The Dissertation Committee for Emilie Destruel
certifies that this is the approved version of the following dissertation:

The French *C'est*-cleft: Empirical studies of its meaning and use

Committee:

<u>David I. Beaver, Supervisor</u>

<u>Richard P. Meier</u>

<u>Scott Myers</u>

<u>Stephen Wechsler</u>

<u>David Birdsong</u>

The French *C'est*-cleft: Empirical studies of its meaning and use

by

Emilie Destruel , DEA; M.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2013

This dissertation is dedicated to my husband Chad, and my parents Anne-Marie and François.

Acknowledgments

I wish to thank the many people who helped me get this far. Of course, my advisor David Beaver whose comments and feedback through this research have been invaluable. I also want to thank him for encouraging me to always be positive about my work even when I felt things appeared to be going wrong. I am so grateful for Richard Meier and Jean-Pierre Montreuil who have been extremely generous with their time, support and feedback all along. Their doors were always open and I know I took advantage of it maybe one too many times. Special thanks to David Birdsong for sharing his expertise on experimental design. Finally, I wish to thank Steve Wechsler and Scott Myers for their useful comments throughout the progress of my work.

I would also like to thank the members of the Beaver Lab, Christopher Brown, Justin Cope, Joey Frazee, Daniel Velleman and Tony Wright for their feedback on the various pilots involved in this research. Of course, I also wish to thank many other graduate students in the Linguistics Department but also in the French & Italian Department for the numerous discussion sessions where we shared issues on our research: Qiuana Lopez, Gabriela Garcia Salido, Stacy Chen, Nick Gaylord, Elias Ponvert and Nicolas Bacuez. And of course my cube mate and dear friend Stéphanie Villard for sharing the past seven and a half years of graduate life with me. Special thanks to my favorite québecoise Julie Rinfret for her advice and encouragement while she was at UT.

Further thanks to all the participants of my experiments, especially the wonderful moms Isabelle Bossart, Isabelle Bettler, Claire DeCoatpont and Claire Graby, and all the people who gave me precious feedback at the Chicago Linguistic Society conference in

2010 and at the CSSP conference in 2011.

Last but not least, I would like to thank my husband Chad Johnson for his endless support and going through this process by my side, and my parents for being so present even five thousand miles away.

The French *C'est*-cleft: Empirical studies of its meaning and use

Publication No. _____

Emilie Destruel , Ph.D.
The University of Texas at Austin, 2013

Supervisor: David I. Beaver

This dissertation contributes to a fuller description of the French *c'est*-cleft by reporting on three empirical studies on its meaning and use, and presenting a unified account of the cleft couched in Stochastic Optimality Theory. The first two studies in this dissertation explore the meaning of the cleft, more specifically the *exhaustive* meaning. First, the results from a forced-choice task, designed to test the level of exhaustivity of the cleft compared to exclusive sentences and canonical sentences, show that the cleft does not behave like the other two sentence forms. This is taken to indicate that the exhaustivity associated with the cleft is not truth-conditional. Instead, I argue that exhaustivity arises from a pragmatic constraint on the way speakers use language. This argument is supported further in the second study, a corpus study that shows there is no categorical ban on the type of NP that can occur in post-copular position in a cleft. In fact, the cleft interacts felicitously with a number of expressions such as universal quantifiers and additives, which have been claimed to never appear in post-copular position. This corpus study further shows that the primary aspect of the cleft is not to convey exhaustivity, but instead to convey contrast or correction. Finally, the third study, a semi-spontaneous

production experiment, helps make precise the situations in which an element is clefted. The results demonstrate that there is a clear asymmetry between the way grammatical subjects or non-subjects are marked: focused subjects are mostly clefted whereas focused non-subjects generally remain in situ. Moreover, the experiment shows that there exists some amount of free variation: subjects can be realized via prosody and non-subjects can be clefted. I conclude my research by proposing that the non-random alternation cleft/canonical is not a categorical phenomenon, but is gradient and explained by a set of constraints on French' syntax, prosody and pragmatics. The cleft is used to provide contrast or a total answer to the question under discussion.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

This dissertation provides a unified account of the meaning and use of the French *c'est*-cleft construction couched in the framework of Stochastic Optimality Theory (Boersma and Hayes, 2001). This construction, to which the term "cleft" was originally applied by Jespersen (1927), is illustrated by the examples in (1)-(3).[1]

(1) C'est LE PETIT   qui est tombé dans l'   escalier!
     It-is THE LITTLE-ONE who is fallen in   the stairs.
     'It's THE LITTLE ONE who fell down the stairs.'

(2) C'est DEMAIN   que je partirai.
     It-is TOMORROW that I   will-leave.
     'It's TOMORROW that I'll leave.'

(3) C'est DANS CETTE MAISON que nous avons vécu pendant dix ans.
     It-is IN   THIS HOUSE   that we   have lived during   ten years.
     'It's IN THIS HOUSE that we've lived for ten years.'

The *c'est*-cleft can be linearly analyzed as in (4).

(4) pronoun - copula - focused XP (pivot) - relative clause

---

[1]Throughout this dissertation, I will indicate the intended focus of the sentence with small capitals, partly in order to distinguish between clefts where the pivot is focused from non-cleft copular sentences involving restrictive modification of the postcopular element.

The goal of this dissertation is twofold: to add to the body of French and cross-linguistic literature about the marking of focus, and to elucidate the nature of the meaning and use of the *c'est*-cleft. In very broad terms, the overarching research questions addressed in this dissertation are: (1) What is the nature of the exhaustivity associated with the cleft, (2) What does the cleft add to a sentence that a canonical sentence alone cannot do, (3) What are the factors governing the alternation cleft/canonical sentence in the marking of focal information?

Over several decades, researchers have focused much of their attention on different aspects of the cleft construction, both in French but also cross-linguistically for similar constructions such as the English *it*-cleft. One of the questions that led to a large amount of work concerns the meaning of the cleft. In other words, what does the cleft contribute semantically, if anything, that sets it apart from a simple SVO sentence? Under many previous analyses, the cleft is argued to semantically contribute *exhaustivity*, such that the cleft sentence has the same truth-conditions as a sentence with an exclusive like 'seulement/only' (Kiss, 1998; Clech-darbon *et al.*, 1999). This is illustrated in (5), where the two sentences are argued to be semantically equivalent and to assert 'Nobody else than Jean came to the party'.

(5) a. Seul JEAN est venu à la soirée.
         Only JEAN is came to the party.
         'Only JEAN come to the party'.

     b. C'est JEAN qui est venu à la soirée.
         It-is JEAN who is came to the party.
         'It's JEAN who came to the party.'

     = *Nobody else than Jean came to the party.*

However, I argue that this analysis is not sustainable. Instead, I follow Horn's (1981) analysis according to which *exhaustivity* is implicated, and arises due to pragmatic factors in the language. It is interesting to note that, while never dealing directly with the issue, the definition that Lambrecht (2001) adopts for clefts implicates that *exhaustivity* is not truth-conditional, by stating that canonicals and clefts have the same truth-conditions. Indeed, in his definition, Lambrecht takes the cleft to be constituted of a matrix and a relative, which together "express a logically simple proposition, which can also be expressed in the form of a single clause **without a change in truth conditions**." In this work, I provide experimental evidence showing that the French *c'est*-cleft does not behave like exclusive-sentences . . . but nor does it behave exactly like canonical sentences. The results from the experiment, along with a corpus study on distributional patterns of certain expressions within the cleft lead me to argue that the cleft contributes use-conditional instead of truth-conditional *exhaustivity*. This use is governed by a pragmatic constraint which demands that pivots be interpreted as total answers (exhaustively). Ranked high, such a constraint will require pivots to be interpreted exhaustively, whereas it will allow non-exhaustive readings when ranked lower, or over-ranked by other constraints.

The second type of research question on which scholars have concentrated is the condition on the cleft's usage. The seminal work of Lambrecht started an hypothesis about the use of this construction, which is that the cleft occurs because spoken French disallows lexical subjects in sentence-initial positions (Lambrecht, 1994, 2001). The author also argues that French must resort to different types of syntactic reordering when marking focus. Taken together, these claims have influenced French scholars but also scholars working cross-linguistically and who compare French to other languages (Dufter,

2009). Thus, French is widely assumed to be a language that resorts to clefting as its main focus marking strategy. However, this view has been challenged by various recent studies on French. Notably, Féry (2001) and Claire *et al.* (2004) argue that subject-focus can be realized via prosody: the subject-focus is wrapped into an independent Intonational Phrase delimited by a low boundary tone on its right edge. Hupet and Tilmant (1986) and Beyssade *et al.* (2011) show that non-subjects do not require syntactic reordering, and are mainly realized in situ via prosody. Finally, Hamlaoui (2008) argues that a *c'est*-cleft is only required in demotic French to mark focus when answering a Qui/Who-question. She argues, contra Lambrecht, that the cleft is in fact prosodically motivated; when used, it creates an independent Intonational Phrase (IP), and thus fulfills the language's requirement for having main stress occur on the right edge of IPs. Using semi-spontaneous eliciting methods, the second experiment in this dissertation is designed to settle the debate on focus realization by providing empirical evidence for some of the conditions under which the cleft occurs.

To the best of my knowledge, this dissertation is one of the very few controlled experimental studies on the use of the cleft in focus marking (Hupet and Tilmant, 1986, 1990; Féry, 2001), and one of the first to empirically examine the exhaustive inference associated with it. Such studies help establish the link between observed linguistic phenomena and recordable spontaneous data. In addition, the production evidence adds to the discussion on the marking of focus information in French, and the debate on the role of clefting in that phenomenon. Moreover, the interpretation evidence will add to the cross-linguistic debate on the nature of exhaustivity by providing data for another language. Overall, the studies in this dissertation add to our greater knowledge of the way

information structure interacts with syntax, prosody and pragmatics, and how these domains interact with each other.

1.1 Motivation for Research

Why another study on French? French is typologically interesting because it does use clefting and syntactic reordering to a much greater extent than related Romance languages and English (Dufter, 2009). Within a typological classification of focus realization such as the one proposed in Büring (2009), many scholars would probably argue that French should be classified under the '*Egde*' category characterized by (6) in Büring (2009, 14):

(6) In an 'Edge' category, focus is marked by non-standard constituent order, with the focus in left- or right-peripheral position.

