sup>32</sup>

As I said above, I account for contrast – and its different kinds – in terms of (contrastive) discourse relations between segments. I look at these relations – and at the link between them and focus/topic – through the lenses of the QUD structure in which the contrasting segments occur. In particular, Riester et al’s (2018) principle of (simple and complex) PARALLELISM will be relevant to identify the QUD structure of the contrasting segments.

## 5.2 Contrastive relations

Discourse relations are interpretative relations between utterances (see Hobbs 1985; Mann and Thompson 1988; Webber et al. 2019, among many others). For instance, an utterance can express the cause, the result, or the goal of what is described in the preceding utterance(s); it can elaborate on or explain what is said in the preceding utterance(s), etc. Discourse relations can be marked by discourse markers or subordinating/coordinating conjunctions, such as ‘therefore’, ‘however’, ‘in order to’, ‘but’, ‘yet’, ‘because’, etc. Among discourse relations, various types of contrastive ones have been proposed, and analyses differ with respect to their number and their exact definition, as well as their names (see Repp 2016 for a summary). In Brunetti (2024), I borrow from Repp (2016) the classification of contrastive relations into SIMILAR, OPPOSE, and CORR(ection), and I add a CONCESSION relation. Below I will provide some examples of each of them and discuss their properties in the light of the QUD structure.

The first relation discussed by Repp (2016) is SIMILAR, which has also been called PARALLEL or LIST in the literature. Indeed, its main function is to compare or list the pieces of information provided by two or more discourse segments. More precisely, SIMILAR is a relation where the contrasting segments “make the same kind of contribution to the current question under discussion” (Repp 2016 :8). In (66), for instance, A<sub>23</sub>’ and A<sub>24</sub>” provide two answers to the QUD ‘Where specifically, within the 7<sup>th</sup> district?’. Notice that the shared content is (almost) identical. That helps to identify the two contrasting segments.

- (66) French, CFPP2000 corpus (7<sup>ème</sup>), Brunetti (2024:18)  
 A<sub>23</sub> : alors (...) j’aime beaucoup le septième euh où qu’ce soit  
           ‘so ehm I love very much (...) the 7th district, wherever it may be’  
 Q<sub>24</sub> : *Where specifically, within the 7th district?*  
 >A<sub>24</sub>’ : [que ça soit [vers la Tour-Maubourg]<sub>F</sub>]~  
           that DEM be towards the Tour-Maubourg

<sup>31</sup> In fact, they discuss *corrective* focus only. See more on other types of contrast below.

<sup>32</sup> See the discussion on ‘purely implicational’ topic in the next section.

‘be it towards the Tour-Maubourg’  
 >A<sub>24</sub>” :[que ce soit [ici]<sub>F</sub>]~  
 that DEM be here  
 ‘be it here’

SIMILAR can also occur with *complex* PARALLELISM, as is (67):

- (67) French, CFPP2000 corpus (13<sup>ème</sup>), Brunetti (2024:6)  
 > Q<sub>1</sub> Spk1: euh comment est-c’que euh toi ou tes parents vous êtes arrivés dans l’quartier (...)  
 ‘How did you or your parents arrive in the district (...)?’  
 >> Q<sub>1.1</sub> : *How did you arrive in the district?*  
 >>> A<sub>1.1</sub> : Spk2: alors, donc [[moi]<sub>CT</sub> j’suis arrivé à Paris [j’étais tout petit]<sub>F</sub>]~ (...)  
 so then me I am arrived at Paris I was all little  
 ‘well, as for myself, I arrived in Paris as a little child’  
 >> Q<sub>1.2</sub> : *How did your parents arrive in the district?*  
 >>> A<sub>1.2</sub> : et [[mes parents]<sub>CT</sub> sont venus à Paris [pour le boulot]<sub>F</sub>]~  
 and my parents are come to Paris for the job  
 ‘and my parents came to Paris for work’

The two utterances are two partial answers to the (here, explicit) superordinate question ‘In what circumstances did your family arrive in the district?’. The CILDED strong pronoun *moi* ‘me’ in A<sub>1.1</sub> is recognized as one of the two topic alternatives (*moi* vs *mes parents*). Syntactic identity between the background part of the two segments (*j’suis arrivé à Paris* ‘I came to Paris’ / *mes parents sont venus à Paris* ‘my parents came to Paris’) helps to identify the SIMILAR relation by pointing out their shared content.

SIMILAR can be seen as the “vanilla” contrastive relation: it has the minimal requirements for two segments to be recognized as contrastive, that is: shared content and a syntactic (coordination) and discourse structure (PARALLELISM) that favors the comparison between them. The relation OPPOSE is enriched with an additional requirement: not only must the two segments answer the same QUD, but they also must make opposing contributions to it. My definition of OPPOSE is narrower and rather different than the one given by Repp. By ‘opposing contributions’ Repp means that “the first conjunct serves as an argument for some background assumption whereas the second conjunct serves as an argument against it”. In other words, OPPOSE for Repp implies violated expectations (Repp 2016:8). I do not assume that OPPOSE requires violated expectations, but propose that what is crucial for this relation is that each sentence makes a predication of opposite polarity with respect to the other one. A complex PARALLELISM configuration is required with OPPOSE, so that the predications of opposite polarity can be made about two different entities and the two statements are not contradictory. Consider (68). The speaker is explaining that he did not find it difficult to find an apartment of social housing, because the apartments with a suitable price for him – namely those of an intermediate price – were easy to find. (Recall that CT means ‘contrastive topic’ and it indicates the part of the sub-QUD that is new with respect to the preceding context.)

- (68) French, CFPP2000 spoken corpus (13<sup>ème</sup>), Brunetti (2024:19)  
 >> Q<sub>42</sub> : *Is the waiting time long to get an apartment of social housing?*  
 >>> Q<sub>42.1</sub> : *What is the waiting time to get really affordable social housing?*  
 >>>> A<sub>42.1</sub> : c’est-à-dire que [[les HLM vraiment pas chers là]<sub>CT</sub> [y a des queues et une attente incroyable]<sub>F</sub>]~

‘That is, for really cheap social housing, there are never-ending lines and an incredibly long waiting time’

>>> Q<sub>42.2</sub> : *What is the waiting time to get intermediate cost social housing?*

>>>> A<sub>42.2</sub> : [[dans la gamme intermédiaire]<sub>CT</sub> comme ici [on a un petit peu plus de chance euh d’aboutir]<sub>F</sub>]~

‘in the intermediate (price) range (for apartments of social housing), like here, we have a little more chance to succeed’

Contrast is interpreted as OPPOSE because on one hand, *y a des queues et une attente incroyable* ‘there are never-ending lines and an incredibly long waiting time’ obviously means that the waiting time is long; on the other hand, *on a un petit peu plus de chance euh d’aboutir* ‘we have more chance to succeed’ implies that the waiting time is *not* long. In other words, a predication of opposite polarity is inferred in A<sub>42.2</sub>.

Note that Q<sub>42</sub> is a polar question. In previous work on the QUD model by Riester, De Kuthy, and myself, no use is made of polar QUDs. Indeed, when polar questions have a narrow focus on some constituent, they have the same function as wh-questions, since they help to identify the narrow focus in the corresponding answer:

- (69) - Did John give a JACKET to Mary? / What did John give to Mary?  
- He gave her [a HAT]<sub>F</sub>.

However polar questions, unlike wh-questions, can focus on the polarity of the sentence:

- (70) - Is the waiting time long (or not long)?  
- The waiting time is NOT long. / The waiting time IS long.

It is for that purpose – that is for cases of an OPPOSE relation – that in Brunetti (2024) I introduce polar questions as QUDs.

The relation OPPOSE can be inferred by linguistic marking. We have seen above that a CILDED expression may be contrastive, though not all the time: either it can just express a change of QUD that does not evoke an alternative set (cf. (58), (60)); or because the CILD is just used to remove an argument from the sentence final position, so that the verb can bear the focal accent (cf. (63), (65)). Nevertheless, a contrastive interpretation of a CILD seems the one most naturally inferred in the absence of contextual cues that indicate otherwise. Such case is described by Büring (1997) under the name of ‘purely implicational’ topic: the topic has no explicit alternative in the discourse, but an inference is made, which is taken to be a conversational implicature (Grice 1975) by Lambrecht (1994), Büring (2003), and others, either that the same predicate does *not* hold for the implicit contrasting topic (strong implicature), or that the speaker does not know whether the same predicate holds or not for that implicit contrasting topic (weak implicature, cf. Neeleman and Vermeulen 2013).<sup>33</sup> In other words, an OPPOSE contrasting relation with an implicit discourse segment is inferred. Consider the example from Garassino and Jacob (2018) below:

- (71) French, Garassino and Jacob (2022: 9)  
a. Monsieur le Président, (...) il faut rendre honneur à la présidence française, il faut rendre honneur au président Chirac, qui a été au charbon, qui a combattu et qui a vaincu sur sa vision de l’Europe

---

<sup>33</sup> Hara (2006), who analyzes how this interpretive effect is triggered by the Japanese particle *wa*, takes it to be a *conventional* implicature.



‘Mr. President (...) we should honor the French Presidency, we should honor President Chirac. He was at the coalface, he fought and conquered for his vision of Europe’

- b. parce que [[lui]<sub>CT</sub> il a une vision]<sub>F</sub>]~  
 because he CL.NOM.3SG has a vision  
 ‘because *he* does have a vision’

The dislocated *lui* ‘he’ (*lui il a une vision*) triggers the inference that the other politicians do *not* have a vision. The CLLD is thus used as a rhetorical strategy to have an OPPOSE relation inferred by the addressee.<sup>34</sup>

Moving now to the CORR(ECTION), this relation occurs when two utterances are not compatible, namely one is negated to assert the other, as in (72).

- (72) French, Rhapsodie (Mertens corpus) (Brunetti 2024:15)  
 >Q<sub>2</sub> : *What about your father ?*  
 >>A<sub>2</sub>': Spk1: [votre père [était riche]<sub>F</sub>]~  
 ‘Your father was rich’  
 >>Q<sub>3</sub>: *Was he rich?*  
 >>>A<sub>3</sub>: Spk2: [riche [c’est un grand mot]<sub>F</sub>]~ [=he was not rich]  
 ‘rich’ is a big word’  
 >>A<sub>2</sub>": mais enfin disons qu’[il [appartenait à cette bourgeoisie euh qui n’a pas de problèmes d’argent]<sub>F</sub>]~  
 ‘but well let’s say that he belonged to that bourgeoisie ehm who does not have money problems’

The writer corrects the journalist who says that her father was rich, by explaining that ‘rich’ is not the right word to define her father’s financial situation. We can interpret the speaker’s comment in A<sub>3</sub> as a denial of the interviewer’s statement that her father was rich, while statements A<sub>2</sub>' and A<sub>2</sub>", which both answer Q<sub>2</sub>, constitute the two contrasting segments (one correcting the other).

We see in this example, that the two contrasting alternatives are explicitly mentioned, each being uttered by a different speaker. In Brunetti’s (2024) data, however, the relation CORR mostly occurs between an explicit utterance and an implicit one, corresponding to some shared assumption between speaker and hearer whose truth the speaker denies. Consider (73), from the same French interview. The writer is talking about her poor childhood. (We assume that *ce qui est dur* ‘what is hard’ and *ce qui est horrible* ‘what is horrible’ are in this context synonymous: the speaker is simply varying her language for stylistic reasons):

- (73) French, Rhapsodie (Mertens corpus) (Brunetti 2024:16)  
 >A<sub>7</sub> : c’est une expérience ça que je n’ai jamais oubliée  
 ‘That is an experience I’ve never forgotten’  
 >Q<sub>8</sub> : *What is hard, in this experience?*  
 >>A<sub>8</sub>': [~~What is hard is [to be poor]<sub>F</sub>]~~~  
 >>Q<sub>9</sub> : *Is it hard to be poor?*  
 >>>A<sub>9</sub> : [ce qui est dur [ (...) ce n’est vraiment pas]<sub>F</sub> d’être pauvre (...)]~  
 ‘What is hard (...), it is not really to be poor’  
 >>A<sub>8</sub>": [ce qui est horrible [c’est de se dire je n’en sortirai jamais]<sub>F</sub>]~  
 ‘What is horrible, it is to tell oneself: “I’ll never get out of it”’

<sup>34</sup> See also the discussion about example (56).

The speaker is correcting the shared assumption (given world knowledge) that being poor is a hard experience. The CORR relation is between that implicit statement (*A<sub>8</sub>'*) (written in italics with a strike-through, and worded in English) and the explicit answer *A<sub>8</sub>'* ('It is hard to tell oneself: I'll never get out of it'). The implicit answer is easily reconstructed thanks to the denial of its truth in *A<sub>9</sub>* ('What is hard is not to be poor'). Notice that the implicit positive statement (*A<sub>8</sub>'*) and the statement that replaces it (*A<sub>8</sub>'*) are both answers to the same QUD (*Q<sub>8</sub>*).

Summarizing, the difference between OPPOSE and CORR lies in that OPPOSE only concerns segments in a complex PARALLELISM configuration, where the same speaker asserts two utterances of opposite polarity (without being contradictory, since the two utterances answer different sub-QUDs). CORR(ECTION), on the contrary, concerns segments in a simple PARALLELISM configuration, which provide different and incompatible answers to the same QUD (if one is true, the other is not). Notice that the QUD they answer is not a polar question, since the correcting segment replaces the focus constituent of the first one.

Finally, CONCESSION is a relation such that one conjunct goes against the expectations triggered by the other, in that it contradicts the default inference coming from the other conjunct (Umbach 2004). In Webber et al's (2019) Penn Discours TreeBank annotation guidelines, "Concession is meant to be used when a causal relation expected on the basis of one argument is cancelled or denied by the situation described in the other" (Webber et al 2019:23). An Italian example is (74). The CONCESSION relation is lexically marked by *anche se* 'even if'.

- (74) Italian, blog interview (*Bilingue per gioco*), Brunetti (2024:17)  
 >Q<sub>10</sub> : *What about your linguistic background?*  
 >>A<sub>10</sub>' : Io (per fare un esempio) parlo fluentemente inglese,  
           'As an example, I speak English fluently'  
           (...)  
 >> A<sub>10</sub>' : sono senz'altro bilingüe  
           'I am definitely bilingual'  
 >>A<sub>10</sub>' : nel senso che [[ho pieno controllo di due codici linguistici]<sub>F</sub>~,  
           in.the sense that have.1SG full control of two codes linguistic  
           'in the sense that I have full control over two linguistic codes'  
 >>A<sub>10</sub>' : anche se [[non sono bilingue precoce]<sub>F</sub>~.  
           even if not am bilingual early  
           'even though I am not an early bilingual'

The speaker is arguing that having full master of two languages and therefore being bilingual does not necessarily mean to have acquired both languages early in life; on the contrary, and against what one might expect, someone who speaks fluently two languages may not be an early bilingual. In order to provide an example, the speaker says that she speaks fluent English though she did not acquire it early in life. The contrasting utterances are the last two in the example: the first one (*ho pieno controllo di due codici linguistici* 'I have full control over two linguistic codes') triggers the inference contradicted by the second (*non sono bilingue precoce* 'I am not an early bilingual').<sup>35</sup>

An example of CONCESSION with complex PARALLELISM is (75), where the speaker talks about his apartment.

- (75) French, CFPP2000 corpus (13<sup>ème</sup>)

<sup>35</sup> Notice that in this example, there are four utterances that answer the same QUD *Q<sub>10</sub>*: the first two are in a SIMILAR relation, the last two are in a CONCESSION relation.

- A<sub>48</sub>: surtout que pour des immeubles de l'OPAC on a la chance de pas être sur l'Periph  
 'above all, given that we are in social housing, we are lucky that we are not on the beltway'
- Q<sub>49</sub>: *Where are buildings of social housing?*
- >Q<sub>49.1</sub>: *Where are many of them?*
- >>A<sub>49.1</sub>: parce que [[beaucoup d'immeubles (...) des HLM]<sub>CT</sub> sont quand même [en périphérie]<sub>F</sub>]~  
 'because many buildings of social housing are actually in the suburbs'
- >Q<sub>49.2</sub>: *(Despite that) where is yours?*
- > A<sub>49.2</sub>: et [[nous]<sub>CT</sub> on est euh quand même [dans l'centre du treizième]<sub>F</sub>]~  
 'and we are actually in the center of the 13th district'

The fact that the building where the speaker and his family live is inside Paris (*dans le centre du treizième* 'in the center of the 13th district (of Paris)') is unexpected, knowing that most buildings of social housing are in the periphery. The second alternative denies the expectation triggered by the first one: 'We are in the center of the 13<sup>th</sup> district' means that they are not in the suburbs. The adverb *quand meme* 'actually' marks the CONCESSION relation.

The contrasting segments seen so far are all coordinated from both a syntactic and a discourse point of view. From a discourse point of view, they are parallel segments in the QUD tree (they follow PARALLELISM). CONCESSION, however, tends to deviate from this generalization and often involves discourse subordination (in the sense of Asher and Vieu 2005). Consider the Italian example in (76) from Brunetti (2024).<sup>36</sup>

- (76) Italian, blog interview (Brunetti 2024, 28)
- La mamma (...) parla la propria lingua ai figli  
 'The mother (...) speaks her own language to her children'
- >>Q<sub>25</sub>: *What do children often do (concerning their mother's language)?*
- >>>A<sub>25</sub>... : (...) molto spesso [i bambini  
 '(...) very often the children'
- >>>Q<sub>26</sub>: *Despite what?*
- >>>>A<sub>26</sub>: [pur [capendola perfettamente]<sub>F</sub>]~  
 'though understanding it perfectly'
- >>>...A<sub>25</sub>: non parlano attivamente]<sub>F</sub> la sua lingua]~  
 'very often the children, even if they understand it perfectly, do not proactively speak her language'

The subordinate clause in A<sub>26</sub> is introduced by the conjunction *pur* 'even if', explicitly marking CONCESSION: the expectation driven from A<sub>26</sub> is that the children, who perfectly understand the language, also speak it, while this is denied by the following assertion. The verb in A<sub>26</sub> has a non-finite form (gerundive) and the clause is nested inside the matrix clause. In Brunetti, De Kuthy, and Riester's (2021) terms (see Section 2.3.1), this clause is an IS-peripheral discourse segment, namely a segment that does not contribute to answer the matrix's QUD, but answers its own QUD, subordinated to the matrix's QUD (see Q<sub>26</sub>).

Discourse subordination can also occur when two segments are still *syntactically* coordinated. In (77), the interviewer is closing her blog interview by asking the interviewee, who is the author of an e-book on bilingualism, why one should buy her book.

<sup>36</sup> Concerning the choice of representing A<sub>25</sub> in two parts (A<sub>25</sub>... and ...A<sub>25</sub>) in the QUD tree, see discussion in Section 2.3.1 on ex. (30).

- (77) Italian, blog interview (*Bilingue per gioco*) (Brunetti 2024, 28)
- Q<sub>57</sub> :Spk1 : Concludendo, perché comprare *In che lingua giochiamo*?  
 ‘Concluding : why should one buy *In che lingua giochiamo* ?  
 (...)’
- >A<sub>57</sub> : Spk2 : [Perché [abbiamo a disposizione tantissime risorse ,]F]~  
 ‘Because we have plenty of resources at our disposal’
- >Q<sub>58</sub> : *Despite that, what happens?*
- >>A<sub>58</sub> : ma [[nemmeno ce ne rendiamo conto]F]~  
 ‘but we don’t even notice them’

The CONCESSION relation holds between A<sub>57</sub> and A<sub>58</sub>: since there are plenty of resources (to raise a bilingual child), one would expect that they are visible to everybody; on the contrary, we do not see them, and therefore don’t use them. The QUD is explicit and is made by the interviewer. In order for this question to have an appropriate answer, A<sub>58</sub> must be included in the answer: one should buy the book because it is not easy to realize how many resources there are at our disposal. Note that an answer made of A<sub>57</sub> alone would not make sense in this context: “We must buy the book because we have many resources”. Therefore, A<sub>58</sub> cannot be an independent answer to the question, but must be part of A<sub>57</sub>, the answer to Q<sub>57</sub>; at the same time, it answers a subordinated QUD, Q<sub>58</sub>.

CONCESSION can be seen as a special type of the OPPOSE relation: there is still opposite polarity, but the second utterance is denied by the causal inference triggered by the first. With OPPOSE, the first utterance does not trigger any inference: contrast simply arises by the fact that a predication of opposite polarity is made by the two utterances. In (68), for instance, the fact that there is a long line for cheap social housing does not trigger the inference that the line should be long for housing of intermediate cost. On the contrary, in (74), the fact that one has full control over two linguistic codes may be considered by most people as a definitory property of an early bilingual. Finally, unlike CORR, with CONCESSION all alternative propositions are true in the evaluation world, though one triggers the inference that the other *should* be false.

Summarizing, contrast requires the presence of alternatives, which are triggered by focus or topic. It also demands that such alternatives be instantiated by actual utterances in the discourse, so that a contrastive relation can be established between them. The kind of contrastive relation depends on particular contextual and semantic restrictions, which can be identified thanks to the QUD-structure associated with the contrasting pairs. SIMILAR and OPPOSE differ from each other in that with SIMILAR, the utterances provide different answers to the QUD (directly or indirectly) but they are not restricted in any particular way, while OPPOSE requires that the contrasting utterances make predications of opposite polarity, and that they answer two different sub-QUDs (that is, OPPOSE is only compatible with *complex* PARALLELISM). With CORR, both utterances answer the same QUD but one is rejected as false, while the other replaces it as the correct one. Finally, CONCESSION is a special type of OPPOSE in that one alternative denies the expectation triggered by the other.

## 6 Future developments on IS modeling and the QUD-based model

In the Introduction to Part 1, I’ve mentioned the discussion about the independence of information structure from other domains of grammar, in particular syntax. Some scholars such as Vallduvì, Leonetti, Neeleman have particularly insisted on the idea that information structure is governed by its own rules. In much of their work, as well as my own (see e.g. Brunetti 2004;

2009a), IS phenomena are explained in such a way that the existence of dedicated syntactic positions (*à la* Rizzi 1997, within the so-called ‘cartographic’ approach) is not necessary nor desirable, neither for focus and topic nor for specific types of focus and topic elements. Several arguments can be put forward against a cartographic approach to IS, such as the fact that focus and topic elements can be non-constituents; that their informational status can be ambiguous (which is hard to explain if they occupy dedicated positions, reached via feature-checking operations); that they are always additionally marked by intonation, etc. The idea put forward by these studies is that the syntax of the language is related to the expression of information structure only in as much as it can be exploited to fulfil IS-specific rules, such as for instance the requirement to have a transparent separation between focus and background or topic and comment.

The denial of a clear one-to-one mapping between form (and syntactic form in particular) and focus or topic interpretations goes a step farther in Matic and Wedgwood’s (2013) reflection on the notion of focus. In their thought-provoking paper, the authors give a series of examples from different languages in order to provide evidence not only against the idea of a one-to-one mapping between form and focus interpretation, but against the very existence of the informational category of focus. The authors claim that the large variety of semantic/pragmatic phenomena associated with focus eventually unveils the non-existence of such a category. As they say: “The category decomposes into a potentially unlimited number of encoded meanings, which for different reasons happen to be compatible with the effects usually associated with the putative discrete focus”. Once such a “decomposition” is made, nothing is left, in their view, of the focus category in itself. To the light of their analysis, focus is eventually only worth considering as a heuristic or descriptive tool. Though Matic and Wedgwood’s proposal is worth taking seriously, if only for the accuracy with which they revise certain focus phenomena that were imprecisely described in the literature, and for their illuminating, typologically-based, bird-eye perspective on the phenomenon, I will not adopt it here and will hang on the idea that a possible common semantic/pragmatic denominator for focus phenomena can be found. I will make the working hypothesis that such a common denominator resides in the fact that focus represents the part of the utterance that expresses the utterance’s speech act. This idea is the starting point of part of my ongoing and future research. Its implementation however can only be carried out through a long-term project.

The correspondence between focus and assertion has long been acknowledged for various typologically unrelated languages (see for instance Lambrecht 1994; Jary 2010; and Matic and Wedgwood 2013 themselves). Undeniably, however, non-assertive utterances (non-declarative sentence types) show IS marking too. In Romance languages, for instance, y/n interrogatives mark focus by prominence, in the same way as declarative sentences do:

(78) Italian

- a. Hai dato la BORSA, a Leo (o lo zaino)?  
‘Did you give Leo the BAG (or the backpack)?’
- b. Hai dato a LEO la borsa (o a Linda)?  
‘Did you give the bag to LEO (or to Linda)?’
- c. Gliel’hai DATA la borsa (o no)?  
‘Did you (or did you not) give him the bag?’

In order to account for the IS of non-declarative sentences, the position I would like to adopt is the one proposed by Beyssade and collaborators (Beyssade, Marandin, and Rialland 2003; Beyssade et al. 2004; Beyssade 2013; 2017): following Jacobs (1984), these authors propose that focus is “the part of the content that is specifically affected by the illocutionary operator associated with the sentence (...)” (Beyssade et al. 2004, 458), whether the utterance

accomplishes an assertive speech act or a different one, such as a question, a command, an advice, a promise, etc.

In order to explain this proposal, let us make a short step back. Starting from the name itself, the notion of “information” has been central in the information structure domain of grammar. Focus for instance can be defined as the answer to a question, hence as the *new piece of information* provided by an utterance at a certain point of discourse. Beyssade and collaborators suggest that the illocutionary definition of focus has gone unnoticed because the definition of focus as new information and the illocutionary definition “overlap in the general case since in standard assertions, Focus contributes a proposition that is informative and specifically asserted. However, the illocutionary definition is more general than the informative one since it covers all illocutionary types of utterances.” (Ibid.). In (79) the same sentence, with the same IS partition into focus (*pour Chirac*) and background (*Mathilde a voté pour x*) can be used to accomplish different functions, depending on the illocutionary operator:

- (79) a. C’est pour Chirac que Mathilde a voté. (Beyssade 2013, 221)  
 b. ASSERT  $\langle \lambda x$  Mathilde a voté pour x, Chirac  $\rangle$   
 c. QUEST  $\langle \lambda x$  Mathilde a voté pour x, Chirac  $\rangle$   
 d. SURPRISE  $\langle \lambda x$  Mathilde a voté pour x, Chirac  $\rangle$

In all cases, *Chirac* is what is asserted (Mathilde voted for Chirac and not for other candidates), what is questioned (the question is whether Mathilde voted for Chirac or for other candidates), is the source of surprise for the speaker (it is surprising that she voted for Chirac and not for other candidates), etc. Within this illocutionary perspective, the notion of information only makes sense for assertions, since with assertions indeed the speaker provides some information about the world, the truth of which she commits to and has evidence for. The notion of information is on the contrary not relevant for non-assertive utterances.