Yet, while writing this dissertation, I came across empirical evidence that makes the membership to this category harder to maintain. The major problem is that French exhibits a clear marking asymmetry between subject and non-subject focus which makes it difficult to commit to one specific category in a classification. For subject-focus, my work, along with past studies (Lambrecht, 1994; Hupet and Tilmant, 1990; Katz, 1997) would agree on classifying French as an '*Edge*' language because clefting is used, in that case, more than 85% of the time.[2] But things become trickier when looking at non-subject focus. The results from the production experiment presented in this dissertation, in line with

---

[2]In Chapter 5, I provide exact percentages for the use of clefts in different conditions.

previous studies such as Hupet and Tilmant (1986), Vion and Colas (1995) and Beyssade *et al.* (2011), suggest that French behaves more like '*Boundary*' languages where focus is marked by insertion of a prosodic phrase boundary to the left or right of the focus.

Now more specifically, let me motivate each part of the dissertation. First, the study on the nature of exhaustivity in the *c'est*-cleft, found in Chapter 3, is motivated by the lack of empirical discussion about this issue in French. While there is an extensive body of literature on the issue in languages like English and Hungarian, very few studies propose an account in the French literature, and yet, ironically, almost all of the studies on *c'est*-clefts take exhaustivity as being associated with the construction. To my knowledge, Clech-darbon *et al.* (1999) is the only study that gives a formalization of this meaning, arguing for a truth-conditional contribution. This formalization is opposite to the one proposed in this dissertation.

The corpus study presented in Chapter 4 is motivated by two factors. First, the study was designed to evaluate the claim that certain expressions (e.g. quantifiers, numerals, additives, indefinite pronouns, etc.) are banned from the cleft pivot position (Katz, 1997; Kiss, 1998). The quantitative study aims to expand on previous accounts by providing empirical evidence evaluating this claim, and to demonstrate that it is too strong. Second, the motivation for the qualitative study lies in the lack of in-depth analysis of empirical data at the discourse level. Various studies on French rely too often on constructed examples, or examples involving the cleft sentence alone.[3] In my dissertation, I

---

[3]Hupet and Tilmant (1990) is one of the few studies that analyze the role of context in their experimental design. Their methodology differs from the one I use in this work since they propose a forced-choice task where the participants choose which form, either a cleft or a canonical sentence, is the most appropriate given the context.

expand on past studies by looking at the use of the cleft in a larger discourse context. This allows me to better understand the situation in which the cleft is introduced, importantly the mental state of the interlocutors. By analyzing the surrounding context, I demonstrate that all cases where the expressions claimed to be banned in fact occur in the cleft pivot position, are pragmatically special: they all convey a (explicit) contrast between the speaker's expectations and the speaker's beliefs about the hearer's expectations. Additionally, in all these cases, the pivot element is taken to be the most pragmatically important element of the sentence, often because of its low discourse expectation (i.e. the pivot is not expected by the addressee). Relating these observations to the results from the study on exhaustivity, I conclude that *contrastiveness* plays a more important role than other factors, and that a constraint on that pragmatic factor will overrule the constraint on exhaustive interpretation.

A further note: register has been argued to have a clear effect on the sentence form produced (Hamlaoui, 2009). My corpus analysis includes data both from formal register (taken from the EUROPARL corpus, which is extracted from the proceedings of the European Parliament), and from informal/conversational register (taken from Google hits on a variety of personal blogs, and from the PFC corpus which includes transcripts of spontaneous conversations). However, I do not make any claims about the role of register.

Thirdly, the production experiment reported on in Chapter 5 is motivated by the lack of controlled semi-spontaneous data in the literature on French. Many studies have applied experimental methods to test focus realization in French (Hupet and Tilmant, 1986; Vion and Colas, 1995; Féry, 2001; Beyssade *et al.*, 2011), but to my knowledge, my experiment benefits from two design characteristics: it is the first study that looks at a

large range of grammatical focus types including subject, direct object, indirect object, predicate and sentence-focus. Past studies have often restricted their attention to narrow focus on subjects and objects, or broad focus on sentences. It is also one of the first studies that elicits semi-spontaneous data. Many of the past studies have used designs such as forced-choice judgment tasks or reading tasks.

Finally, the use of Stochastic Optimality Theory to develop an account of the alternation cleft/canonical is motivated by the fair amount of variation observed in the data collected in my work. Here again, my dissertation is the first study to account for default focus realizations but also for the (free) variation. Hamlaoui (2008) proposes an analysis of the emergence of the cleft couched in Standard Optimality Theory, but her analysis relies on constructed examples and does not explain the variation that actually occurs in empirical data.

1.2   Research Questions

Taking past studies into account, the goal of this dissertation is to address the following research questions:

(1) What is the nature of the exhaustivity associated with the *c'est*-cleft? Put differently, is the exhaustivity truth-functional or does it arise from pragmatic effects on the way speakers use the cleft in language?

(2) If the cleft does not add *exhaustivity* semantically, what other factors influence a speaker's choice when he make a choice between a cleft or a canonical form? When do speakers use a cleft? And relatedly, what does the cleft add to a sentence that the canonical alone does not do?

(3) Is it the case that the cleft is always associated with one type of focus? In other words, is there a strict one-to-one relationship between focus and its realization? (4) How are the factors governing the alternation cleft/canonical ranked with respect to each other? (5) In French, are there constraints on the type of expressions that may be realized either in pivot or in subject position?

1.3   Hypotheses

In relation to the research questions cited above, the following hypotheses are tested in this work:

• The exhaustivity associated with the French *c'est*-cleft is not truth-conditional. The *c'est*-cleft will not behave like a sentence that includes an exclusive like 'seulement/only'.

• The cleft and the canonical are, for the most part, in complementary distribution, where there are contexts in which a cleft is infelicitous and a canonical is the option used. The cleft emerges in special pragmatic cases that involve the need for a total answer and/or a contrast in the expectations of speaker and addressee.

• There is no strict one-to-one relationship between focus and its realization. The cleft is used mostly, but is not restricted to marking subject-focus and contrast.

• Pragmatic constraints over-rank constraints on prosodic and syntactic well-formedness.

• There is no categorical ban on the type of expression that can occur in pivot position. All the expressions that are argued, in prior literature, to semantically clash with the cleft do occur in pragmatic cases where contrast is conveyed.

1.4 Outline of Dissertation

Chapter 2 provides a more in-depth look at the background literature on notions that are central to the topic of this dissertation, including more discussion on Information Structure, focus and contrastiveness, but also a background discussion on the phonology of French.

Chapter 3 presents the experiment on exhaustivity and discusses previous analyses of that inference. The chapter ends by proposing a use-condition constraint for the cleft which will be later integrated into the full OT analysis developed in Chapter 6.

Chapter 4 presents the corpus study. This study includes both a quantitative analysis examining the ratio of occurrence of given expressions in the cleft vs. canonical, and a qualitative analysis investigating in which discourse contexts a cleft occurs, and why a canonical sentence is pragmatically infelicitous in these contexts. The chapter ends by proposing pragmatic constraints derived from the findings of the corpus study.

Chapter 5 turns to the production experiment and includes a discussion of past syntactic and prosodic analyses of focus marking in French.

Chapter 6 develops the Stochastic Optimality Theoretic account of the alternation cleft/canonical integrating constraints on prosody and syntax as well as the pragmatic constraints posited through the chapters.

Chapter 7 synthesizes the results of the experiments, concludes and discusses ways to expand the work in new directions, including in the processing field.

Chapter 2

Theoretical Background

The goal of this chapter is to give an overview of three domains that interact in the study of the French *c'est*-cleft, namely Information Structure, Syntax and Prosody.

2.1   Information Structure

The primary notion that this dissertation calls on is the notion of *Information Structure* (IS). The role and status of IS have been extensively studied, which led to different conceptualizations and conventions. Chafe (1976) refers to IS as 'information packaging', which he views as the speaker's assessment of the hearer's ability to process the new information conveyed to that hearer in contrast to the background information. Prince (1981a) also uses the term 'information packaging', and takes it to be the tailoring of an utterance by a sender to meet particular assumed needs of an intended receiver. Thus, IS in natural languages reflects the sender's hypotheses about the receiver's assumptions, beliefs and strategies. Lambrecht (1994) claims that IS is the formal expression of the pragmatic structuring of a proposition in a discourse, taking IS to be an independent component within the grammar of a language, and defined as in (2.1). Therefore, studying IS involves studying the relationship between linguistic forms and the mental states of discourse participants.

(7) '… that component of sentence grammar in which propositions as conceptual representations of states of affairs are paired with lexicogrammatical structures in accordance with the mental states of interlocutors who use and interpret these structures as units of information in given discourse contexts.' (Lambrecht, 1994, 5)

This dissertation is directly concerned with the study of IS, since the goal is to show how the form of sentences and the discourse contexts in which sentences are used to communicate pieces of propositional information are linked. Following Lambrecht (1994), three important components of IS must be distinguished: (1) *presuppositions and assertions*, which concern the structuring of propositions into portions which a speaker assumes an addressee already knows or takes for granted, or which the addressee does not know yet or is not expected to know; (2) *identifiability and activation*, which concern a speaker's assumption about the status of the mental representations of discourse referents in the addressee's mind at the time of utterance; and (3) *topic and focus*, which relate to a speaker's assessment of the relative predictability vs. unpredictability of the relations between the propositions and their elements in given discourse situations. These three categories correspond to the three senses of "givenness" discussed in Prince (1981a). In this dissertation, the notions most relevant to the study of the meaning and use of the *c'est*-cleft are the notions of *focus, presupposition* and *assertion*. In the next subsection, I define these notions in more details, and show how they apply to the French *c'est*-cleft construction.

12

2.1.1 Presupposition and Assertion

The notion of presupposition is defined differently depending on whether it is understood as a semantic notion or a pragmatic one. In semantic approach of presupposition, following the Frege-Strawson tradition, one sentence presupposes another if and only if whenever the first is true or false, the second is true.[1] Under that view, presupposition is found at the word and sentence level. One of the well-known examples to illustrate the notion of presupposition is (8):

(8) The king of France is bald.
→ There exists a king of France.

The major counterpoint to a semantic approach à la Frege-Strawson is proposed by Stalnaker (1974), who argues for a pragmatic approach of presupposition, claiming that it has to do with what people take for granted when they are speaking. In uttering a sentence, the speaker expresses a pragmatically structured proposition which reflects what she expects to hold of the common ground between her and the addressee. In other words, in uttering a sentence the speaker makes assumptions about the addressee's state of mind and beliefs at the time of utterance. In a sentence, the specific piece of information which is assumed to be part of the command ground to all discourse participants is understood as being the *pragmatic presupposition*. It is important to note that Stalnaker does not perceive pragmatic presupposition to be in direct opposition with a semantic approach to presupposition. However, whenever something is semantically presupposed, it

---

[1]I refer the reader to Strawson (1950) and Frege (1892).