If we adopt this perspective, it becomes obvious that the QUD model presented above must be revised, since it is precisely based on the idea that focus is the *informative* part of the utterance, the one that answers an implicit question (where a question is an information-seeking device). The question-answer context does no longer make sense if the utterance realizes an act of questioning, requesting, advising, menacing, thanking, and so on. I will sketch below an idea of how to enrich the QUD-model so that it becomes capable to give the information structure of non-assertive utterances. The idea is beginning to be explored in a paper in progress in collaboration with Arndt Riester, to be published as part of a volume entitled *Annotating Text with Questions Under Discussion*.

The first non-assertive speech act that must be considered is the one associated with interrogatives, namely the act of questioning, since interrogatives play a special role in the QUD model. In Riester, Brunetti, and De Kuthy (2018) it is argued that interrogatives can function as overtly realized QUDs.<sup>37</sup> However, unlike implicit QUD – which must follow Q-GIVENNESS – overtly realized questions can be made of *new* content. The idea that I put forward here is that by consequence, like any other utterance of the text, interrogatives too must be preceded by an implicit QUD, which reveals what part of the interrogative is new and what part is given. See the constructed mini-contexts below, illustrating this point, where the two interrogatives have different ISs:

- (80) Italian  
 a. Francesca è andata a Firenze.

---

<sup>37</sup> Interrogatives then violate the rule stated in Riester et al (2018) that the utterances of a text are represented in the QUD-tree as terminal nodes.

- ‘Francesca went to Florence’  
*Q<sub>1</sub> : What about her going to Florence?*  
 >Q<sub>2</sub>: [[Come]<sub>F</sub> ci        è andata?]<sub>~</sub>  
           how    CL-LOC is gone  
 ‘How did she go there?’
- b. Francesca è andata a Firenze.  
 ‘Francesca went to Florence’  
*Q<sub>1</sub> : What are the consequences of this event?*  
 >Q<sub>2</sub>: [[Come ha reagito Paolo?]<sub>F</sub>]<sub>~</sub>  
           how has reacted Paolo  
 ‘How did Paolo take it?’

The reconstructed *Q<sub>1</sub>* in (80a) indicates that the only new content of the explicit interrogative *Q<sub>2</sub>* is the *wh*-word, while in (80b) it is the whole sentence. In this way, the focus part of an interrogative still corresponds to new content and is still obtained according to the principles stated in Riester et al (2018); however, what the focus indicates is the part of the sentence that is responsible for the questioning, rather than the informative, function.<sup>38</sup>

As for utterances expressing non-questioning acts, the idea is to assume that the QUD that precedes them asks about what part of the sentence is affected by the illocutionary operator. In case of assertions, this move is not very informative and eventually coincides with asking a question that is formulated according to the QUD principles:

- (81) Italian  
 Sono le nove.  
 ‘It’s nine’  
*Q: What (does the speaker **state** that) happens next?*  
 >A: [[Francesca va a lavarsi i denti]<sub>F</sub>]<sub>~</sub>  
 ‘Francesca goes to brush her teeth’

In case of non-assertive acts, the QUD will name the act that is at stake, for instance:

- (82) Italian  
*Q: What does the speaker **order**?*  
 >A: [[Vai a lavarti i denti!]<sub>F</sub>]<sub>~</sub>  
 ‘Go brush your teeth !’

If the imperative sentence is made of given content, as in (83), then the QUD will contain it:

- (83) FRANCESCA: Non mi piacciono i miei denti.  
 ‘I don’t like my teeth’  
*Q: What does the speaker order, concerning Francesca’s teeth?*  
 >A: [Vai a lavarte-li]<sub>F</sub>,    i    tuoi denti!  
           go to brush-CL.PL    the    your teeth  
 ‘Go to BRUSH your teeth!’

<sup>38</sup> Beyssade (2006) already suggests a speech-act analysis of the IS of (French) interrogatives. Concerning the fact that focus may not only include the *wh*-word but also other parts of the sentence, she suggests that this would correspond to the interrogative in fact asking two questions. For instance, all-focus *Come ha reagito Paolo?* ‘How did Paul react?’ Would correspond to ‘If Paul reacted, how did he do it?’.

Other sentence types (ex. exclamatives) and other speech acts should be treated the same way, though the details of the analysis need to be carefully spelled out. The difficulty may come from the possibility itself to identify the illocutionary properties of an utterance. Indeed, for some sentences the identification of the act they accomplish is not evident and there is room for ambiguity. Furthermore, complex speech acts may be realized by one single utterance.

While my present research is not yet at the point to tackle these issues, I have been dedicating part of my recent research to the study of interrogatives that accomplish ‘non-canonical’ speech acts, namely that do not (or not only) realize an act of questioning (do not seek for information). The studies of non-canonical questions that I summarize in Part 2 give an idea of how complex the identification of the act realized by these utterances can be. Firstly, a sentence with a particular form can be ambiguous and appropriate for different uses; secondly, it is not always trivial to understand what exactly one use consists of. Within the line of research described above, which tries to account for non-assertive utterances, an IS analysis of these non-canonical questions might not be trivial. Some help to accomplish the task will perhaps come from prosody, since we expect it to be particularly apt to make the utterance function transparent and to identify the parts that are under the scope of the illocutionary operator.



## Part 2 Non-canonical questions

### 1 Introduction

I will present below a line of research started in recent years in collaboration with various scholars, mostly from my own university (Lucia Tovenà, Hiyon Yoo, Agnès Celle) but also from Germany (Jan Fliessbach), and which is still active, concerning the pragmatic and to a lesser extent prosodic properties of certain interrogatives in Romance that have non-canonical uses, that is, that do not have (or not only) an information-seeking function. These questions can be called ‘non-canonical’ in the sense of Farkas (2022). Indeed, a canonical question act for Farkas (2022) is one that assumes a) the speaker’s ignorance concerning the issue she raises; b) the speaker’s assumption of the addressee’s competence to resolve the issue she raises; c) the speaker’s assumption that the addressee is cooperative and will therefore provide the missing information in the immediate future; d) the speaker’s goal to resolve the issue in the immediate future. The interrogatives I am going to present have uses that lack some of these properties, as we will see below.

Interrogatives that do not (or not fully) have an information-seeking function include rhetorical questions, but also mirative (or surprise) questions (Celle and Pélissier 2022), biased questions, challenging questions (Dekhissi 2016; Dekhissi and Coveney 2021), exam questions, and others. In Brunetti et al. (2021), Brunetti, Tovenà, and Yoo (2022) we investigate French interrogatives introduced by the *wh*-word *comment* ‘how’ that have a reading similar to ‘why’. These questions can be partly considered as information-seeking in that they ask for the reasons of some unexpected fact. However, as we will see below, the information they ask is of a particular kind; moreover, they also have an expressive function in that they convey the speaker’s surprise, disapproval, deception or other emotions triggered by the disconfirmation of the speaker’s assumptions.

In Fliessbach, Brunetti, and Yoo (under review), we look at French *comment ça* ‘how that<sub>demonstrative</sub>’ and Spanish *como que* ‘how that<sub>complementizer</sub>’ interrogatives, and propose three possible pragmatic uses for them, which only in part can be described as information-seeking. These interrogatives may ask for a particular kind of information, that is information about the meaning of some previous utterance (clarification requests); they can have basically the same use as *comment* questions mentioned above (ask for the reason of an unexpected fact); and they can have a challenging function (not information-seeking at all).

In an ongoing project with Jan Fliessbach and Gabriel Thiberge (see Thiberge, Brunetti, and Fliessbach submitted) we take a larger perspective on non-canonical questions and compare four French interrogatives that all ask for the reason, but that also convey other (non-canonical) readings. The idea is to better understand the degree of overlap among different interpretations and the degree of association between interrogative forms and the various interpretations. All these studies are described more in details below.

## 2 *Comment* interrogatives in French

The wh-word *comment* in French can be interpreted as a manner (a) or means (b) adverb that is dependent on the verb, or as a sentence adverb (c) that inquires about the reasons for the possibility of the ‘prejacent’ (in (84), the fact of refusing somebody’s proposition):<sup>39</sup>

(84) French (Brunetti et al. 2021, 248)

- Comment tu pourrais refuser sa proposition?  
how you could refuse.INF POSS.SG proposition  
‘How could you refuse his/her proposition?’
- a. Avec tact. (=⇒ manner of an event)  
‘With tact’
- b. En envoyant une lettre. (=⇒ means)  
‘By sending a letter’
- c. Je sais, mais j’ai déjà accepté un poste aux Etats Unis... (=⇒ reason)  
‘I know, but I’ve already accepted a position in the USA...’

It is not rare to have in a language a wh-expression with the meaning of ‘how’ that alternates with an expression whose meaning is ‘why’, see par example *how come* in English (Collins 1991; Ochi 2004). Two more examples of such interrogatives in French are provided in (85), from Desmets and Gautier (2009):

(85) French (Desmets and Gautier 2009, 107)

- a. Comment pouvez-vous vivre ainsi ?  
‘How can you live like this?’
- b. Comment oses-tu sortir comme ça?  
‘How dare you to go out like that?’

Note that *comment* with such a reason reading always triggers an inference that the fact the speaker asks about is unexpected, while this inference is not necessarily found in a question introduced by *pourquoi* ‘why’. But even when it is found, the meaning of the two interrogatives is still different: with *pourquoi*, the speaker just wants to know the cause of the unexpected fact; with reason-*comment*, the speaker asks the addressee in what way she should reconsider her expectations, so as to accept the prejacent. It is perhaps due to such a difference that a question with *pourquoi* can be answered by a sentence introduced by *parce que* ‘because’ (see 86a), while the answer of a *comment* question cannot be introduced by *comment*, or any other conjunction (86b).

- (86) Q: Pourquoi tu quittes un homme aussi adorable?  
‘Why are you going to break up with such a sweet man?’  
A: Parce qu’il m’a trompée.  
‘Because he cheated on me’
- b. Q: Comment peux-tu quitter un homme aussi adorable?  
‘How can you break up with such a sweet man?’  
A: \*Comment il m’a trompée.  
‘\*How he cheated on me’

---

<sup>39</sup> On the choice of the term ‘prejacent’ (or ‘pseudo-prejacent’), see Fleury and Tovenà (2018).

Note that the interrogatives introduced by *comment* with a reason reading generally contain a modal verb, often the modal *pouvoir* ‘can’; *pourquoi* ‘why’ on the contrary is not acceptable with it, cf. (86b) with (87):

- (87) ?? Pourquoi peux-tu quitter un homme aussi adorable?  
 ‘Why can you break up with such a sweet man?’

In (87) *pouvoir* gets an abilitative reading, so the question has the unlikely meaning ‘Why are you able to break up with such a sweet man?’; on the contrary, with *comment*, the modal gets an epistemic reading: ‘How is it possible that you broke up with such a sweet man?’. Indeed, *comment* is itself interpreted epistemically, so the epistemic reading of the modal naturally arises in such questions.

From the additional meaning of unexpectedness derives the emotional load that is commonly associated with these questions. Disbelief for the violation of expectations typically triggers surprise, but can also border on negative emotions such as concern, disapproval, or contempt. Indeed, these questions share some properties with the so-called ‘challenging questions’ (Koshik 2003) or ‘rhetorical opposing questions’ (Gruber 2001; Dekhissi 2016; Dekhissi and Coveney 2021), which are claimed to express the speaker’s negative reaction towards the listener’s words or actions via an implicit emphatic assertion, as in the interrogative introduced by *qu’est-ce que* in (88):

- (88) French (Dekhissi and Coveney 2021, 289)  
 Djamel: Au fait / j’ai vu Malik / Il te cherchait pour une histoire de thune là  
 ‘By the way / I saw Malik / He was looking for you for some money matter’  
 Nordine : Qu’est-ce t(u) as été te mêler de ça, toi?! / C’est pas ton problème  
 ‘Why did you meddle in that?! / That’s not your problem’

If these interrogatives are called ‘rhetorical’ by Dekhissi and others, what such a term exactly refers to must be looked at more closely. Rhetorical questions are claimed to have an assertive illocutionary force, which corresponds to the question’s negative answer (Sadock 1971). In this sense, challenging questions like (88) are rhetorical. The inferred meaning of the *qu’est-ce que* interrogative in (88) is ‘You should not have meddled in that’. However, rhetorical questions are also associated in the literature with no degree of uncertainty as to the answer, which is obvious for both speaker and addressee (Sadock 1974; Rohde 2006; Caponigro and Sprouse 2007). For instance, (89) has the obvious answer for both speaker and addressee, given their shared common knowledge: ‘Nobody would steal a newspaper’, which corresponds to the assertion that the interrogative eventually conveys.

- (89) Who would steal a newspaper? (Rohde 2006, 140)

Challenging questions cannot be called rhetorical in this sense, because clearly, since they are not used to point out or recall some shared knowledge to the addressee, but rather to do the opposite, namely to point out that some statement that is assumed to be true by the addressee is not true for the speaker. Similar observations can be made for *comment* interrogatives with a reason reading. The meaning of a question like *Comment peux-tu refuser sa proposition?* ‘How can you refuse his/her proposition?’ may be understood as its negative answer (‘You cannot refuse his proposition’). The question can be seen as having a challenging function, but like in challenging questions discussed above, the speaker alone commits to the negative assertion, not the addressee. In other cases of *comment* interrogatives, the speaker’s goal is not to challenge the addressee at all, but rather to get help from the addressee to revise her expectations and

eventually to accept the possibility of the prejacent. For instance, in (90), the speaker's goal is to get help from her friend to come to terms with the possibility of her friend joining the oil company, despite what the speaker knows about her friend's caring about the environment and hating oil companies.

(90) French (Brunetti et al. 2021, 253)

*A friend of yours tells you that he is going to join a big oil company with a bad reputation. You are shocked, since your friend has always been a fervent environmental activist. You say:*

Mais comment pourrais-tu intégrer cette entreprise?! (Tu détestes tout ce qu'ils représentent !)

'But how could you join this company? You hate all that it represents!'

The speaker's goal is to ask for some explanations, some justifications, which will make the addressee's behavior no longer in overt contradiction with the speaker's expectations (he needs money; he wants to discover the secrets of that company to better fight it; the company is not so bad as we thought, etc.). Even in this case, where the speaker has no challenging goal, the negative answer 'It is impossible for you to join this company' cannot be an obvious answer *for the addressee*.

Some *comment* interrogatives may get close to the rhetorical meaning of (89) (that is, the question infers an obvious answer for both speaker and hearer). That happens when they express surprise for the violation of a *generally shared rule*: these interrogatives eventually state the impossibility to break up such a rule. For instance (91), which contains the expression *ce type de* 'this kind of', which refers to a standard, can be used to build a form of consensus (cf. the impersonal subject *on* 'one') with the addressee on condemning the acceptance of a certain compromise (hence the interpretation: "Nobody can accept such a kind of compromise").

(91) Comment peut-on accepter ce type de compromis?

'How can one accept this kind of compromise?'

The rhetorical reading is enhanced by morphosyntactic or lexical features that have been identified in the literature as characteristics of rhetorical questions: the presence of modal verbs, the conditional mode, or the presence of high degree adverbs such as *tel* 'such', *aussi*, *tellement* 'so', which mark the threshold below or above which the existence of the event described by the prejacent is considered as not probable.

The fact that these interrogatives can have a rhetorical use or not, depending on their form, is indicative of the difficulty to establish a unique and unambiguous pragmatic function for them. Indeed, their function seems to heavily depend on a set of morphosyntactic and lexical properties as those just discussed as well as others, like for instance the subject's person (and in particular whether the subject coincides with the addressee or not). The function of these interrogatives will also vary depending on their prosodic characteristics, which are discussed in the following section.

## 2.1 The prosody of *comment* interrogatives

The study of the link between intonation and meaning is nowadays a crucial part of semantic research. This is due in my opinion to two facts. On the one hand, the attitude towards prosody –called the "Cinderella of communication" in Bolinger's (1986) seminal book on English intonation – has changed in the last three decades of linguistic research, as Albert Di Cristo already noticed ten years ago (Di Cristo 2013). On the other hand, what semanticists have been

interested in for many decades by now does not restraint to conventionalized meaning, but extends to all those phenomena that cannot be explained without taking into account language use and context (see the insightful review of the literature made by McNally 2013 on the semantics/pragmatics distinction, or lack of it). Given that suprasegmental aspects of speech and in particular intonation are tightly related to contextually-related aspects of meaning, the interest of semanticists for prosody has grown accordingly. Another set of phenomena where prosody is central is ambiguity, since ambiguous sentences are typically disambiguated by prosody.

For all these reasons, our interest in French *comment*-interrogatives extended to their prosody: we expected that prosody would disambiguate the canonical (manner) and non-canonical (reason) reading, and we hoped that prosodic properties would support the interpretive analysis of *comment* interrogatives with a reason reading presented above.

The prosody of non-canonical questions (rhetorical, surprise, echo ones) and how it differs from canonical information-seeking questions has been the object of some studies on French and other languages (see for instance Beyssade and Delais-Roussarie 2022; Celle and Pélissier 2022; Braun et al. 2019; Shochi, Rilliard, and Aubergé 2009). Nevertheless, a specific study on *comment* interrogatives in French was missing. We aimed to fill this gap by conducting a self-paced production study whose experimental procedure was inspired by Braun et al. (2019). Participants read silently a given context favoring either a manner or a reason interpretation, presented on a screen, and then pronounced aloud the target interrogative sentence, using the prosodic pattern that mostly suited the context. In addition to sentences that were identical, that is only disambiguated by context, we also asked participants to read sentences whose morphosyntax and lexicon favored either a manner or a reason reading. The annotation included initial accents, because they can have an emphatic or expressive function, and final tunes, because in French they are relevant for indicating the illocutionary force. We followed the ToBI annotation procedure for French (É. Delais-Roussarie et al. 2015) and also looked at some phonetic parameters: for each target utterance, a specific use of voice quality such as breathy or creaky voice, as well as the presence of laughter, was indicated.

Results of the statistical analysis show a difference in the distribution of initial tunes, presented in Figure 5. In both ambiguous and non ambiguous sentences, we observe a significantly greater proportion of Hi tones produced at the first syllable of *comment* in reason than in manner questions, which can be seen as a cue for the expression of surprise.

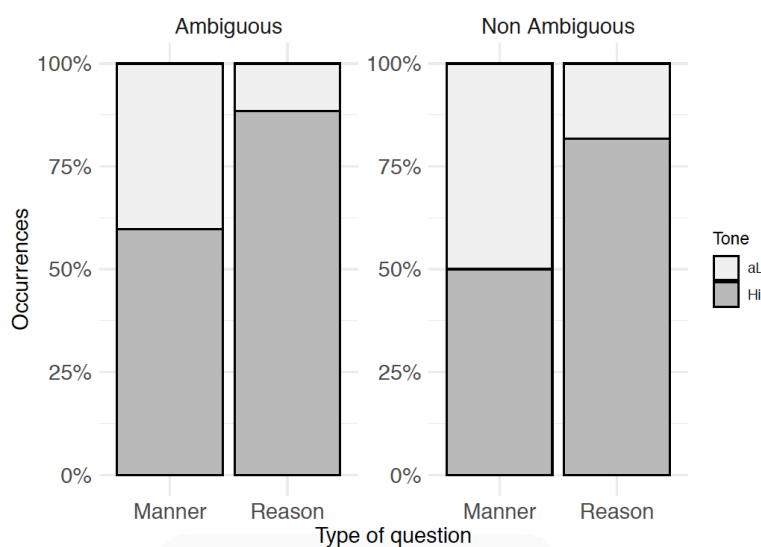


Figure 5: Initial tone produced on *comment*, from (Brunetti et al. 2021).

The distribution of the final tunes is presented in Figure 6. A !H\*!H% prosodic pattern was significantly more frequently produced with the reason than the manner reading in ambiguous sentences (see graph on the left). Interestingly, Delais-Roussarie et al. (2015) associate this contour to epistemic-biased utterances expressing speakers' attitudes such as doubt. They argue that this contour is used in contradiction statements when speaker and hearer do not share the same opinion. An example of the difference between the !H\*!H% pattern and the H\*H% pattern is given in Figure 7.

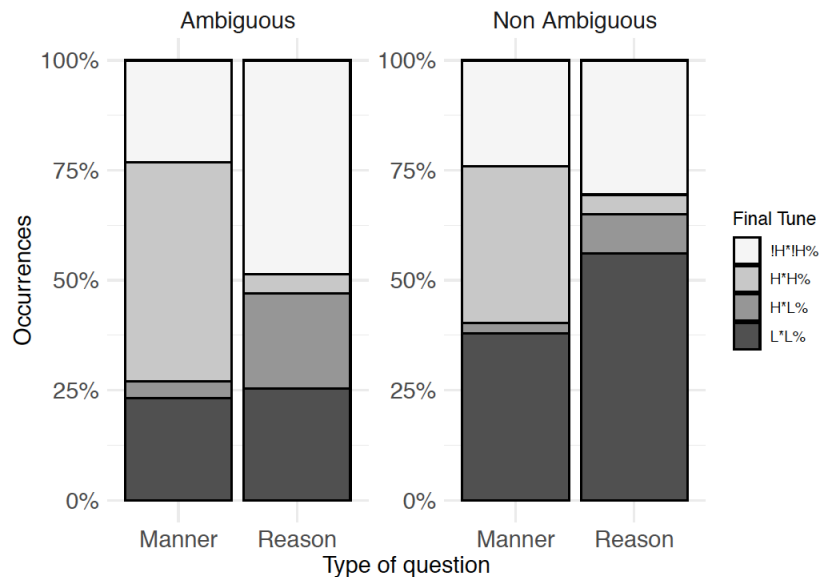


Figure 6: Occurrences of final patterns in ambiguous and non-ambiguous questions.

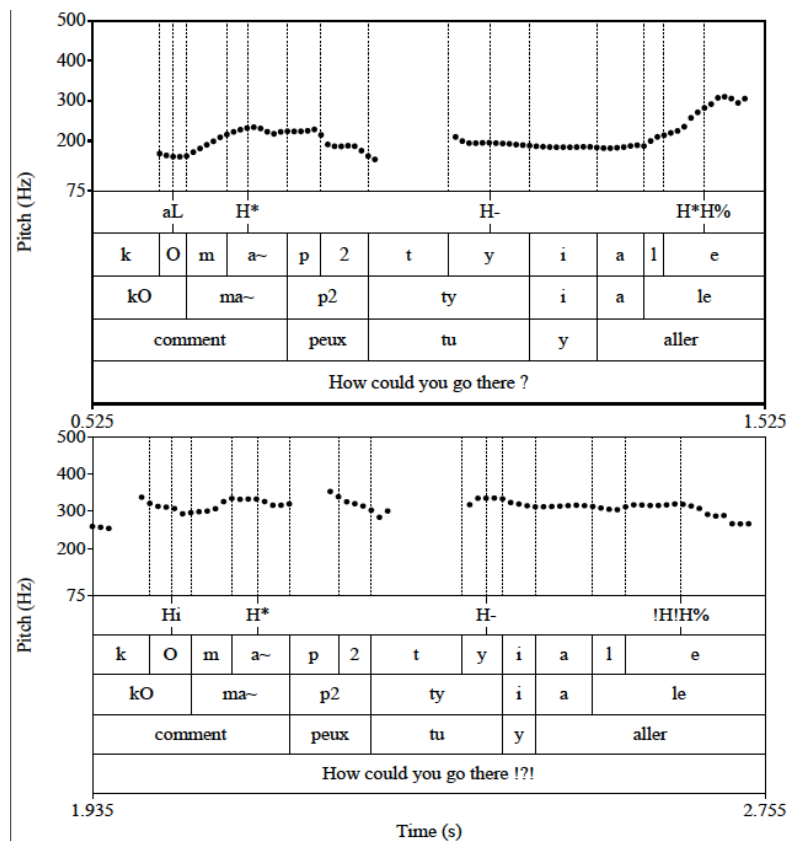


Figure 7: f0 curve and annotation of an ambiguous manner-comment question with a H\*H% final tune (upper picture) and reason-comment question with a !H\*!H% final tune (lower picture).

The final tune distribution is quite different for non-ambiguous sentences (see graph on the right of Figure 6), where the L\*L% patterns are more frequently produced in reason questions. One explanation for this result lies on the greater length of non-ambiguous sentence than that of ambiguous ones: speakers could have readjusted the prosodic contour because of rhythmic constraints (Martin 2009). Another possible explanation takes into account the linguistic material that was used in non-ambiguous interrogatives in order to make the reason interpretation transparent, typically high degree expressions and modals. A qualitative appreciation of the prosody of these expressions shows that they often bear a focal accent. It could then be the case that the accentuation of these parts has the effect of producing a low final contour. The accentuation of these crucial parts for the identification of the reason reading is interesting in itself: the speaker's strategy seems to be that of marking, through focus, the pragmatic function of the interrogative: for instance, a focal accent on the modal *pouvoir* highlights the fact that the question is about the *possibility* for the prejacet to be true.<sup>40</sup>

Results also showed that there is a difference in speech rate: reason questions were significantly slower than manner ones, a property that is shared with other French non-canonical questions (surprise or rhetorical, see Celle and Pélissier 2022; Beyssade and Delais-Roussarie 2022). A longer speech rate could then be ascribed to what all these questions have in common, namely the non-canonicity of their reading.

Finally, we also found that speakers use special phonation modes, such as breathy voice or laughter, more frequently in reason questions than in questions with a manner reading. This result goes along with the results found by Shochi et al. (2009) for Japanese, which associate this voice quality to surprise, and Braun et al. (2019) for German, where it is an important phonetic cue in distinguishing rhetorical questions. Importantly, Braun et al. (2019) consider breathy voice to signal an exasperated attitude of the speaker when uttering rhetorical questions.

Eight target utterances were finally produced with laughter. They all had a reason reading, which is an interesting point if we consider that laughter's main function is to highlight some incongruity (Mazzocconi, Tian, and Ginzburg 2022). The presence of laughter could then signal the incongruity between the speaker's expectations and the content of the prejacet.