13

should also be expected to be pragmatically presupposed. The Stanford Encyclopedia[2] entry on presupposition notes: "if a definite description semantically presupposes the existence of a suitable referent, then it follows that speakers using definites will pragmatically presuppose the existence of such referents."

Lambrecht (1994, 2000), and Lambrecht and Michaelis (1996) follow this pragmatic approach to presupposition and propose the definition in (9), whereby the part of discourse which is the assumed knowledge state of an addressee at the time of utterance is presupposed:

(9) PRAGMATIC PRESUPPOSITION: The set of propositions lexico-grammatically evoked in a sentence which speaker assumes the hearer already knows or believes or is ready to take for granted at the time the sentence is uttered (the 'old' information).[3]

On the other hand, the information in the sentence which is not taken as presupposed, but instead which is 'new', constitutes what Lambrecht refers to as *pragmatic assertion*. He proposes the following definition:

(10) PRAGMATIC ASSERTION: The proposition expressed by a sentence that the speaker expects the hearer to know or believe or take for granted as a result of hearing the utterance (the 'new' information).

---

[2]Beaver, David I. and Geurts, Bart, "Presupposition", The Stanford Encyclopedia of Philosophy (Summer 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2011/entries/presupposition/>.

[3]This 'Pragmatic Presupposition' is also referred to as 'Knowledge Presupposition' by Lambrecht (2000).

So in the exchange in (11) adapted from Lambrecht (2000, 475), the pragmatic presupposition corresponds to the set of propositions evoked by the clause 'what I need is', in other words the clause expressing that the speaker needs 'x', while the pragmatic assertion corresponds to the proposition that 'x = a sheet of paper and a pencil'.

(11) a. A: De quoi as   tu   besoin?
          A: Of what have you need?
          'A: What do you need?'

      b. B: Ce   dont   j' ai   besoin est d' un papier et   d' un stylo.
          B: That rel-pro I have need   is of one paper and of one pencil.
          'B: What I need is a sheet of paper and a pencil.'

Of course, (11b) is one way the speaker could structure her answer, but the possible answers in (12) are just as felicitous, and also carry the presupposition and assertion that (11b) carries.

(12) a. J' ai besoin d'un papier et d'un stylo.

      b. C'est d'un papier et d'un stylo dont j'ai besoin.

      c. D'un papier et d'un stylo, j'ai besoin.

Lambrecht (1994) proposes that one way to test what is part of the pragmatic assertion and what is presupposed is for the addressee to contradict the speaker's sentence by continuing the discourse with 'That's not true'.[4] The part affected by the contradiction is the assertion (that the things the speaker needs are paper and a pencil), while the presupposition survives (it is still true that the speaker needs something).

---

[4]The lie-test was formulated as a test to distinguish assertion from the rest of the sentence by Erteschik-Shir (1997).

2.1.2 Topic and Focus

2.1.2.1 Topic

Topic is often thought of as the complement of focus. Indeed, a common distinction in the literature is to talk about 'old/new' or 'ground/focus' information. Here, I will briefly introduce the notion of topic as it is discussed in Lambrecht (1994) but will concentrate on discussing the notion of *focus* in much greater details.

Topic has often been loosely defined as "what the sentence is about", and the focus as what is being said about that topic. In his Ph.D dissertation, Lambrecht (1986) argues that a pragmatic definition of topic is necessary because it cannot be identified on the basis of grammatical form alone. In other words, there is no direct correlation between the grammatical role of a constituent and its topicality. To give a concrete example, not all grammatical subjects are topics. Thus, the knowledge of the discourse context and the mental states of speakers are required to identify the topic of a sentence. Lambrecht distinguishes between *topic* and *topic expression* in order to not confuse the information-structural notion with its grammatical realization:

(13)   a. TOPIC: A referent is interpreted as the topic of a proposition if in a given situation the proposition is construed as being about this referent, i.e. as expressing information which is relevant to and which increases the addressee's knowledge of this referent.

   b. TOPIC EXPRESSION: A constituent is a topic expression if the proposition expressed by the clause with which it is associated is pragmatically construed as being about the referent of this constituent.

16

Prince (1983) shows that aboutness is not the only property of topichood, but that *givenness* and *recoverability* (i.e. being previously mentioned in discourse) is also associated with topics. Her 1981 scale of Assumed Familiarity (Prince, 1981b) provides means of assessing the accessibility of a given referent according to how easily it can be recovered from the previous context. In her model, illustrated in Table (3.1), a referent can be new, inferable or evoked depending on whether the discourse context allows for its activation in the mind of the addressee.

New	Brand-new, anchored	*A guy that I used to work with* won the lottery.
	Brand-new, unanchored	I had to borrow *a pen*.
	Unused	*Barack Obama* is running for president.
Inferable		I went back to the store but *the clerk* could't help me.
Evoked	Textual	Jane had lunch at a new restaurant and *she* hated the food.
	Situational	The woman on the phone wanted to ask *you* something.

Table 2.1: Prince's Assumed Familiarity

In his 1994 book, Lambrecht builds on Prince (1981b)'s scale and on Chafe (1976)'s notion of activation states to propose a scale of topic accessibility illustrated in Table (2.2). According to this scale, discourse referents that are the most active in discourse will constitute better topics than some that are brand-new and unanchored in the previous discourse.

Active	most acceptable
Accessible	
Unused	
Brand-new anchored	
Brand-new unanchored	least acceptable

Table 2.2: Lambrecht's scale of Topic Accessibility

In her 2003 Ph.D dissertation, Cowles summarizes these scales of familiarity by stating that a topic must have a referent whose existence is presupposed, which is accessible, and which the sentence is 'about'.

2.1.2.2 Focus

The literature on *focus* can be confusing because scholars have used this term to refer to a wide variety of phenomena, such as prominence (a psychoacoustic notion), focus marking (a prosodic or intonational notion), F-marking (a syntactic notion) and semantic focus (a notion concerned with the meaning of focus). Moreover, the term focus has been associated with a wealth of other terms including 'new', 'emphasis', 'stress', 'rheme', 'comment', 'accented', 'prominent', 'stress', 'informative' and 'contrast'. Thus, some scholars failed to clearly distinguish between focus as a category at the information-structure level, focus referring to the way it is interpreted in language, and focus referring to the way it is realized in language (either via syntactic, prosodic and/or morphological means). While this dissertation does examine all three aspects of focus in French, the present section intends to give a working definition of focus at the information-structural level. I refer the reader to Chapter 3 for a richer background discussion on the interpretation of focus (specifically the exhaustive interpretation associated with the cleft construction), and Chapter 5 for details on focus realization (specifically the realization of different grammatical and pragmatic focus types).

So what exactly is *focus*? Let's consider the example in (14) for a moment.

(14)   a. Qu'   est-ce-que Jean a   mangé pour le   petit-déjeuner?
          What is-is-that   John has eaten   for   the breakfast?

'What did John eat for breakfast?'

b. Jean a mangé DES OEUFS.
John has eaten SOME EGGS.
'John ate some eggs.'

c. *JEAN a mangé des oeufs.
*JOHN has eaten some eggs.
'*JOHN ate some eggs.'

If we consider the SMALL CAPS to signal the fact that this element is somehow marked more saliently than the rest of the sentence, then it is reasonable to rule out (14c) as a felicitous answer to the question in (14a). The most common way to identify a focus is by ways of question-answer pair: focus is that part of the sentence which resolves the variable instantiated by the wh-item in a previous (explicit) question. Yet linguists have offered a variety of definition for focus.

In generative linguistics, Chomsky and Halle (1968) proposed that focus is a feature associated to a single word in a sentence. They posited the Nuclear Stress Rule which demanded that this single word in the sentence carried prosodic prominence, i.e. the main stress in the sentence. In later syntactic analysis, scholars such as Jackendoff (1972) and Selkirk (1995) argue that focus is a feature associated with a structural position in the syntactic representation of a sentence. Another approach to focus and focus interpretation is found in the field of semantics, especially in Rooth (1992, 1996) and von Stechow (1991). These researchers concentrate on how focus directly affects meaning. In his seminal paper on focus interpretation, Rooth (1992) proposes a theory of focus known as Alternative Semantics, which defines focus is in terms of a focus semantic value which contrasts with the ordinary semantic value of a sentence. The focus semantic value is fur-

19

ther defined in terms of a set of relevant alternatives from which the ordinary meaning value is drawn. To give a concrete example, in (14b), the ordinary semantic meaning is the binary relation represented in (15a) and the set of relevant alternatives for the interpretation of the sentence is the set in (15b). By uttering (14b), the speaker indicates that the ordinary meaning is true and rejects the other alternatives.

(15) Jean a mangé DES OEUFS.

   a. mangé⟨ Jean, oeufs⟩

   b. {mangé⟨ Jean, oeufs⟩, mangé⟨ Jean, pain⟩, mangé⟨ Jean, céréales⟩}

   Krifka (2007) formulates the definition in (16):

(16) Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions.

However, Krifka notices that the definition in (16) is silent about the exact nature of the alternatives relevant for interpretation. In the literature on focus, calculating the alternatives, that is defining what exactly is their form and nature have proven to be more challenging than expected. But that discussion goes beyond the scope of this work, and I refer the reader to (Krifka, 2007).

Pragmatic approaches to focus such as Lambrecht (1994, 2001) define focus independently of structure and prosody. What is concentrated on are the ways in which mental states of interlocutors influence language use. Lambrecht (1994) proposes the definition of focus in (17) based on the pragmatic notions of presupposition and assertion:

(17) FOCUS: That component of a pragmatically structured proposition whereby the pragmatic assertion differs from the presupposition. The focus component is by definition an unpredictable part of the proposition.

Under this view, the focus of a proposition is that denotatum whose presence in the sentence makes the utterance into an assertion, that is, makes it possible for the sentence to convey information unknown to the addressee at the time of utterance (Lambrecht, 2001, 474). But this definition suffers from a shortcoming. Indeed, many sentences that have a focus are not assertions: for example imperatives as in (18) and questions as in (19). The definition Lambrecht is adopting seems to confuse how languages mark what is *at-issue*, that is what is the main point of a sentence,[5] and how they mark what speech act is being performed.

(18) Listen to what YOUR MOTHER is telling you!

(19) Who FINISHED the beans?

In this work, I will follow a pragmatic definition of focus based on the notion of *predictability* but also on the notion of *importance*, as discussed in Beaver and Velleman (2011, 1674). This definition is more complete than the definition proposed by Lambrecht (2001).

(20) FOCUS: In a sentence, the most *communicatively significant* expression, which significance is based on a number of different factors, those having to do with the *pre-*

---

[5]Beaver *et al.* (2011) define at-issueness as: A proposition p is at-issue relative to a question Q if and only if ?p is relevant to Q.

dictability of the expression, and those having to do with its *importance* in meeting

the speaker's goals.

Under this definition, a focus will be the part of the sentence that's the most significant for the goal of communication between speaker and addressee, either because that part of the sentence is the least predictable in the mind of the addressee or because it plays a special role in meeting the speaker's goals. Its importance may derive from the fact that it requires the most processing, and thus must be drawn to the addresse's attention, or from the fact that it constitutes an unexpected piece of information.

As introduced at the beginning of this section, the literature is often fuzzy about what is referred to when talking about focus. One notion related to *focus* is that of *focus realization* or *focus marking*. This notion refers to the strategies used by different languages to make sure the focus element is marked in a special way that sets it apart form the rest of the sentence. More specifically, the literature often used the term focus to refer to prosodic prominence; a strategy used to mark focus. However, while focus is a universal phenomena, prosodic marking of focus, which is pervasive in languages like English, is not universal. Yet defining focus in terms of prosodic marking is also problematic because the domain of the F-mark is not always clear-cut. Cowles (2003) points out that, while a single word or part of a word carries the main prosodic stress of a sentence, the F-marked constituent can be much larger than that word, so long as it contains it. For example, this makes focus interpretation ambiguous in cases where the constituent carrying the pitch-accent is more deeply embedded in the surface form (21), whereas cases where the pitch-accent falls on the grammatical subject are unambiguous (22). The sentence in (21) where the pitch-accent falls on the most deeply embedded NP 'Paris' can be

an appropriate answer to the different questions listed in (21a-21d). This is just not the case in (22), which can only answer the question in (22a).

(21) Jean a   acheté un livre sur   PARIS.
John has bought a   book about PARIS.
'John bought a book about PARIS.'

    a. What did John buy a book about ?

    b. What did John buy ?

    c. What did John do?

    d. What happened?

(22) JEAN a   acheté un livre sur   Paris.
JOHN has bought a   book about Paris.
'JOHN bought a book about Paris.'

    a. Who bought a book (about Paris) ?

    b. * What did John buy ?

    c. * What did John do?

    d. * What happened?

Thus, in a sentence, a focus can be a single argument as in (14b), but also the content of a predicate or an entire sentence where there is no presupposition, as in (23) and (24) respectively. In the literature, this led many scholars to distinguish between what is often called *narrow* focus vs. *broad* focus. Lambrecht (1994) defines three different kinds of focus structure that a sentence can have based on the focus domain: (1) an argument-focus where the focus is limited to a single DP, (2) a predicate-focus where the focus falls

23

onto the V, and (3) a sentence-focus where there is no topic at all and the focus covers the entire proposition.

(23) a. What did you do last night?

    b. Je suis allé   à   une soirée dansante.
       I   am   gone to a   party   dancing.
       'I went to a dance party.'

(24) a. What happened?

    b. Il y   avait   une soirée dansante hier   soir.
       It there have.imparfait.3sg a   party   dancing   yesterday night.
       'There was a dance party last night.'

Cross-linguistically, many studies have documented how various languages differ with respect to focus marking. Of the many languages studied, French is often cited and has been of great interest to linguists because it uses syntax to a greater extent than for example English. What English expresses via prosody, French seems to express via syntax. French has several constructions at its disposal to mark focus, most of which involve clefting, dislocation or other types of non-canonical word order. I will come back to discussing Lambrecht's focus structures and French focus marking in detail in Chapter 4.

The literature on focus can also be confusing because it doesn't always differentiate between focus as a information-structural category and *focus interpretation*. Indeed, many versions of sub-classification of focus have been proposed in the literature such as *contrastive* focus, *exhaustive* focus, *informational* focus, *quantified* focus, etc. A most common one is to classify focus into two subtypes according to whether the focused item is in contrast with other alternatives in a limited set. This is exemplified in Kiss (1998),

which distinguishes between information focus and identification focus, and in Roberts (1998), which differentiates informational focus from operational focus. To look at Kiss's account, she argues that it is necessary to make such a distinction because identification focus differs from information focus in that it is associated with two unique properties Exhaustiveness and Exclusiveness. Moreover, identification focus is systematically realized differently from information focus; it is generally agreed that English cleft sentences 'it is/was … that/who …' is a construction for identification focus (Rochemont, 1986). I will consider the exhaustivity interpretation and clefts in great detail in Chapter 3.

2.1.3 Contrastiveness

The last notion relevant for this work is the notion of *contrastiveness*. While both topic and focus can be contrastive, this dissertation will only look at contrastive focus, as it is more relevant to the data in the corpus study in Chapter 4 and the production experiment in Chapter 5. We must note that, in the pragmatic theory of IS offered by Lambrecht (1994), contrast is not considered to be part of IS because it lacks distinct formal encoding in the grammar. Under Lambrecht's view, contrast differs from topic and focus because the latter two have direct grammatical form correlates in the grammar, but the former doesn't: Semantically identical sentences may have a contrastive or a non-contrastive interpretation, depending on the context. Consider the example in (25):

(25) (Let's assume John used to date Melissa but after an awful separation, they cannot be in the same room without starting a fight).
A: Who came to the party, John or Melissa?
B: MELISSA came to the party.

We saw that the old referent 'Melissa' constitutes the focus of the sentence by virtue of being the relevant answer to the previous QUD.[6] This type of focus is often referred to as *contrastive* because the focal constituent is part of a set of alternatives that was explicitly stated in the QUD.

Halliday (1967) defines contrastiveness as referring to something 'contrary to some predicted or stated alternative'. However, while it is true that a contrastive sentence often carries a corrective reading, it is not required. Chafe (1976) proposes a more general definition based on three factors: (i) shared background knowledge, (ii) a set of possible candidates for the contrasted element, and (iii) the assertion that one of the possible candidates is the correct one, excluding all others. Applied to our example in (25), the participants in the conversation are aware of the open proposition that someone came to the party; John or Melissa are listed as possible candidates; and Melissa is then asserted as the actual guest and John is excluded as a possibility. Under Lambrecht's view, B's response in (25) only differs from the other possible answer 'JOHN came to the party' inasmuch as the context allows speakers to make certain inferences, whereas Chafe argues that these two contrastive sentences activate different sets of referents in the speakers' minds, and therefore reflect different processes.

Kiss (1998) differs from Lambrecht's account by arguing that contrastiveness in fact has a specific encoding in grammar: the English *it*-cleft is used to mark contrastive/ exhaustive focus. She claims that contrastiveness is a subclass of what she refers to as *identificational* focus, defined in (26):

---

[6]The QUD is a term coined by Roberts (1996) to denote the question that has been accepted to become the immediate topic of discussion by the participants of a conversation

(26) An identificational focus represents a subset of the set of contextually or situation-ally given elements for which the predicate phrase can potentially hold; it is iden-tified as the exhaustive subset of this set for which the predicate phrase actually holds.

This claim is particularly relevant to this dissertation since I will provide evidence against it, showing that, in French, there is no strict one-to-one relationship between a syntac-tic encoding of focus and its interpretation. In his paper, Zimmerman (2008) similarly argues that contrast should not be analyzed in familiar semantic terms as involving the introduction and subsequent exclusion of alternatives. Rather, he proposes that an ade-quate analysis of contrast must take into account discourse semantic notions like *hearer expectation* or *discourse expectability* of the focused expression in a given discourse situ-ation.

2.2 Prosodic Phonology

The goal of this section is to present background information on the basic prosodic characteristics that apply to French and that are relevant for this work.

Prosodic phonology deals with how the mental representation of speech is divided into hierarchically organized phonological units (Nespor and Vogel, 1986). While many variants of the prosodic hierarchy have been proposed, the major assumed units com-posing it are illustrated below:

(27) **Prosodic Hierarchy** (Selkirk, 1978):

Utterance (U)

Intonational Phrase (IP)

Phonological Phrase (PHP or ϕ)

Word (ω)

Foot (Σ)

Syllable (σ)

2.2.1 Accentual Phrases

In the majority of studies on French prosody, the prosodic hierarchy includes at least two levels of constituency: the Accentual Phrase (AP)[7] - corresponding to the PHP in (27) - and the Intonational Phrase (IP). At the word level, French differs from other Romance languages and English in that it has no *lexical* stress. In other words, stress does not have a distinctive function on individual words. To give a concrete example, while English distinguishes the verb "to reCORD" from the noun "a REcord", French does not have this distinctive capacity. Instead, stress has a *demarcative* function; it indicates the end of a phrase. The location of stressed syllables occurs at the phrase level as *prosodic* stress. Jun and Fougeron (2002) argue that prosodic stress is typically placed on the final full syllable of a word (realized with longer duration and higher intensity than non-final syllables) only if it is the last full syllable of an (accentual) *phrase*. In this case, the phrase-final syllable is said to carry *primary* stress. To illustrate, consider the example in (28) taken

---

[7]In the French literature, this level of the prosodic structure is also referred to as 'mot prosodique' (Martin, 1987), 'mot rythmique', 'unité rythmique' (Hirst and di Cristo, 1992), 'syntagme phonologique' (Delais-Roussarie, 1996) or finally 'phonological phrase' (PHP) by (Post, 1993).

from Delais-Roussarie and Rialland (2007), in which (28b) shows the possible syllables that can receive primary stress in bold, and (28c) shows the primary accents occurring at the AP level. Moving up one phrase level, the sentence in (28d) shows where the primary stress falls within an IP.

(28)   a. Le jeune frère de François est venu ce matin.

    b. Le **jeune frère** de Fran**çois** est ve**nu** ce ma**tin**.

    c. (Le jeune **frère**)$_{AP}$ (de Fran**çois**)$_{AP}$ (est ve**nu**)$_{AP}$ (ce ma**tin**)$_{AP}$.

    d. (Le jeune frère de François est venu ce ma**tin**)$_{IP}$.

Regarding realization, Delais-Roussarie (1996) and Jun and Fougeron (2002), argue that the phrase-final syllable carrying primary stress is often realized via a rising pitch movement. Moreover, Delais-Roussarie (1996) proposes that a low peripheral tone is associated with the first syllable of the AP; the tonal representations associated with an AP are either L LH* or LH*. Jun and Fougeron (2002) offer a variant proposal by arguing that an AP is demarcated by an initial rising and a final rising tone, giving the underlying tonal pattern LHiLH*, where the final H* aligns with the stressed syllable of an AP. Both studies consider the final rise (H*) to be a "pitch accent" because of its association with the prosodically prominent syllable.