### 3 French *comment ça* and Spanish *cómo que* interrogatives

Interrogatives introduced by *pourquoi* and *comment* are not the only interrogatives with a reason reading in French. In Fließbach, Brunetti, and Yoo (under review) we study another interrogative of the same semantic family, the one introduced by *comment ça* 'how that<sub>demonstrative</sub>', and we compare it to the Spanish interrogative introduced by *cómo que* 'how that<sub>complementizer</sub>' (see (92) and (93) respectively).<sup>41</sup>

#### (92) French, Film 'Le Petit Spirou'

PONCHELOT : Madame Suzette, Monsieur Spirou me charge de vous faire passer ce message. Comme il nous quitte dans deux semaines, il tient à vous...

'Madame Suzette, Monsieur Spirou asks me to pass you this message. Since he leaves us in two weeks, he cares about you...'

SUZETTE (elle le coupe) : Comment ça, il nous quitte dans deux semaines ?  
(she cuts him short): 'HOW THAT, he leaves us in two weeks?'

<sup>40</sup> Cf. discussion on the illocutionary role of focus in Section 6 of Part 1.

<sup>41</sup> In the paper, both wh-expressions are translated literally in the examples with capital 'HOW THAT'.



- (93) Spanish, Film ‘Estoy vivo’  
 DIRECTORA : Lo cierto es que usted debería estar vivo.  
 ‘The truth is that you should be alive.’  
 VARGAS : Espere, espere... un momento por favor  
 ‘Wait, wait... one moment please,’  
 ¿Cómo que debería estar vivo?  
 ‘HOW THAT I should be alive?’  
 ¡Si es así exijo que inmediatamente me devuelvan!  
 ‘If that is the case, I demand that you bring me back immediately!’

Such interrogatives are syntactically marked: both are characterized by a particular element following the *wh*-word (the demonstrative *ça* and the complementizer *que*), whose behavior is not the one expected from a demonstrative or a complementizer, respectively. Therefore, the form of these interrogatives makes them good candidates to express non-canonical question acts. Indeed, the literature, though not rich, points to this direction.

We have seen above that French *comment* interrogatives can ask about the reasons for an unexpected fact. This interpretation is also available with interrogatives introduced by *comment ça*, with the difference that what is unexpected must be mentioned in an immediately preceding utterance (that we call the ‘provocation’). The provocation can fully or in part be quoted after the *wh*-expression (we call this quote the ‘remnant’). In Fliessbach, Brunetti and Yoo (under review) we call this reading a ‘mirative’ one, and (92) and (93) are two examples of it. Smirnova (2021) finds that *comment ça* interrogatives ask for clarification, as in (94). The interrogative in this example is indeed followed by a reformulation of B’s question in (94d).

- (94) French, (Smirnova 2021, 52)  
 a. A : Surtout, je te le répète, cela n’intéresse personne  
 ‘Above all, I repeat it, that does not interest anybody’  
 b. B : Et moi ?  
 ‘And me?’  
 c. A : Comment ça ?  
 ‘HOW THAT ?’  
 d. B : Je t’intéresse, moi ?  
 ‘Do I interest you ?’

Lefeuve (2009) mentions yet another reading, where *comment ça* questions the validity of its provocation, as in (95), where the truth of the sentence ‘We didn’t do anything (wrong)’ is questioned:

- (95) French, (Lefeuve 2009, 82)  
 a. A: Mamie !... Madame !... On n'a rien fait !  
 ‘Grandma!... Madam!... We didn’t do anything (wrong)!’  
 b. B: Comment ça, vous n'avez rien fait ? (...) et mon pouce alors ?  
 ‘HOW THAT, you did not do anything? (...) and what about my thumb?’  
 c. Il agite sous votre nez le doigt en question empaqueté dans un gros pansement  
 ‘He moves under your nose the finger in question, packed inside a big plaster’

As for Spanish, Rosemeyer (2021) argues that *cómo que* interrogatives may have an epistemic reading and can be paraphrased with “How is it possible that ...?” (cf. Smirnova 2021). Rosemeyer further argues that such “interactional challenges are (...) not information-seeking



in the sense that they make an answer to the question relevant. Rather, they signal that the listener needs to justify her or his previous move.” (Rosemeyer 2021, 132). A “clarification request” use has only been discussed for French, but it also seems plausible for Spanish, and is supported by a search that we have done on the OPUS subtitles French-Spanish parallel corpus (Lison and Tiedemann 2016).

Given such a background, in Fliessbach, Brunetti and Yoo (under review) we looked at a corpus made of film scripts and sociolinguistic interviews for both French and Spanish.<sup>42</sup> We annotated *comment ça* and *como que* interrogatives according to three readings: mirative reading, clarification request, and disagreement. In the mirative reading, the speaker asks for the reasons of some unexpected event reported in the provocation, reasons that will help her to revise her expectations, cf. (92) and (93). In the case of clarification requests, the speaker asks for clarifications about the meaning of the provocation or a part of it, as in the Spanish example in (96), where the interrogative is followed by an explicit reference to what the speaker meant when he said *algo* ‘something’. With disagreements, the interrogative is used to challenge or contradict the addressee’s provocation, as in (97). Disagreements lead to a conversational crisis that can only be resolved by retraction of a discourse commitment by one of the interlocutors.

- (96) Spanish, Film ‘Estoy vivo’  
 a. ENLACE: (...) El auténtico Marvin murió hace dos días. Algo invadió su cuerpo entonces.  
     ‘The real Marvin died two days ago. Something has invaded his body since then.’  
 b. MÁRQUEZ : Algo... ¿Cómo qué algo?  
     ‘Something... HOW THAT something?’  
 c. ENLACE: Un ente hostil, ya se lo he dicho. (...) Un espíritu. Un ser incorporeo. (...)  
     ‘A hostile body, I already told you that. (...) A spirit. An incorporeal being.’
- (97) French, Film ‘Un homme idéal’  
 a. ALAIN FURSAC : (...) J’ai appelé mon ami Cassagnol, le préfet. Il m’a promis de faire le maximum.  
     ‘I called my friend Cassagnol, the prefect. He promised me he’ll do his best.’  
 b. MATHIEU : C’est gentil, mais c’était pas la peine...  
     ‘That’s nice of you, but it was not worth it’ (meaning: ‘you didn’t have to’)  
 d. ALAIN FURSAC: Comment ça pas la peine ! Mais on tient à vous, Mathieu!  
     ‘HOW THAT not worth it! But we care about you, Mathieu!’

Both mirative questions and clarification requests have an information-seeking component in that they make an answer relevant (Rosemeyer 2021) and require the addressee to provide for it (Caponigro and Sprouse 2007). In the case of disagreement uses, on the contrary, there is

<sup>42</sup> For French, we included 97 film scripts that were accessible on the website of the association *Lecteurs anonymes* (<http://lecteursanonymes.org/scenario/>). To this, we added sentences extracted from two corpora of sociolinguistic interviews: the ESLO corpus ((Abouda and Skrovec 2018) and the CFPP2000 (Branca-Rosoff et al. 2012). For Spanish, we could search through 117 film scripts accessible on the *Aprendercine* and *Premio Goya* websites. To this, we added 70 examples from the COSER corpus of sociolinguistic interviews (Fernández-Ordóñez 2005).

no such expectation of a reaction from the addressee; the speaker does not accept the provocation and therefore does not ask for any explanation: her goal is to convince the addressee that what he states is not true. This reading is comparable to that of challenging questions discussed in Section 2 above (cf. (88)). However, *comment ça/como que* interrogatives require a provocation in the immediately preceding context, which is partially or fully quoted in the interrogative as the remnant. Moreover, the remnant does not need to be propositional, so the speaker can challenge non-propositional parts of the provocation.

The annotation of the three functions in our corpus of film scripts and sociolinguistic interviews has fed a statistical analysis that has confirmed both the validity of our classification and of the attribution of the three functions to the interrogatives of *both* languages. Moreover, we tested the relative importance of some linguistic criteria for the attribution of the three functions. In particular, we found that a remnant that is an exact quote of (part of) the provocation, and in particular that does not present person shift, is an indicator of a clarification request use, while the presence of adversative markers indicates uses different from clarification requests.

The most important predictor of the three functions however is whether there is a change of turn after the interrogative or not. Clarification requests are more likely followed by a change of turn, which is expected since this function is the only truly information-seeking one. Furthermore, disagreement and mirative questions differ from each other in that the turn continuation of the former is more often an utterance that does not ask for information. Indeed, given that miratives accept the proffered content and evaluate it as unexpected, it seems reasonable that the speaker may seek information that would solve the mismatch between such content and her expectations. Disagreements, on the contrary, take a diverging stance that may require an explanation about the falsity of the proffered content, and this can be done using an assertive continuation.

The surface phenomena that we investigated allowed us to reach a predictive accuracy of 63%, meaning that there are still important factors to be accounted for.

At the discourse level, a more fine-grained classification of the addressee's response to the turn containing the interrogative might improve the accuracy of our model. Furthermore, an annotation of discourse relations holding between the interrogative and the preceding utterance (cf. Kehler and Rohde 2017) and between the interrogative and the utterance immediately following it (cf. the taxonomy of response space in Ginzburg et al. 2022) might improve the predictability of the three discourse functions. We expect that, depending on the use of the interrogative, recurring discourse relations will be found between the interrogative and the surrounding utterances. This future research will benefit from the experience that I am gathering thanks to my contribution to the creation of an annotation scheme for interrogatives in English and French TED Talks, within the project Transquest led by Agnès Celle. The annotation scheme precisely includes the annotation of discourse relations between the interrogative and the preceding discourse and the annotation of the interrogative's response space.

At the construction level, prosody is a factor that was not included among the linguistic cues for the three readings, but it is undoubtedly worth considering. Previous prosodic descriptions of echo questions (cf. for instance É. Delais-Roussarie et al. 2015 for French, ; Hualde and Prieto 2015 for Spanish) are problematic in that they do not assume the same fine-grained functional distinctions of *comment ça* or *como que* interrogatives that we have seen above.

In an attempt to gain some first insights, we performed a pilot recording of 10 dialogues per language from the script corpus with the surrounding dialogue with one French and two Spanish native speakers. Our Spanish informants always produced L\*L% nuclear configurations, with variability in the scaling of a consistent prenuclear rise to an H target that aligns either with the right edge of *cómo* or with the vowel in *que*. The strategy of the French informant is different. Instead of a rise from low as on *cómo*, the realizations on *comment* all start on an initial high

target and fall to low on the second syllable. Moreover, our French informant shows different nuclear configurations. Disagreement uses of *comment ça* interrogatives are realized with an L% boundary tone, whereas mirative and clarification request uses show mid-high boundary tones. One further possible distinction between mirative and clarification request uses can be found in phrasing: in mirative uses, no phrasing between *comment ça* and the remnant was found, while in clarification requests, phrasing was marked by the presence of a pause, as if the remnant was considered a citation. A future study, including a representative sample of speakers, should test these first impressions.

## 4 Future developments on ‘reason’ interrogatives

An ongoing larger project, in collaboration with Jan Fliessbach and Gabriel Thiberge, is dedicated to get a complete picture of “reason” interrogatives in French and their interpretative nuances. The idea is to gather experimental evidence for the different interpretations (and the interplay between them) of the various interrogative constructions that can be said to share at least a reason reading.

Other than *pourquoi*, which can be considered as the default reason question in French, we have seen above that *comment* and *comment ça* are used with a reason reading. To these, the wh-expressions *comment se fait-il que* and its informal counterpart *comment ça se fait que* must be added (see (98)). Two more wh-expressions must finally be considered, namely *qu’est-ce que* ‘what is it that’ with no direct-object function (cf. the challenging question seen in (88) and the one in (99) below, where interestingly *qu’est-ce que* is translated with *why*); and *c’est quoi* ‘it is what’, discussed for instance in Celle and Pélissier (2022), see (100):

- (98) (Abeillé et Godard 2021, 1114)  
Comment se fait-il que les épisodes de certaines séries télévisées sont diffusés dans le désordre? (csa.fr, 7 nov. 2012)

- (99) (Dekhissi and Coveney 2021, 127)  
Ce n’est pas la peine, il est foutu, je le sais bien ! ... et c’est de ta faute, Pierre!  
Qu’est-ce que tu avais besoin de lui dire toutes ces horreurs ? (Sapilla, *Le Sac de Pierre*, 2011)

‘It’s not worth it, he is fucked up, I know it very well!... and it’s your fault, Pierre!  
Why did you have to tell him all those horrible things?’

- (100) (Dekhissi 2018, 223)  
C’est quoi, cette chance de malade, là ?!  
it is what this luck of sick there  
‘How can you be so lucky?!’

The first two experiments that we have conducted within this project aimed to compare a subset of such wh-expressions: *pourquoi*, *comment ça*, *comment se fait-il que*, and non-argumental *qu’est-ce que* (see Thiberge, Brunetti, and Fliessbach submitted). *Comment* alone was excluded because reason-*comment* interrogatives, unlike the other reason interrogatives, strongly favor the presence of a modal verb, to the point that some speakers we informally consulted do not find such sentences grammatical without one. We also excluded *comment ça se fait que*, which we considered as synonymous to the more formal *comment se fait-il que*, and we excluded *c’est quoi*, which can only be followed by a noun phrase. By making such choices,

we aimed at obtaining a set of interrogatives that were experimentally comparable and therefore whose interpretative differences and similarities could be identified more clearly.

A choice was also made with respect to the pragmatic functions that were to be tested in the experiments. The starting assumption was that all of these interrogatives have a reason reading, that is, that they can be used to ask for the reasons of some fact. However, the work presented in the sections above shows that the reason reading may be linked to the need of the speaker to reconcile her assumptions with an unexpected fact. We therefore made the hypothesis that these interrogatives could also express surprise, due to the disconfirmation of the speaker's assumptions. We have also seen above that these interrogatives can be similar to challenging questions, namely they can be used to express a discrepancy between speaker and addressee. We therefore tested this reading too, that we called 'reproach'. Finally, since *pourquoi* and *qu'est-ce que* have been described as being used to ask for the goal (other than the reason) behind some action or behavior (see Dekhissi and Coveney 2021), this reading was included too.

Our hypotheses were that *pourquoi* mainly has information-seeking uses, that is, it asks for the reason or for the goal, though it can also be used in contexts where it expresses the surprise for some unexpected fact, or in order to make an interactional challenge. *Comment se fait-il que*, *comment ça* and non-argumental *qu'est-ce que* can ask for reasons for the possibility of an unexpected fact, hence they can express surprise (Celle and Pélissier 2022; Fleury and Tovenà 2018; Desmets and Gautier 2009; Smirnova and Abeillé 2021). A challenging reading is expected to be possible with these interrogatives, especially with non-argumental *qu'est-ce que*.

The two experiments aimed to experimentally test these hypotheses. In the first, we constructed sequences of two sentences: an interrogative introduced by one of the four wh-expressions followed by another sentence forcing a particular interpretation of the interrogative, see (101).

- (101) Q: Pourquoi/Comment ça/Comment se fait-il que/Qu'est-ce que tu as planté des oliviers dans ton jardin?
- |       |  |            |
|-------|--|------------|
|       | 'Why did you plant olive-trees in your garden?'      |            |
| a.    | Ils étaient en promotion ?                           | (reason)   |
|       | 'Were they on sale?'                                 |            |
| a'.   | C'est pour produire de l'huile d'olive ?             | (goal)     |
|       | 'To produce olive oil?'                              |            |
| a''.  | On doit planter des fleurs pour aider les abeilles ! | (reproach) |
|       | 'One should plant flowers to help the bees!'         |            |
| a'''. | Tu as toujours parlé d'acheter des pommiers !        | (surprise) |
|       | 'You always spoke about buying apple trees!'         |            |

Participants were asked to judge the coherence of the continuation using a slider. We expected that the comparison of all the experimental conditions with each other would potentially reveal, on one hand, a hierarchy of interpretations for each of the interrogative forms studied and on the other, the preferred associations between interrogative form(s) and interpretation(s).

Results of a Bayesian statistical analysis show, from the perspective of the linguistic form, that some wh-expressions are specialized for expressing certain interpretations. For instance, *qu'est-ce que* and *comment ça* are better with continuations of reproach and surprise than reason or goal, showing that they tend to specialize in a non-canonical reading of the expressive type, whereas *pourquoi* seems to prefer an information-seeking reading, questioning about the reason or the goal. From the perspective of the interpretation, the expression of surprise seems more coherent with *comment ça* or *comment se fait-il que*, the expression of reproach with *comment*

*ça*; the function of asking for the goal seems to be better expressed by *comment se fait-il que* and *pourquoi*, and finally asking for the reason seems to be better expressed by *comment se fait-il que*. *Qu'est-ce que* was judged less coherent than the other three interrogative adverbials across all continuation types, suggesting an external factor, presumably sociolectal variation. Moreover, continuations favoring a reason reading were judged less coherent overall. We attribute this result to the experimental setting, in that a continuation after an interrogative with an 'unmarked' reason reading is not expected: not only this reading is information-seeking, so the speaker is supposed to wait for an answer from her interlocutor, but it is the default reading, so we do not expect it to require a continuation that disambiguates it.

Experiment 2 tried to limit these potential side effects of Experiment 1's setting. Experiment 2 was designed so that participants made a metalinguistic reflection on the relative "weight" of each of the four interpretations per interrogative type. They were asked to evaluate on a scale to which extent each of the four possible readings was present in each interrogative type (presented alone, with no continuation).

A Bayesian linear regression model was calculated independently for each of the response scales. The results complement those of Experiment 1. The interpretation of reason is present with the forms *comment ça* and *qu'est-ce que*, though it favors the use of *comment se fait-il que* and *pourquoi* (which are also favored when a goal reading needs to be expressed). *Comment ça* seems to mainly convey surprise, and to a lesser extent reproach. Conversely, *comment se fait-il que* and *pourquoi* seem to be among the forms less compatible with an expressive interpretation, i.e. reproach or surprise.

These results corroborate many of the intuitions found in the literature and mentioned above. They do, however, provide a much clearer and more precise idea of the similarities and differences in interpretation between the four interrogatives. Experiment 2 confirmed that *pourquoi* is the most appropriate element for questioning about the reason and, more generally, that it is less good/appropriate than the other forms in non-information-seeking uses (surprise and reproach). The results also show that *comment ça* and *qu'est-ce que* are less appropriate than *pourquoi* for asking for the goal, again confirming their less optimal use as genuine questions. The results for *comment ça* show particularly well its preference for expressive uses.<sup>43</sup> However, we found no significant difference between *comment se fait-il que* and *pourquoi* in the association with reason, which is unexpected given the marked form of *comment se fait-il que*. The use of this form in a formal register could explain its association with *pourquoi*.

Several follow up studies need to be conducted in order to enrich and complete this picture. One study is in preparation that will compare *comment se fait-il que* with the more colloquial *comment ça se fait que*. The goal of this study is to understand whether these two wh-expressions only differ from a sociolinguistic point of view or also occupy different functional positions in the usage space of "reason" interrogatives.

Another follow up study concerns *qu'est-ce que*, which was very poorly judged. It will be useful in the future to test this form only after selecting participants who fully accept it in their idiolect. Alternatively, a way to reduce the negative bias towards non-argumental *qu'est-ce que* is to present the sentences to participants in a spoken rather than written form. A version of Experiment 1 with audio recordings of the items (and without their written counterpart) is being prepared at present. Other general effects are expected from the presence of intonation. We predict that intonation will help out participants in the recognition of the required interpretations (the ones indicated by the continuations) and we therefore expect results to be more reliable (for instance, we expect less variance among speakers). Intonation may also favor a better

---

<sup>43</sup> It must be kept in mind that the clarification request reading (see Section 3 above) was not tested in the two experiments.

judgment of all non-canonical readings, given that prosody typically contributes to disambiguate expressive content (see Brunetti et al 2021).

The different experimental settings of the two experiments finally revealed us the importance, for the interpretation of interrogatives, of the interactional context in which they occur. Experiment 1 involved a single speech turn, whereas the questioning readings would have greater visibility if the turn changed and if there was a response from another person. Future studies will therefore explicitly test the role of the interactional context by integrating it in the experimental protocol.

## References

- Abouda, Lotfi, and Marie Skrovec. 2018. "Pour Une Micro-Diachronie de l'oral : Le Corpus ESLO-MD." Edited by F. Neveu, B. Harmegnies, L. Hriba, and S. Prévost. *SHS Web of Conferences*, no. 46: 11004. <https://doi.org/10.1051/shsconf/20184611004>.
- Arregi, Karlos. 2003. "Clitic Left Dislocation Is Contrastive Topicalization." In *Proceedings of the 26th Annual Penn Linguistics Colloquium*, 9.1:31–44. Penn Working Papers in Linguistics. Philadelphia: Penn Linguistics Club.
- Ashby, William J. 1988. "The Syntax, Pragmatics, and Sociolinguistics of Left- and Right-Dislocations in French." *Lingua* 75 (2–3): 203–29. [https://doi.org/10.1016/0024-3841\(88\)90032-0](https://doi.org/10.1016/0024-3841(88)90032-0).
- Asher, Nicholas, and Laure Vieu. 2005. "Subordinating and Coordinating Discourse Relations." *Lingua* 115 (4): 591–610. <https://doi.org/10.1016/j.lingua.2003.09.017>.
- Avanzi, Mathieu. 2012. *L'interface Prosodie/Syntaxe En Français: Dislocations, Incises et Asyndètes*. GRAMM-R. Études de Linguistique Française / GRAMM-R. Studies of French Linguistics 13. Bruxelles: P.I.E. Peter Lang.
- Belletti, Adriana. 2004. "Aspects of the Low IP Area." In *The Structure of CP and IP*, edited by Luigi Rizzi, 16–51. Oxford University Press New York, NY. <https://doi.org/10.1093/oso/9780195159486.003.0002>.
- Benincà, Paola, Gianpaolo Salvi, and Lorenza Frison. 1988. "L'ordine degli elementi della frase e le costruzioni marcate." In *Grande grammatica italiana di consultazione. 1: La frase, i sintagmi nominale e preposizionale*, edited by Lorenzo Renzi, Nuova ed, 115–225. Strumenti Linguistica e critica letteraria. Bologna: Il Mulino.
- Beyssade, Claire. 2006. "La Structure de l'information Dans Les Questions : Quelques Remarques Sur La Diversité Des Formes Interrogatives En Français." *Linx*, no. 55 (December): 173–93. <https://doi.org/10.4000/linx.470>.
- . 2013. "Sous Le Sens : Recherche Sur La Modélisation Des Aspects Dynamiques de La Communication et de l'interprétation." Mémoire d'Habilitation à diriger des Recherches, Université de Paris 8.
- . 2017. "Chapitre 6. Sens et intonation : accent initial et stratégies discursives." In *Sous le sens*, 159–82. Sciences du langage. Saint-Denis: Presses universitaires de Vincennes. <https://www.cairn.info/sous-le-sens--9782842925895-p-159.htm>.
- Beyssade, Claire, and Elisabeth Delais-Roussarie. 2022. "The Prosody of French Rhetorical Questions." *Linguistics Vanguard* 8 (s2): 277–86. <https://doi.org/10.1515/lingvan-2020-0121>.
- Beyssade, Claire, Elisabeth Delais-Roussarie, Jenny Doetjes, Jean-Marie Marandin, and Annie Rialland. 2004. "Prosodic, Syntactic and Pragmatic Aspects of Information Structure. An Introduction." *Handbook of French Semantics*, 455–75.

- Beyssade, Claire, Jean-Marie Marandin, and Annie Rialland. 2003. "Ground/Focus Revisited. A Perspective from French." In *Proceedings of LSRL 2001*. Benjamins.
- Bolinger, Dwight. 1986. *Intonation and Its Parts: Melody in Spoken English*. Stanford, Calif: Stanford University Press.
- Branca-Rosoff, Sonia, Serge Fleury, Florence Lefevre, and Matthew Pires. 2012. "Discours Sur La Ville. Présentation Du Corpus de Français Parlé Parisien Des Années 2000." <http://cfpp2000.univ-paris3.fr/CFPP2000.pdf>.
- Braun, Bettina, Nicole Dehé, Jana Neitsch, Daniela Wochner, and Katharina Zahner. 2019. "The Prosody of Rhetorical and Information-Seeking Questions in German." *Language and Speech* 62 (4): 779–807.
- Brody, Michael. 1990. "Remarks on the Order of Elements in the Hungarian Focus Field." *Approaches to Hungarian* 3 (95–121): 148.
- Brunetti, Lisa. 2004. *A Unification of Focus*. RGG Monografie. Padova: Unipress.
- . 2009a. "Discourse Functions of Fronted Foci in Italian and Spanish." In *Focus and Background in Romance Languages*, edited by Andreas Dufter and Daniel Jacob, 43–81. Studies in Language Companion Series 112. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/slcs.112.04bru>.
- . 2009b. "On the Semantic and Contextual Factors That Determine Topic Selection in Italian and Spanish." *The Linguistic Review* 26 (2–3). <https://doi.org/10.1515/tlir.2009.010>.
- . 2009c. "On Links and Tails in Italian." *Lingua* 119 (5): 756–81. <https://doi.org/10.1016/j.lingua.2008.10.005>.
- . 2024. "Contrast in a QUD-Based Information-Structure Model." In *On the Place of Contrast in Information Structure: Definition, Types, Encoding and Annotation*, edited by Jorina Brysbaert and Karen Lahousse, Ch. 8. Trends in Linguistics. Studies and Monographs [TiLSM]. Ch. 8, 191–224, De Gruyter. <https://doi.org/10.1515/9783110986594-008>
- Brunetti, Lisa, Stefan Bott, Joan Costa, and Enric Vallduví. 2011. "A Multilingual Annotated Corpus for the Study of Information Structure." In *Grammatik Und Korpora 2009. Dritte Internationale Konferenz, Mannheim, 22. - 24.09.2009. Grammar & Corpora 2009*, edited by M. Konopka, J. Kubczak, C. Mair, F. Sticha, and U.H. Wassner. Mannheim: Gunter Narr Verlag.
- Brunetti, Lisa, Kordula De Kuthy, and Arndt Riester. 2021. "The Information-Structural Status of Adjuncts: A Question-under-Discussion-Based Approach." *Discours*, no. 28 (October). <https://doi.org/10.4000/discours.11454>.
- Brunetti, Lisa, Laia Mayol, and Xavier Villalba. 2020. "Bridging Strength, Monotonicity, and Word Order Choices in Catalan." *Discourse Processes* 57 (8): 703–24. <https://doi.org/10.1080/0163853X.2020.1768623>.
- Brunetti, Lisa, Lucia M. Tovená, and Hiyon Yoo. 2022. "French Questions Alternating between a Reason and a Manner Interpretation." *Linguistics Vanguard* 8 (s2): 227–37. <https://doi.org/doi:10.1515/lingvan-2020-0130>.
- Brunetti, Lisa, Hiyon Yoo, Lucia Tovená, and Rachel Albar. 2021. "French Reason-Comment 'How' Questions: A View from Prosody." In *Expressive Meaning Across Linguistic Levels and Frameworks*, by Lisa Brunetti, Hiyon Yoo, Lucia Tovená, and Rachel Albar, 248–78. Oxford University Press. <https://doi.org/10.1093/oso/9780198871217.003.0012>.
- Büring, Daniel. 1997. *The Meaning of Topic and Focus: The 59th Street Bridge Accent*. Routledge Studies in Germanic Linguistics 3. London New York: Routledge.
- . 2003. "On D-Trees, Beans, and B-Accents." *Linguistics and Philosophy* 26 (5): 511–45. <http://www.jstor.org/stable/25001898>.