In addition to the primary stress, an AP in French has also been claimed to have an optional secondary stress (or initial stress). The majority of the relevant studies claim that the initial stress occurs on the first or the second syllable of a word (Rossi, 1985), suggesting that the (optional) secondary stress of AP must occur leftmost.

The next issue to examine are the criteria for an AP's well-formedness. Knowing what a well-formed AP "looks like" will inform us on how to correctly parse a sentence into APs. Many studies have proposed that the prosodic structure would be primarily influenced by the syntactic structure of the sentence. Within "*end-based mapping*" approaches (Selkirk, 1986; Selkirk and Shen, 1996), the algorithms responsible for the syntactic-prosodic mapping are defined in terms of the *ends* of syntactic constituents, imposing a left or right-edge phonological boundary to occur at the left or right-edge of a certain syntactic constituent. These ideas are further refined within the Optimality Theory framework and transformed into *alignment* constraints (Selkirk, 1995; Truckenbrodt, 1999; Selkirk, 2000):

(29) ALIGN-XP, R: Align (XP,R ; P,R)

"The right edge of any XP in syntactic structure must be aligned with the right edge of a phonological phrase in prosodic structure."

(30) ALIGN-XP, L: Align (XP,L ; P,L)

"The left edge of any XP in syntactic structure must be aligned with the left edge of a phonological phrase in prosodic structure."

Yet, for French, Delais-Roussarie (1996), Jun and Fougeron (2000, 2002) and Féry (2001) notice that there can be a great deal of variation in the way speakers actually pair up elements into an AP, and some metrical configurations can be strongly disfavored even though they respect the rightmost stress assignment constraint (Delais-Roussarie and Rialland, 2007). Consider the example in (31).

(31) Le président serbe.

a. *(Le prési**dent**) (**serbe**).

b. (Le **pré**sident) (**serbe**).

Despite exhibiting primary stress rightmost, the realization in (31a) is unlikely to be produced because it has two primary stress occurring next to each other. (31b), where the first syllable of " président" receives secondary stress. This example amongst many others led scholars to argue that the stress pattern assigned to an AP in French results from distinct types of constraints, notably constraints on the phonological size of the phrase.

In that sense, French differs from English, Spanish and Italian, since an AP will not always directly correspond to a single XP. The constraints posited by Truckenbrodt (1999) for English WRAP-XP, whereby each XP must be contained within its own AP will not apply to French since not every constituent will form its own phrase. Following Delais-Roussarie (1996, 19), determining what counts as a AP is governed by the alignment constraint ALIGN (P, R; HEAD(P), R) derived from (29), but also by the three other constraints in (32):

(32) **Phonological constraints on French AP**:[8]

a. ALIGN (P, R; HEAD(P), R): Align the right boundary of every accentual phrase with its head.

b. MIN$_{ap}$: An accentual phrase must be composed of at least three syllables.

c. MAX$_{ap}$: An accentual phrase must be composed of at most six syllables.

d. *STRUC$_{ap}$: Avoid accentual phrases.

---

[8]Jun and Fougeron (2000) find similar numbers on the length, arguing that an AP tends to contain an average of 2.3-2.6 words (or 1.2 content words), and 3.5-3.9 syllables.

These constraints illustrate how (33a) is preferred over (33b) because the latter violates $\text{M{\scriptsize IN}}_{ap}$ and *$\text{S{\scriptsize TRUC}}_{ap}$.

(33) a. (Les petits enf**ants**) (font un de**ssin**)

b. (Les petits enf**ants**) (**font**) (un de**ssin**)

2.2.2 Intonational Phrases

The next level of prosodic constituency is the Intonational Phrase (IP). While the AP level is the domain of pitch accents (T*), the IP level is the domain of boundary tones (T%). In the literature, the rules for forming an IP are much less defined than the ones for an AP, but the majority of studies take an IP to be the domain of a single prosodic contour. There exist two main proposals; scholars such as Edmonds (1976) and Nespor and Vogel (1986) who argue that an IP is derived from syntactic information, and others such as Selkirk (1984) who argue that it is derived from semantic information. Amongst the syntactic approaches, studies generally argue that an IP corresponds to a *root* sentence, a sentence which is not dominated by another node other than an S.

One issue with these approaches is pointed out by Selkirk (2005) and concerns the treatment of parentheticals, appositives and non-restrictive relative clauses. Consider the example in (34):

(34) Les Romains, qui sont   arrivés avant 100 ans   avant J-C, trouvèrent une
      The Romans,   who be.3pl arrived before 100 years before J-C, found   a
      terre de collines boisées.
      land of hills   wooded.
      'The Romans, who arrived before 100 AD, found a land of wooded hills.

While Edmonds (1976) does not treat the expression between comma - '*who arrived before 100 AD* - as a root sentence but as dominated by the sentence containing them, McCawley (1989) argues that it is found completely "outside" of the sentence. Potts (2002, 2003) proposes that this kind of expression is in fact an adjoined phrase carrying a [+comma] feature. Selkirk (2005) follows Potts' proposal by arguing that an expression with a [+comma] feature forms a Comma Phrase, which is constrained by the alignment constraint in (35):

(35) ALIGN R (COMMAP, IP): Align the R edge of a constituent of type Comma Phrase in syntactic representation with the R edge of a corresponding constituent of type IP in phonological representation.

This constraint will require the prosodic representation in (36):

(36) (The Romans)$_{IP}$ (who arrived before 100 AD)$_{IP}$ (found a land of wooded hills)$_{IP}$

Truckenbrodt (2005) offers a slightly different account by extending the constraints he proposed for AP to the IP level:

(37) ALIGN-CP = ALIGN(CP, R; I, R): The right edge of a CP must coincide with the right edge of an intonation phrase.

(38) WRAP-CP: Each CP is contained in a single intonation phrase.

Both constraints are in conflict in examples of coordinated embedded clauses, where two propositions are expressed within a single IP, illustrated in (39).

(39) a. [Billy thought his father was a merchant and his mother was a secret agent]$_I$

b. *[Billy thought his father was a merchant]$_I$ [and his mother was a secret agent]$_I$

In English, the realization of (39a) over (39b) is correctly predicted by having the alignment constraint ALIGN-CP ranked lower than the wrapping constraint WRAP-CP: This latter constraint limits the formation of IP and ensures that the sentence is derived into a single large IP. However, Truckenbrodt (2005) argues that, in German, the opposite ranking is observed.

Concerning French, the account offered by Selkirk (2005) is interesting because it makes it possible to treat non-canonical word-orders such as dislocations as independent IPs. Some scholars such as Clech-darbon *et al.* (1999) and Hamlaoui (2007) further apply this analysis to clefts, where the relative-like clause is taken to carry a [+comma] and form its own IP. These studies and their motivation for positing two IP in a cleft are discussed in section 2.3.1.

2.2.3 Summary

Following the discussion above, the generalizations for French phonology are described below.

(40) **French Stress Rule**:

a. Assign primary stress to the rightmost phonological word in the AP.

b. Assign main stress to the rightmost AP in the IP.

c. Assign main stress to the IP.

(41) **Syntax-Prosody Mapping Rule**:[9]

   a. Align the right edge of a syntactic phrase with the right edge of a AP phrase.

   b. Align all the left-edges of the largest extended projection of the verb with the left-edge of an intonation phrase.

   c. Align all the left-edges of the intonation phrase with the left-edges of the largest extended projection of the verb.

   d. Align all the right-edges of the largest extended projection of the verb with the right-edge of an intonation phrase.

   e. Align all the right-edges of the intonation phrase with the right edges of the largest extended projection of the verb.

2.3 The *C'est*-Cleft

The term 'cleft-sentence' generally refers to the sentences like (42) for French. These sentences are characterized by a bi-clausal structure divided into a matrix-clause and a relative-like clause or cleft-clause.[10] The matrix clause contains a copula and a clefted constituent, the focus, and the relative-like clause generally contains the pragmatic presupposition. Depending on the context, the relative clause can simply be omitted, yielding what is known as a 'reduced' cleft-sentence (i.e. '*C'est Paul*' for the first sentence, '*C'est en forgeant*' for the second, etc).

---

[9]Adapted from Hamlaoui (2007, 3).

[10]Note that the second clause of a cleft sentence is often taken as not behaving like a relative clause since many authors take a relative clause to be primarily restrictive (Comrie, 1981; Keenan, 1985). However, Lambrecht (2001) argues against such a distinction and refers to the second clause as a relative clause, arguing that the fundamental property of all relative clauses is that they are *predicates*.

(42) a. C' est Paul qui boit   du   vin.
             It is   Paul who drinks some wine.

             'It's Paul who is drinking wine'.

     b. C' est en forgeant qu' on   devient   forgeron.
             It is   by forging   that one becomes blacksmith.

             'It's by forging that one becomes a blacksmith = Practice makes perfect.'

     c. C 'est là-bas   que je me   sens le mieux.
             It is   over-there that I myself feel the best.

             'It's over there that I feel the best.'

2.3.1   Syntactic properties

In the literature on French, the syntactic structure of the *c'est*-cleft is still debated. Two major kinds of accounts exist. On the one hand, there are accounts following the definition proposed by Lambrecht (2001), arguing that a cleft is a complex sentence structure consisting of a matrix clause and a relative clause collectively expressing a single proposition, and on the other hand there are accounts arguing in favor of two structurally independent propositions such as Clech-darbon *et al.* (1999), claiming that cleft constructions do not exist as such. I review both approaches below.

In the traditional approaches, or *cartographic* approaches, the cleft is associated with an extraction operation by which the clefted constituent moves into a dedicated VP-peripheral focus position where it can receive main stress and be interpreted as the focus of the sentence (Katz, 1997; Lambrecht, 2001; Belletti, 2005). In their *Grammaire Méthodique du français*, Riegel *et al.* (2006) describe a cleft as a constituent extracted from the sentence and placed at the beginning of that sentence, surrounded by *c'est* and the relative pronoun *qui* or *que*. Belletti (2005) compares the French *c'est*-cleft to Italian

subject inversion, arguing that both strategies have the same informational content and serve the same purpose. Thus, the author proposes a unified account for the analysis of these strategies positing the structure in (43) for French, and a similar structure in (44) for Italian:

(43) [TP C'est [TOP [FOC Paul [TOP [VP <être> [SC <Paul> [CP qui boit du vin]]]]]]]

(44) [TP *pro* … ha parlato … [TOP [FOC Gianni [TOP [VP …]]]]]

In French, the relevant preverbal subject position is filled by the pronoun "ce/c'" , while the focus element is originally merged as the subject of the small clause complement of the copula, and then moved into Focus position. Clefting is selected in French because the language disallows null (preverbal) subjects. Under Belleti's account, answering with a cleft is syntactically motivated because it properly activates the VP periphery in a way compatible with the non-null subject nature of the language. Along the same lines, Lambrecht (1994, 25) argues for a functional motivation of cleft-sentences where they are a way for the language to "have its cake and eat it too". Clefting results in the placement of syntactic constituents and prosodic accents in cognitively preferred positions from which the grammar of the language normally bans them, without causing ungrammaticality.