- Butt, Miriam. 2014. "Questions and Information Structure in Urdu/Hindi." In *Online Proceedings of the LFG14 Conference*, edited by Miriam Butt and Tracy Holloway King. Stanford: CSLI.
- Caponigro, Ivano, and Jon Sprouse. 2007. "Rhetorical Questions as Questions." *Proceedings of Sinn Und Bedeutung*, 121-133 Pages. <https://doi.org/10.18148/SUB/2007.V11I10.635>.
- Celle, Agnès, and Maud Péliissier. 2022. "Surprise Questions in Spoken French." *Linguistics Vanguard* 8 (s2): 287–302. <https://doi.org/10.1515/lingvan-2020-0109>.
- Chomsky, Noam. 1972. "Deep Structure, Surface Structure and Semantic Interpretation." In *Studies on Semantics in Generative Grammar*, 62–119. De Gruyter. <https://doi.org/10.1515/9783110867589-004>.
- Cimmino, Doriana. 2016. "La topicalizzazione in italiano in prospettiva contrastiva con l'inglese : il caso della scrittura giornalistica online." <https://doi.org/10.5451/UNIBAS-006769929>.
- Cinque, Guglielmo. 1993. "A Null Theory of Phrase and Compound Stress." *Linguistic Inquiry* 24 (2): 239–97. <http://www.jstor.org/stable/4178812>.
- Collins, Chris. 1991. "Why and How Come." *MIT Working Papers in Linguistics* 15: 31–45.
- Cresti, Emanuela, and Massimo Moneglia, eds. 2005. *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*. Studies in Corpus Linguistics 15. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/sci.15>.
- 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). <https://doi.org/10.5334/gjgl.1100>.
- De Cat, Cécile. 2007. *French Dislocation: Interpretation, Syntax, Acquisition*. Oxford Studies in Theoretical Linguistics 17. Oxford: Oxford university press.
- De Kuthy, Kordula, Lisa Brunetti, and Marta Berardi. 2019. "Annotating Information Structure in Italian: Characteristics and Cross-Linguistic Applicability of a QUD-Based Approach." In *Proceedings of the 13th Linguistic Annotation Workshop, LAW@ACL 2019, Florence, Italy, August 1, 2019*, edited by Annemarie Friedrich, Deniz Zeyrek, and Jet Hoek, 113–23. Stroudsburg: The Association for Computational Linguistics.
- De Kuthy, Kordula, Nils Reiter, and Arndt Riester. 2018. "QUD-Based Annotation of Discourse Structure and Information Structure: Tool and Evaluation." In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Miyazaki, Japan: European Language Resources Association (ELRA). <https://aclanthology.org/L18-1304>.
- Dekhissi, Laurie. 2016. "Qu'est-Ce t'as Été Te Mêler de Ça ?! Une « nouvelle » Structure Pour Les Questions Rhétoriques Conflictuelles." *Journal of French Language Studies* 26 (3): 279–98. <https://doi.org/10.1017/S0959269515000253>.
- Dekhissi, Laurie, and Aidan Coveney. 2021. "Le Contexte Linguistique Des Questions Rhétoriques Conflictuelles et La Variation Entre Pourquoi et Qu'est-Ce Que?" *Langue Française* N° 212 (4): 123–37. <https://doi.org/10.3917/lf.212.0123>.
- Delais-Roussarie, Elisabeth, Jenny Doetjes, and Petra Sleeman. 2004. "Dislocation." In *Handbook of French Semantics*, edited by Francis Corblin and Henriëtte de Swart, 451–78. CSLI Lecture Notes 170. CSLI Publications.
- Delais-Roussarie, Élisabeth, Brechtje Post, Mathieu Avanzi, Caroline Buthke, Albert Di Cristo, Ingo Feldhausen, Sun-Ah Jun, et al. 2015. "Intonational Phonology of French: Developing a ToBI System for French." In *Intonation in Romance*, edited by S Frota and P Prieto. Oxford: Oxford University Press.
- Desmets, Marianne, and Antoine Gautier. 2009. "« Comment n'y Ai-Je Pas Songé plus Tôt ? », Questions Rhétoriques En Comment?" *Travaux de Linguistique* n° 58 (1): 107–25. <https://doi.org/10.3917/tl.058.0107>.



- Di Cristo, Albert. 2013. *La prosodie de la parole*. Voix, parole, langage. Bruxelles [Marseille]: De Boeck Solal.
- Doetjes, Jenny, Georges Rebuschi, and Annie Rialland. 2004. "Cleft Sentences." In *Handbook of French Semantics*, edited by Francis Corblin and Henriëtte de Swart, 529–52. CSLI Lecture Notes 170. CSLI Publications.
- Dufter, Andreas. 2009. "Clefting and Discourse Organization: Comparing Germanic and Romance." In *Focus and Background in Romance Languages*, edited by Andreas Dufter and Daniel Jacob, 83–121. Studies in Language Companion Series 112. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/slcs.112.05duf>.
- Dufter, Andreas, and Daniel Jacob. 2009. "Introduction." In *Focus and Background in Romance Languages*, edited by Andreas Dufter and Daniel Jacob, 1–18. Studies in Language Companion Series 112. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/slcs.112.02duf>.
- Endriss, Cornelia. 2009. *Quantificational Topics A Scopal Treatment of Exceptional Wide Scope Phenomena*. Studies in Linguistics and Philosophy 86. Springer.
- Fabricius-Hansen, Cathrine, and Dag Haug, eds. 2012. *Big Events, Small Clauses: The Grammar of Elaboration*. DE GRUYTER. <https://doi.org/10.1515/9783110285864>.
- Faller, Martina. 2014. "Reportativity, (Not-) at-Issue-ness, and Assertion." *Annual Meeting of the Berkeley Linguistics Society* 40 (December): 62. <https://doi.org/10.3765/bls.v40i0.3133>.
- Farkas, Donka F. 2022. "Non-Intrusive Questions as a Special Type of Non-Canonical Questions." *Journal of Semantics* 39 (2): 295–337. <https://doi.org/10.1093/jos/ffac001>.
- Fernández-Ordóñez, Inés. 2005. "Corpus Oral y Sonoro Del Español Rural." [www.corpusrural.es](http://www.corpusrural.es).
- Fery, Caroline. 2001. "Focus and Phrasing in French." In *A Festschrift for Arnim von Stechow*, edited by Caroline Féry and Wolfgang Sternefeld, 153–81. Berlin, Boston: Akademie Verlag. <https://doi.org/doi:10.1515/9783050080116.153>.
- Fleury, Damien, and Lucia M. Tovená. 2018. "À Propos Des Lectures Des Questions En Comment." Edited by F. Neveu, B. Harmegnies, L. Hriba, and S. Prévost. *SHS Web of Conferences* 46: 12003. <https://doi.org/10.1051/shsconf/20184612003>.
- Fliessbach, Jan, Lisa Brunetti, and Hi-Yon Yoo. under review. "On the Overlapping Discourse Functions of Spanish 'cómo Que' and French 'Comment Ça' Interrogatives." *Open Linguistics*, no. Special issue: Subjectivity and Intersubjectivity in Language.
- Garassino, Davide, and Daniel Jacob. 2018. "Polarity Focus and Non-Canonical Syntax in Italian, French and Spanish: Clitic Left Dislocation and Sì Che / Sí Que-Constructions." In *The Grammatical Realization of Polarity Contrast: Theoretical, Empirical, and Typological Approaches*, edited by Christine Dimroth and Stefan Sudhoff, 227–54. *Linguistik Aktuell/Linguistics Today* 249. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/la.249.08gar>.
- Ginzburg, Jonathan. 1995. "Resolving Questions, I." *Linguistics and Philosophy* 18 (5): 459–527. <http://www.jstor.org/stable/25001601>.
- . 2012. *The Interactive Stance: Meaning for Conversation*. Oxford Linguistics. Oxford: Oxford university press.
- Ginzburg, Jonathan, Zulipiye Yusupujiang, Chuyuan Li, Kexin Ren, Aleksandra Kucharska, and Pawel Lupkowski. 2022. "Characterizing the Response Space of Questions: Data and Theory." *Dialogue & Discourse* 13 (2): 79–132. <https://doi.org/10.5210/dad.2022.203>.
- Givón, Talmy. 1983. *Topic Continuity in Discourse*. Amsterdam: John Benjamins.
- Grice, H. P. 1975. "Logic and Conversation." In *Speech Acts*, edited by Peter Cole and Jerry L. Morgan, 41–58. BRILL. [https://doi.org/10.1163/9789004368811\\_003](https://doi.org/10.1163/9789004368811_003).

- Gruber, Helmut. 2001. "Questions and Strategic Orientation in Verbal Conflict Sequences." *Journal of Pragmatics* 33 (12): 1815–57. [https://doi.org/10.1016/S0378-2166\(00\)00083-7](https://doi.org/10.1016/S0378-2166(00)00083-7).
- Halliday, M. A. K. 1967. *Intonation and Grammar in British English*: DE GRUYTER. <https://doi.org/10.1515/9783111357447>.
- Hamblin, C. L. 1973. "Questions in Montague English." *Foundations of Language* 10 (1): 41–53. <http://www.jstor.org/stable/25000703>.
- Hobbs, Jerry R. 1985. "On the Coherence and Structure of Discourse". CSLI-85-37. Center for the Study of Language and Information, Stanford University.
- Hooper, Joan B., and Sandra A. Thompson. 1973. "On the Applicability of Root Transformations." *Linguistic Inquiry* 4 (4): 465–97. <http://www.jstor.org/stable/4177789>.
- Hualde, José Ignacio, and Pilar Prieto. 2015. "Intonational Variation in Spanish. European and American Varieties." In *Intonation in Romance*, edited by Sonia Frota and Pilar Prieto, 350–91. Oxford: Oxford University Press.
- Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Studies in Linguistics Series 2. Cambridge, Mass: MIT Press.
- Jacobs, Joachim. 1984. "Funktionale Satzperspektive Und Illokutionssemantik." *Linguistische Berichte* 91: 25–58.
- . 2001. "The Dimensions of Topic-Comment." *Linguistics* 39 (4). <https://doi.org/10.1515/ling.2001.027>.
- Jary, Mark. 2010. *Assertion*. London: Palgrave Macmillan UK. <https://doi.org/10.1057/9780230274617>.
- Karssenber, Lena, and Karen Lahousse. 2018. "The Information Structure of French Il y a Clefts and c'est Clefts: A Corpus-Based Analysis." *Linguistics* 56 (3): 513–48. <https://doi.org/10.1515/ling-2018-0004>.
- Katz, Jonah, and Elisabeth Selkirk. 2011. "CONTRASTIVE FOCUS VS. DISCOURSE-NEW: EVIDENCE FROM PHONETIC PROMINENCE IN ENGLISH." *Language* 87 (4): 771–816. <http://www.jstor.org/stable/41348860>.
- Kehler, Andrew, and Hannah Rohde. 2017. "Evaluating an Expectation-Driven Question-Under-Discussion Model of Discourse Interpretation." *Discourse Processes* 54 (3): 219–38. <https://doi.org/10.1080/0163853X.2016.1169069>.
- Koshik, Irene. 2003. "Wh-Questions Used as Challenges." *Discourse Studies* 5 (1): 51–77. <https://doi.org/10.1177/14614456030050010301>.
- Kuno, Susumu. 1973. *The structure of the Japanese language*. Current studies in linguistics series 3. Cambridge, Mass.: MIT.
- Kuppevelt, Jan Van. 1995. "Discourse Structure, Topicality and Questioning." *Journal of Linguistics* 31 (1): 109–47. <http://www.jstor.org/stable/4176301>.
- Lacheret, Anne, and Jacques François. 2003. "De La Notion de Détachement Topical à Celle de Constituant Thématique Extrapositionnel." *Cahiers de Praxématique*, no. 40 (January): 167–98. <https://doi.org/10.4000/praxematique.2730>.
- Lacheret, Anne, Sylvain Kahane, and Paola Pietrandrea, eds. 2014. *Rhapsodie: A Prosodic and Syntactic Treebank for Spoken French*. Studies in Corpus Linguistics. Amsterdam: Benjamins.
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. 1st ed. Cambridge University Press. <https://doi.org/10.1017/CBO9780511620607>.
- Lefevre, Florence. 2009. "Catégorisation de Comment Subordonnant." *Travaux de Linguistique* n° 58 (1): 63–88. <https://doi.org/10.3917/tl.058.0063>.
- Leonetti, Manuel. 2014. "Spanish VSX." In *Romance Languages and Linguistic Theory 2012*:

- Selected Papers from "Going Romance" Leuven 2012*, edited by Karen Lahousse and Stefania Marzo, 6:37–64. Romance Languages and Linguistic Theory. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/rllt.6.02leo>.
- Lison, Pierre, and Jörg Tiedemann. 2016. "OpenSubtitles2016: Extracting Large Parallel Corpora from Movie and TV Subtitles." In *Proceedings of the Tenth International Conference on Language Resources and Evaluation LREC 2016, Portorož, Slovenia, May 23-28, 2016*, 923–29. Portorož, Slovenia: ELRA.
- Liu, Yingdong, Elodie Winckel, Anne Abeillé, Barbara Hemforth, and Edward Gibson. 2022. "Structural, Functional, and Processing Perspectives on Linguistic Island Effects." *Annual Review of Linguistics* 8 (1): 495–525. <https://doi.org/10.1146/annurev-linguistics-011619-030319>.
- López, Luis. 2009. *A Derivational Syntax for Information Structure*. Oxford Studies in Theoretical Linguistics 23. New York, NY: Oxford University Press.
- Mann, William C., and Sandra A. Thompson. 1988. "Rhetorical Structure Theory: Toward a Functional Theory of Text Organization." *Text - Interdisciplinary Journal for the Study of Discourse* 8 (3). <https://doi.org/10.1515/text.1.1988.8.3.243>.
- Marandin, Jean-Marie. 2006. "Structure Informationnelle." In *Sémanticlopédie*. [https://dicosem.llf-paris.fr/wiki/Structure\\_informationnelle](https://dicosem.llf-paris.fr/wiki/Structure_informationnelle).
- Marandin, Jean-Marie, Claire Beyssade, Elisabeth Delais-Roussarie, and Annie Rialland. 2002. "Discourse Marking in French: C Accents and Discourse Moves." In *Proceedings of Speech Prosody 2002*, 471–74.
- Martin, Philippe. 2009. *L'intonation Du Français*. Paris: Ophrys.
- Matić, Dejan, and Daniel Wedgwood. 2013. "The Meanings of Focus: The Significance of an Interpretation-Based Category in Cross-Linguistic Analysis." *Journal of Linguistics* 49 (1): 127–63. <https://doi.org/10.1017/S0022226712000345>.
- Mazzocconi, Chiara, Ye Tian, and Jonathan Ginzburg. 2022. "What's Your Laughter Doing There? A Taxonomy of the Pragmatic Functions of Laughter." *IEEE Transactions on Affective Computing* 13 (3): 1302–21. <https://doi.org/10.1109/TAFFC.2020.2994533>.
- McNally, Louise. 1998. "On Recent Formal Analyses of Topic." In *The Tbilisi Symposium on Language, Logic, and Computation: Selected Papers*, edited by Jonathan Ginzburg, Zurab Khasidashvili, Carl Vogel, Jean-Jacques Lévy, and Enric Vallduví, 147–60. Stanford, CA: CSLI Publications.
- . 2013. "Semantics and Pragmatics." *WIREs Cognitive Science* 4 (3): 285–97. <https://doi.org/10.1002/wcs.1227>.
- Mertens, Piet. 1987. "L'intonation Du Français: De La Description Linguistique à La Recon-naissance Automatique." Université de Louvain.
- Molnár, Valeria. 2002. "Contrast - from a Contrastive Perspective." In *Information Structure in a Cross-Linguistic Perspective*, edited by Hilde Hallelgard, Stig Johansson, Bergljot Behrens, and Cathrine Fabricius-Hansen, 147–61. Amsterdam/New York: Rodopi.
- Neeleman, Ad, and Hans Van De Koot. 2008. "Dutch Scrambling and the Nature of Discourse Templates." *The Journal of Comparative Germanic Linguistics* 11 (2): 137–89. <https://doi.org/10.1007/s10828-008-9018-0>.
- Neeleman, Ad, and Reiko Vermeulen. 2013. "Chapter 1. The Syntactic Expression of Information Structure." In *The Syntax of Topic, Focus, and Contrast*, edited by Ad Neeleman and Reiko Vermeulen, 1–38. DE GRUYTER. <https://doi.org/10.1515/9781614511458.1>.
- Ochi, Masao. 2004. "How Come and Other Adjunct Wh-Phrases: A Cross-Linguistic Perspective." *Language and Linguistics*, no. 5 (1): 29–57.
- Onea, Edgar. 2016. *Potential Questions at the Semantics-Pragmatics Interface*. Current Research in the Semantics/Pragmatics Interface 33. Leiden ; Boston: Brill.

- Portner, Paul, and Katsuhiko Yabushita. 1998. "The Semantics and Pragmatics of Topic Phrases." *Linguistics and Philosophy* 21 (2): 117–57. <https://doi.org/10.1023/A:1005311504497>.
- Potts, Christopher. 2004. *The Logic of Conventional Implicatures*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199273829.001.0001>.
- Ramm, Wiebke. 2011. "Discourse-Structural Salience from a Cross-Linguistic Perspective: Coordination and Its Contribution to Discourse (Structure)." In *Salience*, edited by Christian Chiarcos, Berry Claus, and Michael Grabski, 143–72. DE GRUYTER MOUTON. <https://doi.org/10.1515/9783110241020.143>.
- Reinhart, Tanya. 1981. "Pragmatics and Linguistics: An Analysis of Sentence Topics." *Philosophica* 27 (0). <https://doi.org/10.21825/philosophica.82606>.
- Repp, Sophie. 2016. *Contrast*. Edited by Caroline Féry and Shinichiro Ishihara. Vol. 1. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199642670.013.006>.
- Rialland, Annie, and Stéphane Robert. 2004. "La Focalisation En Wolof: Morphosyntaxe et Intonation." In *Focalisation et Moyens d'expression de La Focalisation à Travers Les Langues*, edited by A. Lacheret-Dujour et J. François, 138–60. Mémoires de La Société de Linguistique de Paris. Peeters. <https://shs.hal.science/halshs-00182177>.
- Riester, Arndt. 2019. "Constructing QUD Trees." In *Questions in Discourse*, edited by Malte Zimmermann, Klaus von Heusinger, and Edgar Onea, 2: Pragmatics:163–92. Current Research in the Semantics / Pragmatics Interface 36. Leiden: Brill.
- Riester, Arndt, Lisa Brunetti, and Kordula De Kuthy. 2018. "Chapter 14. Annotation Guidelines for Questions under Discussion and Information Structure." In *Information Structure in Lesser-Described Languages*, edited by Evangelia Adamou, Katharina Haude, and Martine Vanhove, 403–44. Studies in Language Companion Series 199. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/slcs.199.14rie>.
- Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery." Edited by Liliane Haegeman. *Elements of Grammar: Handbook in Generative Syntax*. Dordrecht: Springer Netherlands. [https://doi.org/10.1007/978-94-011-5420-8\\_7](https://doi.org/10.1007/978-94-011-5420-8_7).
- . ed. 2010. *The Structure of CP and IP*. Repr. The Cartography of Syntactic Structures 2. Oxford: Oxford Univ. Press.
- Roberts, Craige. 1996. "Information Structure: Towards an Integrated Formal Theory of Pragmatics." Edited by Jae Hak Yoon and Andreas Kathol. *OSUWPL, The Ohio State University Department of Linguistics* Volume 49: Papers in Semantics.
- . 2012. "Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics." *Semantics and Pragmatics* 5 (December). <https://doi.org/10.3765/sp.5.6>.
- Rohde, Hannah. 2006. "Rhetorical Questions as Redundant Interrogatives." *San Diego Linguistics Papers*, no. 2: 134–68.
- Rooth, Mats. 1992. "A Theory of Focus Interpretation." *Natural Language Semantics* 1 (1): 75–116. <https://doi.org/10.1007/BF02342617>.
- Rosemeyer, Malte. 2021. "Two Types of Constructionalization Processes in Spanish and Portuguese Clefted *Wh* -Interrogatives." *Studies in Hispanic and Lusophone Linguistics* 14 (1): 117–60. <https://doi.org/10.1515/shll-2021-2042>.
- Sadock, Jerrold M. 1971. "Queclaratives." *Papers from the Regional Meeting of the Chicago Linguistic Society*, no. 7: 223–31.
- . 1974. *Toward a Linguistic Theory of Speech Acts*. New York: Academic Press.
- Samek-Lodovici, Vieri. 2005. "Prosody-Syntax Interaction in the Expression of Focus." *Natural Language & Linguistic Theory* 23 (3): 687–755. <http://www.jstor.org/stable/4048112>.
- Schlenker, Philippe. 2023. "Supplements without Bidimensionality." *Linguistic Inquiry* 54 (2): 251–97. [https://doi.org/10.1162/ling\\_a\\_00442](https://doi.org/10.1162/ling_a_00442).

- Schwarzschild, Roger. 1999. "Givenness, AvoidF and Other Constraints on the Placement of Accent." *Natural Language Semantics* 7 (2): 141–77. <http://www.jstor.org/stable/23748636>.
- Selkirk, Elisabeth. 1996. "Sentence Prosody: Intonation, Stress, and Phrasing." In *The Handbook of Phonological Theory*, edited by John A. Glodsmith, 550–69. Oxford, UK: Blackwell Publishing Ltd. <https://doi.org/10.1111/b.9780631201267.1996.00018.x>.
- Shochi, Takaaki, Albert Rilliard, and Véronique Aubergé. 2009. "Intercultural Perception of English, French and Japanese Social Affective Prosody." In *The Role of Prosody in Affective Speech*, 97:31–60. Linguistic Insights - Studies in Languages and Communication. Peter Lang.
- Simons, Mandy, Judith Tonhauser, David Beaver, and Craige Roberts. 2010. "What Projects and Why." *Semantics and Linguistic Theory* 20 (August): 309. <https://doi.org/10.3765/salt.v20i0.2584>.
- Smirnova, Anastasiia. 2021. "Les Questions En Ça et Donc: Études Empiriques." Université Paris Cité.
- Smirnova, Anastasiia, and Anne Abeillé. 2021. "Question Particles Ça and Donc in French: A Corpus Study." *Linguistic Research* 38 (2): 239–69. <https://doi.org/10.17250/KHISLI.38.2.202106.003>.
- Steedman, Mark. 2000. "Information Structure and the Syntax-Phonology Interface." *Linguistic Inquiry* 31 (4): 649–89. <http://www.jstor.org/stable/4179127>.
- Tellier, Christine, and Daniel Valois. 2006. *Constructions méconnues du français*. Presses de l'Université de Montréal. <https://doi.org/10.4000/books.pum.10250>.
- Thiberge, Gabriel, Lisa Brunetti, and Jan Fliessbach. submitted. "Différences fonctionnelles entre quasi-synonymes: les interrogatives de raison en français."
- Tonhauser, Judith, David Beaver, Craige Roberts, and Mandy Simons. 2013. "Toward a Taxonomy of Projective Content." *Language* 89 (1): 66–109.
- Truckenbrodt, Hubert. 2007. "The Syntax-Phonology Interface." In *The Cambridge Handbook of Phonology*, edited by Paul de Lacy, 435–56. Cambridge: Cambridge University Press.
- Umbach, Carla. 2004. "On the Notion of Contrast in Information Structure and Discourse Structure." *Journal of Semantics* 21 (2): 155–75. <https://doi.org/10.1093/jos/21.2.155>.
- . 2005. "Contrast and Information Structure: A Focus-Based Analysis of But." *Linguistics* 43 (1). <https://doi.org/10.1515/ling.2005.43.1.207>.
- Vallduví, Enric. 1992. *The Informational Component*. Taylor & Francis.
- Vallduví, Enric, and Elisabet Engdahl. 1996. "The Linguistic Realization of Information Packaging." *Ling* 34 (3): 459–520. <https://doi.org/10.1515/ling.1996.34.3.459>.
- Velleman, Leah, and David Beaver. 2016. "Question-Based Models of Information Structure." In *The Oxford Handbook of Information Structure*, edited by Caroline Féry and Shinichiro Ishihara. Vol. 1. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199642670.013.29>.
- Venhuizen, Noortje Joost, Johan Bos, Petra Hendriks, and Harm Brouwer. 2014. "How and Why Conventional Implicatures Project." *Semantics and Linguistic Theory* 24: 63. <https://doi.org/10.3765/salt.v24i0.2392>.
- Vermeulen, Reiko. 2013. "Chapter 3. Word Order Variation and Information Structure in Japanese and Korean." In *The Syntax of Topic, Focus, and Contrast*, edited by Ad Neeleman and Reiko Vermeulen, 77–118. DE GRUYTER. <https://doi.org/10.1515/9781614511458.77>.
- Villalba, Xavier. 2000. *The Syntax of Sentence Periphery*. Cerdanyola del Vallès (ES): Servei de publicacions de la UAB.
- Webber, Bonnie, Rashmi Prasad, Alan Lee, and Aravind Joshi. 2019. "The Penn Discourse Treebank 3.0 Annotation Manual." [https:// catalog.ldc.upenn.edu/docs/LDC2019T05/](https://catalog.ldc.upenn.edu/docs/LDC2019T05/)

PDTB3-Annotation-Manual.pdf.

Williams, Edwin. 1997. "Blocking and Anaphora." *Linguistic Inquiry* 28 (4): 577–628.  
<http://www.jstor.org/stable/4178996>.

Zubizarreta, Maria Luisa. 1998. *Prosody, Focus, and Word Order*. Cambridge, Mass: MIT Press.