Taking a look at other languages, similar accounts have been proposed, notably for the English *it*-cleft. The so-called "expletive" analyses of the *it*-cleft take it to be semantically parallel to focus-fronting. The clefted element is argued to occupy a focus-related position in the left periphery of the cleft clause (Halvorsen, 1989; Kayne, 1994; Kiss, 1998; Cottell, 2002). One major commonality of the accounts discussed so far, both for the French and the English clefts, is that the pronoun "ce/c'/it" is a dummy (it is se-

mantically inert).

Lambrecht (2001, 466) refines its account by proposing a constructional account of cleft-sentences whereby clefts are *grammatical constructions* in the sense of construction grammar, that is, as form-function pairings whose structural and semantic properties cannot, or not entirely, be accounted for in terms of other properties of the grammar of a language or of universal grammar and which therefore require independent explanation. In all of his work on clefts, Lambrecht works with the following definition of a cleft construction:

(45)   Cleft Construction: a complex sentence structure consisting of a matrix clause headed by a copula and a relative or relative-like clause whose relativized argument is coindexed with the predicative argument of the copula. Taken together, the matrix and the relative express a logically simple proposition, which can also be expressed in the form of a single clause without a change in truth conditions.

In direct opposition with Lambrecht (2001), Clech-darbon *et al.* (1999) propose that *c'est*-clefts are a simple "amalgamation of independently occurring types of identificational sentences and relative clauses". In other words, there is nothing in their semantic and syntactic properties that "cannot, or not entirely, be accounted for in terms of other properties of the grammar of a language or universal grammar and which therefore require independent explanation". Under their "maximally simple" account, and *contra* Lambrecht, a cleft sentence comprises two propositions in which the second proposition is base-generated as right-adjoined to an ordinary identificational phrase, thus yielding the structure in (46):

(46) [IP [IP C'est$_i$ [VP t$_i$ [DP Paul]]] [CP Op$_j$ [C' qui [IP t$_j$ boit [DP du vin]]]]]

Thus, in this account neither the focused constituent nor the main stress moves to a position from where it is normally banned by the grammar. In fact, it is the exact opposite: the focused constituent is directly merged in the position where grammar assigns main stress.

According to Clech-darbon *et al.* (1999) and Hamlaoui (2007), the main advantage of this base-generated account is that it makes correct predictions about the prosodic structure of the sentence, especially the prosodic properties of the clefted element. Under Belletti's or Lambrecht's account, it is difficult to see how the mapping between syntax and phonology could account for main stress occurring on the clefted element, since that element does not occur rightmost. Yet, in French, main stress is normally assigned to the rightmost constituent in an Intonational Phrase (IP) (Delais-Roussarie, 1995). Figure 2.1 taken from Hamlaoui (2007) shows that the left boundary of the embedded intonational phrases seems invisible to the rule assigning main stress, and the last constituent of the restrictive relative clause is the constituent that gets assigned the main stress. However, by treating the relative clause as right adjoined to the identificational TP, and following the syntax-prosodic mapping rule stated in (41), Hamlaoui proposes the prosodic representation in Figure 2.2, where each TP aligns with an IP.

Hamlaoui's account is attractive because it gives a new perspective on why clefting appears to be mandatory for realizing French lexical subjects. Unlike Lambrecht's account which lacks explanatory power about why French would ban prosodic marking sentence initially, Hamlaoui's analysis motivates the emergence of the cleft based on prosodic factors. In this dissertation, I will follow her analysis despite a hurdle: on the

IntPs
φs   IntPs
                                  φw   φs
ωs   ωs   ωs

[TP c'est [TOP [FOC [DP Ella] [VP [SC [CP qui a mangé [DP **un biscuit**]]]]]]]]
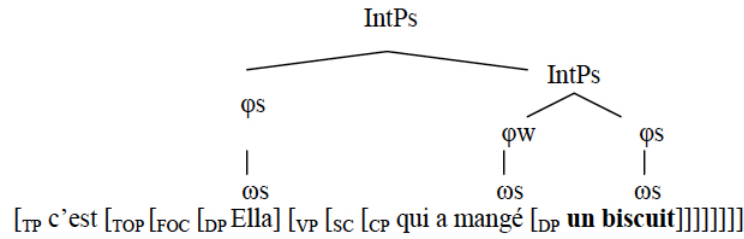
Figure 2.1: *Syntax-Phonology mapping rules for Belletti's account of clefts.*

IntPs
IntPs   IntPw
φs   φw   φs
ωw   ωs   ωs   ωs

[TP [TP C'est [VP [DPElla]]]   [CP qui a mangé [DP un biscuit]]]
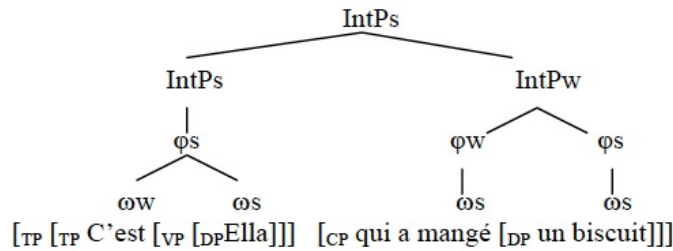
Figure 2.2: *Syntax-Phonology mapping rules for Hamlaoui's account of clefts.*

semantic level, an IP generally corresponds to a proposition, yet it seems hard to treat the cleft as two distinct propositions.

2.3.2   Different types of *c'est*-clefts

Crosslinguistically, Prince (1978) is the first study to argue that there exists more than one type of *it*-cleft. She proposes two distinct types of *it*-clefts for English, which occur under different discourse conditions. She describes the first type - "Stressed Focus" - as a cleft where the focus represents new information and the relative clause represents information which is often, though not always, known from the context. In the other type - "Informative Presupposition" - the focus typically contains an anaphoric item, and the relative clause contains the "message", but marked as a known fact, not as the speaker's

assessment. Each cleft type is illustrated in (47) and (48) below. The two types differ from each other in the informational status of what is found in the relative clause: the relative clause in the stressed-focus type contains presupposed material (i.e. that the addressee is expected to know), whereas the material is a known fact, but not known by the addressee at the time of utterance in the informative presupposition cleft.

(47) *Stressed-Focus:*

So I learned to sew books. They are really good books. It's just the covers that are rotten.

(48) *Informative Presupposition:*

It was just about 50 years ago that Henry Ford gave us the weekend.

Prince' study has sparked much interest in the study of clefts, especially concerning their use in discourse. As a result, many studies on French concentrated on categorizing *c'est*-clefts in various types, showing how they differ from each other in informational and prosodic terms (Lambrecht, 1994; Katz, 1997; Clech-darbon *et al.*, 1999; Rialland *et al.*, 2002; de Cat, 2002; Doetjes *et al.*, 2004).

2.3.2.1   Type 1: Narrow/Identificational cleft

In most studies, a first type of cleft commonly emerges in which the pivot is the only element that receives prosodic prominence (no element in the relative clause is prosodically prominent). The informational structure of this cleft is such that the information in the relative clause is presupposed and the information in pivot position is unknown to the addressee. This type of cleft is similar to the "Stressed-Focus" cleft in Prince, and is

referred to as a "Focus-Ground" partition or a "narrow" focus cleft in the French literature (Beyssade *et al.*, 2001; Doetjes *et al.*, 2004). Example (49) illustrates this type of cleft:

(49) C'est LE PETIT qui est tombé dans l' escalier.
      It-is THE LITTLE ONE who is   fallen in the   stairs.
      'THE LITTLE ONE fell down the stairs'.

Prosodically, studies differ regarding how the element in pivot position is marked. Some claim it is marked via a pitch accent (Katz, 1997), others via emphatic stress (de Cat, 2002) and others via an L% boundary tone present at the right-edge of the focus (Rialland *et al.*, 2002; Doetjes *et al.*, 2004).[11] Functionally however, there is a consensus that this cleft specifies the (unique) value for the variable instantiated by the *qu*-word in the preceding question or discourse. In example (49), "le petit" is being specified as the person who fell down the stairs. Most studies also consider this type of cleft to be associated with an exhaustive reading which implies that no other candidate than X has the property described by the predicate in the relative clause y. The characteristics of this cleft are summarized in Table (2.3).

---

[11]I will detail the prosodic realization within a *c'est*-cleft in Chapter 5.

	C'est X…	*qui/que Y.*
Informational-status of X/Y	focus	presupposed
Mental state of X/Y	unknown/unpredictable by addressee	activated in discourse/retrievable by addressee
X/Y in Discourse	brand-new or discourse referent already present in the preceding discourse	strictly given in preceding discourse or situationally/contextually evoked
X/Y Prosody	prominent	non prominent
X/Y Pragmatic	exhaustive reading	common ground knowledge
Note	When the X is complex, some part can be presupposed and prosodically unmarked	

Table 2.3: Characteristics of a Type 1 cleft ("Focus-Ground" or "Narrow-focus" cleft)

2.3.2.2 Type 2: Contrastive/Corrective cleft

The second kind of cleft discussed in the French literature is often called "Contrastive" (Katz, 2000). This type of cleft has the same informational status as Type 1 clefts described above: the pivot element (or some part of it), as opposed to the information in the relative clause, is not presupposed. Functionally, this cleft is claimed to be used in contexts where the focused element negates the value assigned to a variable, and introduces an alternative value for such a variable (de Cat, 2002, 173). This cleft can be contrastive in the sense that it offers an explicit contrast between a number of entities, but it can also be corrective if it conveys a correction to a faulty assumption from the preceding discourse. Therefore, as Zimmerman (2008) explains, "contrastive marking does not so much indicate the explicit or implicit presence of contrasting alternatives in the linguistic context, although this may be a side effect, but rather a contrast between the information conveyed by the speaker in asserting α and the assumed expectation state of

the hearer". This type of cleft is illustrated in (50) and (51) below.[12] In (50), the speaker is contrasting two entities (the French and the Spanish) and what property holds of each (drink and smoke). In (51), speaker A uses a correction because he considers that B is making a faulty assertion.

(50) C'est LES FRANÇAIS qui boivent et LES ESPAGNOLS qui fument.
It-is THE FRENCH who drink and THE SPANISH who smoke.
'The FRENCH drink and the SPANISH smoke.'

(51) A: Le patron veut te parler.
A: The boss wants you talk.INF
'A: The boss wants to talk to you'

B: Je n' ai rien à lui dire.
B: I neg have nothing to him say.

'B: I have nothing to tell him'

A: C'est LUI qui parlera.
A: It-is HIM who speak.future

'A: HE is the one who will do the talking.'

Prosodically, Katz (1997, 2000) claims that Type 1 (narrow) and Type 2 (corrective/constrative) *c'est*-clefts differ from each other on two points: Type 2 has a "stronger" pitch accent than Type 1, and a high boundary tone is found at the right-edge of the utterance in Type 1 but not Type 2. In Chapter 5 of this work, I show, based on production data, that the difference is in fact that the focused element in a Type 1 cleft is not always marked,

---

[12]Example (51) is taken from the movie *Les Compères*, a 1983 French comedy written and directed by Francis Veber.

whereas it consistently carries prosodic prominence when in contrastive/corrective context.

The characteristics of this type of cleft are summarized in Table (2.4).

	C'est X…	*qui/que Y.*
Informational-status	focus	presupposed
Mental state	unknown/unpredictable by addressee	activated in discourse/retrievable by addressee
In Discourse	brand-new or discourse referent already present in the preceding discourse	strictly given in preceding discourse or situationally/contextually evoked
Prosody	prominent	non prominent
Pragmatic	contrastive or corrective + exhaustive	common ground knowledge

Table 2.4: Characteristics of a Type 2 cleft ("Contrastive" or "Corrective" cleft)

2.3.2.3   Type 3: Broad/Event-related cleft

Finally, a third type of cleft is reported on by some studies (Clech-darbon *et al.*, 1999; Rialland *et al.*, 2002; Claire *et al.*, 2004), but is not as widely assumed as the previous two. This type of *c'est*-cleft is typically called "Broad-focus", "All-focus" or "Explicative". In this type of cleft, the focus is not restricted to the pivot position, but in fact spreads over the whole sentence. Thus, there is no presupposed material and all the information present in the *c'est*-cleft sentence is unknown/unpredictable by the addressee.[13] A couple of examples of such a cleft are given in (52) and (53).[14]

---

[13]This type of information-structural articulation is what Lambrecht (1994) refers to as "Sentence-focus", which I will discuss in Chapter 5, yet Lambrecht claims that this IS does not occur in a *c'est*-cleft but in a *have*-cleft.

[14]These examples are adapted from Clech-darbon *et al.* (1999) and Claire *et al.* (2004).

(52) C'est avec plaisir   que je vous invite à ce   séminaire.
      It-is with pleasure that I   you invite to this conference.

      'It's with pleasure that I invite you to this conference.'

(53) A: You look worried, what's going on?

      B: C'est le   petit qui vient de   tomber dans les escaliers
      It-is the   little one who comes from to-fall in   the stairs.

      'The little one just fell down the stairs.'

Functionally, this type of cleft allows a broad scope reading of the information and mainly has a "presentational" function: it reports to the listener about a state of affairs as a whole. Prosodically, this cleft is argued to have the following features: a continuation rise (H) at the right-edge of the focused element which signals the end of a rhythmic group, and a terminal boundary tone at the right-edge of the utterance (Claire *et al.*, 2004).

	C'est X…	*qui/que Y.*
Informational-status	focus	focus
Mental state	unknown/unpredictable by addressee	unknown/unpredictable by addressee
In Discourse	brand-new or discourse referent already present in the preceding discourse	brand-new or discourse referent already present in the preceding discourse
Prosody	prominent	prominent
Pragmatic	often answers the QUD "What happened?" or is uttered as an out-of-the-blue sentence.	

Table 2.5: Characteristics of a Type 3 cleft ("Broad" or "Explicative" cleft)

2.3.2.4   A few other clefts

A few studies discuss more cleft types, such as the "Causal" type and the "Factual" proposed by Katz (2000). A "Causal" cleft is similar to the "Informative-presupposition" type in Prince's work in that the information in the relative clause does not have to be presupposed, yet differs in that this information is not a generally known fact. As the name indicates, this type of cleft is argued to semantically convey the cause of what is occurring in the relative clause (Katz, 1997, 123). The element in pivot position is always a causal preposition like *"pour çà/for that", "à cause de/because of"* or *"grâce à/thank to"*. Katz (2000) argues that a speaker uses this cleft when she wants to convey that the most important part of the sentence is the relationship between the effect (found in the relative clause) and its cause (expressed in pivot position). Katz and Blyth (2007, 164) argue that this cleft corresponds to the construction "that's why" in English, and provide the example in (54).

(54)   Je ne   suis pas   très contente de vivre ici.   C'est pour çà   que je
       I   neg.not am   neg.not very happy   to live here. It-is for   that that I
       cherche un nouveau poste pour l'   année prochaine.
       look-for a   new   job   for   the next   year.
       'I'm not very happy to live here. That's why I am looking for a new job for next
       year.'

A "Factual" cleft is exactly equivalent to Prince's "Informative-presupposition" since the relative clause contains information which refers to a known fact (unknown by the addressee at the time of utterance), and is accommodated but not truly presupposed. The majority of the examples found in the literature for this cleft involve some kind of historic

event, often well-known by the public, or involving eminent personalities. Prince proposes the example in (55a) and Lambrecht the example in (55b). But this type of cleft is not restricted to generally accepted known facts. Consider the example in (56).[15]

(55) a. It was just about 50 years ago that Henry Ford gave us the weekend.

 b. It was George Orwell who said that the best books are those which tell you what you already know.

(56) C'est en 1981 que je décide de rejoindre mon père   afin qu' il m'   enseigne
It-is in 1981 that I decide to join   my father so that he to-me teaches
son savoir-faire et   ses connaissances, en un mot: son art.
his savoir-faire and his knowledge,   in a   word: his art.
'It's in 1981 that I decide to join my father so that he would teach me his savor-faire and his knowledge, in a single word: his art'.

Even though the information in the relative clause is not presupposed in (56), it seems that the cleft does more than mark that information as a fact. Indeed, here too, I argue that the cleft marks the pivot element, the date 'en 1981', as relevant in the chronology of the story. This type of cleft is more formal and generally occurs in written rather than spoken French, and therefore it could also occur for politeness and style effects. Prince (1978, 904) actually notes that this cleft can be used when self-effacement is sought by the speaker, e.g. for politeness. Lambrecht (1994, 71) comments on this type of cleft by saying that "the utterance did not become unacceptable for the lack of the required presupposition because the speaker could count on the audience's willingness to accommodate it".

---

[15]Example taken from http://www.cordonnerie-dossmann.fr/pages/infos.html

2.3.2.5 Summary and discussion

I argue that despite the tendency to want to classify *c'est*-clefts into different categories, all clefts presented here share the following characteristics:

(57) • They can be unclefted into a canonical mono clausal sentence that has the same truth-conditions.

   • They perform focalization in the sense that the pivot element is the most pragmatically important/relevant element in the sentence. No matter what the information status the element found in the relative clause is, and no matter how the focused element is marked (or unmarked), the information in pivot position is the one foregrounded for some pragmatic reasons.

   • Their basic function is to answer the (explicit) QUD by offering a maximally relevant answer. Put another way, the cleft is used by a speaker to convey the maximally relevant or important piece of information that he is willing to commit to regarding the QUD.

In the literature, the clefts presented earlier are argued to differ in three main ways:

(58) • Prosodically: some *c'est*-clefts have different prosodic pattern.

   • Pragmatically: the presupposed elements differ in the different clefts.

   • Syntactically: some *c'est*-clefts have restrictions on the type of grammatical categories that can occur in pivot position.

Table (2.6) in the Appendix summarizes the different types of clefts discussed in the literature on French with the different terms used across accounts.

In the French literature, it is common to attribute the differences between these clefts to a difference in structure (Clech-darbon *et al.*, 1999). On the contrary, I argue that the clefts described above don't need to be analyzed as different constructions based on the differences in (58). Instead, these difference can be attributed to the QUD: the difference in what element is presupposed comes from the difference in the QUD answered. In some cases, the QUD is more general leading to a broad focus, and in other cases the QUD is more specific leading to narrow focus. The prosodic differences also follow from this difference in information-structure: depending on the context and, therefore, on the QUD, the prosodic pattern will be different. Moreover, the differences between a narrow/informational cleft and a contrastive cleft should be attributed to a difference in focus (informational vs. contrastive/corrective focus) instead to a difference in construction. Even though there is a difference in the meaning conveyed, I do not see the advantage of treating the different clefts as distinct constructions. In fact, I argue the contrastive type is simply an instance of the narrow type under specific contextual conditions. In this dissertation, I offer a unified account of the *c'est*-cleft where there is no distinction amongst cleft types, but the determinant of any semantic, functional or pragmatic differences are the discourse and the participants' mental states. Therefore, the OT account that I develop in Chapter 6 explains the emergence of the *c'est*-cleft under different conditions without having to account for different constructions.

Finally, it is worth mentioning that a couple of studies make a further distinction between the broad focus clefts that are used to answer a direct/explicit question of the type "What happened?" and the ones that are not used to answer a question, which are often found in written texts or formal speeches. Rialland *et al.* (2002) argue that the dif-

ference between these two types is the kind of continuation rise found on the focused element: a minor 'h' for the broad-cleft answering an explicit question, and a major 'H' for the broad-cleft out of the blue. Again, I disagree with their treatment of these clefts. Despite a possible difference in prosody, I believe that a cleft ALWAYS answers a question, even if this question is not explicit in the context.

2.3.3   Cleft lookalikes

Another common trend in studies on the French clefts is to distinguish "real" *c'est*-clefts from cleft lookalikes. Many studies caution that, even though these structures do look like clefts, they should not be confused. The most discussed type of lookalike is the "Restrictive" type, which differs from a *c'est*-cleft because the relative clause is restrictive (Katz, 2000; Lambrecht, 2001; de Cat, 2002). Consider the following two examples:

(59) Tu es en cours avec le garçon qui parle russe?
      You are in class with the boy who speaks Russian?
      'Are you in class with the boy who speaks Russian?'

      (Non). C'est LA FILLE qui parle russe.
      (No.) It-is THE GIRL who speaks Russian.

      'No, It's THE GIRL who speaks Russian.'

(60) Tu as vu la fille là-bas? Qui c'est?
      You have.3sg seen the girl over-there? Who it-is?
      'Have you seen the girl over there? Who is she?'

      C'est la fille qui parle russe.
      It-is the girl who speaks Russian.

      'It's the girl who speaks Russian.

In the first dialogue (59) the sentence 'C'est la fille qui parle russe' is a *c'est*-cleft where the element in pivot position receives prosodic prominence and the information in the relative clause is non-restrictive (i.e. it does not limit the possible referent of the subject). The intended meaning of the cleft-sentence is to say something about the entity for which the predicate holds, and here, to also correct the faulty assertion made by the speaker. Put yet another way, the cleft sentence has the presupposition that "there exists an x who speaks Russian" and the focus "la fille". On the contrary, the same sentence uttered in the second dialogue (60) is a not a cleft but a lookalike: the relative clause is restrictive in the sense that it limits the possible referents of the subject. The important information is found in the restrictive relative clause; the intended meaning of the sentence is to say something about the property of the given pivot element. As for the pivot element itself, it is already active in the discourse and in the participants' mind. Moreover, this sentence doesn't have a simple canonical counterpart. Indeed, while (59) can also be expressed by the canonical form 'LA FILLE parle russe' without altering its truth-conditions, it is impossible to do so for (60). Instead, we would need the presence of a demonstrative such as 'Voici/*here is* la fille qui parle russe'. Prosodically, these two *c'est* constructions differ in that the second one will never bear prosodic prominence on the pivot. Another interesting characteristic of the Restrictive-lookalike is that it allows agreement of a past-participle with the auxiliary 'avoir/to have'. Moreau (1976) argues that such an agreement is not possible with cleft-sentences, as illustrated in (61a) vs. (61b).

(61) a. C'est UNE LETTRE que j' ai   écrit/*écrite.
        It-is A   LETTER that I have written.
        'It's a letter that I wrote.'

b. C'est une lettre que j'ai écrite.
It-is a   letter that I   have   written.
'That's a letter I wrote.'

Another type of evidence for distinguishing between these two *c'est* structures comes from de Cat (2002) who argues that, when X is not coindexed with the subject of the dependent clause, the pivot of a *c'est*-cleft is a PP (62a) while the pivot of a restrictive-lookalike is a NP (62b).

(62) a. C'est à la maman qu' on donne le tablier.
It-is to the mom   that one gives   the apron.
'It's to the mom that ones gives the apron.'

b. C'est la maman à qui   on donne le tablier.
It-is   the mom   to whom one gives   the apron.
'That's the mom to whom one gives the apron.'

The second type of cleft lookalike is called "Accommodated" by Lambrecht (1994) and "*pas parce que*"-structures by Katz (2000). These lookalikes are illustrated by two examples below:

(63) C'est pas parce   que je suis linguiste que je peux expliquer les clivés.
It-is   not because that I am linguist that I can explain   the clefts.
'It's not because I'm a linguist that I can explain clefts.' (From Lambrecht 2001:495)

(64) C'est pas parce   que je te dis   bonjour que je t'   aime bien.
It-is   not because that I you say.1sg hello   that I you like   well.
'It's not because I say Hi to you that I like you.' (From Facebook post)

Here again, even though these clefts look like *c'est*-clefts at first, they do not share some of the basic properties of a cleft. For example, it is impossible to uncleft them into

53

a canonical sentence with the same truth-conditional meaning. The canonical sentences in (65) and (66) derived from (63) and (64) respectively, simply do not mean the same thing:

(65) Je peux pas expliquer les clivés parce que je suis linguiste.
I can not explain the clefts because that I am linguist.
'I cannot explain the clefts because I am a linguist.'

(66) Je ne t' aime pas bien parce que je te dis bonjour.
I neg you like neg well because that I you say hello.
'I don't really like you because I say hello to you.'

(63) and (64) assert that the property X expressed in post-copular position holds of the subject of the relative clause, and imply that the property Y doesn't. Indeed, the speaker could easily continue her statement in (63) by '… *mais j'ai proposé une solution dans ma thèse/ … but "I proposed a solution in my dissertation"*' to cancel the implicature that ¬ Y. However (65) and (66) assert that ¬ Y, and cancellation is not possible. Prosodically, Lambrecht (2001, 496) argues that the focus of the sentence, which receives the primary pitch accent, is expressed in the *que* clause following the adverbial clause. A more appropriate translation for (63) and (64) are :

(67) Just because I'm a linguist doesn't mean I can explain clefts.

(68) Just because I say Hello doesn't mean I like you.

In the next chapter, I move on to discussing one meaning associated with the cleft, namely *exhaustivity*.

	Author	Terminology	Prosody	IS	Pragmatics/Function	Syntax
Type 1	Lambrecht (1994,2000)	Informational	pitch accent on focus	focus-presupposition	exhaustive/identification	constructional
	Katz (1997, 2000)	Variable Fullfilment Cleft (VFL)	"	"	"	"
	Clech-Darbon et al. (1999)	Narrow presentational	C'est HL% qui L%	"	"	bi propositional
	Rialland et al. (2002)	Focus-Ground	C'est HL% qui L%	"	"	n/a
	DeCat (2002)	Dislocated and Non-dislocated clefts	n/a	"	"	n/a
Type 2	Katz (1997, 2000)	Contrastive Cleft (CC)	'stronger' pitch accent	focus-presupposition	contrast/correction	constructional
	Clech-Darbon et al. (1999)	Narrow contrastive	C'est H* que L%	"	"	bi propositional
	DeCat (2002)	Dislocated clefts	n/a	"	"	n/a
Type 3	Clech-Darbon et al. (1999)	Broad-event related	C'est H$_{cont}$ qui L%	all focus	presentational, event-related	n/a
	Rialland et al. (2002)	Broad focus or Out-of-blue	"	"	"	n/a

Table 2.6: Summary of different types of *c'est*-cleft proposed in the French literature

Chapter 3

On the meaning of the c'est-cleft: Exhaustivity

The goal of this chapter is to examine the nature of the exhaustive inference associated with the French *c'est*-cleft. I begin by showing that, similarly to the English *it*-cleft, an exhaustive interpretation is associated with the French cleft. Then, I discuss the lively cross-linguistic debate on the nature of exhaustivity, considering the different accounts proposed to explain it: exhaustivity is an entailment, a presupposition, a conversational or a conventional implicature. In order to shed light on this issue for the French cleft, I present an experiment designed to test its interpretation compared to exclusive-sentences and canonical sentences. The results challenge an account that assumes exhaustivity is an entailment (i.e. truth-conditional). Instead, I argue that the exhaustivity associated with the cleft is a pragmatic phenomenon, and more specifically I propose to analyze it as a scalar implicature: the speaker uses a *c'est*-cleft to convey the maximal answer to the Question-Under-Discussion (QUD) to which he commits. I model this implicature in terms of a constraint on the use of the cleft which states that elements that are associated with an exhaustive interpretation must be clefted. This constraint will be later integrated into the OT analysis proposed in Chapter 6 explaining the alternation cleft/ canonical.

3.1 Introduction

The English *it*-cleft "It is x that P" is argued by many to be associated with exhaustivity effects; the element occurring in pivot position is interpreted as if under the scope of an exclusive like *only* (Horn, 1981; Kiss, 1998; Onea and Beaver, 2008). Surface similarities with *only* sentences are easy to find: both clefts and exclusive sentences require a focus somewhere in *x,* both are used when *x Ps* and no one else *Ps,* and both presuppose that something *Ps.*

(69) Only John sneezed.

(70) It's John who sneezed.

   a. ∃x P(x)

   b. x in some way exhausts the set {x | P(x)}

   c. ⊨ *John and no one else sneezed.*

In the French literature, the *c'est*-cleft "c'est *x* qui *P*" is also described as being exhaustive (Lambrecht, 1994; Clech-darbon *et al.*, 1999; de Cat, 2007). In fact, exhaustivity is argued to be one of the main discourse function of the cleft by many scholars (Katz, 1997; Clech-darbon *et al.,* 1999). Thus, to many French ears, a sentence like (71) suggests (71a) and may be equivalent to (72).

(71) C'est Jean qui a   éternué.
      It-is   John who has sneezed.
      'It's John who sneezed'.

   a. ⊨ *Personne d'autre que Jean n'a éternué.*

(72) Seul Jean a éternué.

Yet, both in English and French, there are significant differences between a cleft and a sentence with an exclusive that make a parallel harder to sustain. A first difference is illustrated by the change in in acceptability in (73), where the prejacent inference "Peter ate pizza" can be strengthened with an exclusive sentence but not a cleft.

(73) a. Pierre a   mangé de la pizza et il a   seulement mangé de la pizza.
          Pierre has eaten of the pizza and he has only   eaten of the pizza.
          'Peter ate pizza and he only ate pizza.'

      b. # Pierre a   mangé de la pizza et   c' est de la pizza qu' il a mangé
          # Pierre has eaten of the pizza and it is of the pizza that he as eaten.
          'Peter ate pizza and it was pizza he ate.'

Other differences are illustrated by the change in acceptability when the focus falls under the scope of negation in (74a) and (74b), and by the fact that the cleft allows quantified pivots (75a), whereas they are ungrammatical in an exclusive-sentence (75b).

(74) a. Pierre n'   a   pas mangé seulement/que de la   pizza, bien   qu' il
          Pierre neg has neg eaten   only   of the pizza, though that he
          ait   mangé de la   pizza.
          have-subjunctive-3sg eaten   of the pizza.
          'Peter didn't only eat pizza, though he did eat pizza.'

      b. # Ce n'   est pas de la   pizza que Pierre a   mangé, bien   qu' il
          # It   not is   not of the pizza that Pierre has eaten,   though that he
          ait   mangé de la   pizza.
          have-subjunctive-3sg eaten   of the pizza.
          '#It's not pizza Peter ate, though he did eat pizza.'

58

(75) a. Ce sont   tous les citoyens qui font l'   objet d' une discrimination.
          It is-3pl-prst all   the citizens that make the object of a   discrimination.
          'It's all the citizens that are discriminated against.'

      b. # Seuls tous les   citoyens font l'   objet d' une discrimination.
          # Only all   the-pl citizens make the object of a   discrimination.
          # 'Only all the citizens make the object of a discrimination.'

Other intriguing facts appear when comparing the cleft to its canonical counter-
part. For example, adding information onto a cleft seems questionable (76b), whereas it
is felicitous for a canonical sentence (76a). Finally, there are cases where the exhaustive
inference is trumped as illustrated in (77).

(76) a. Pierre a   embrassé Marie, et   il a   aussi embrassé Lucie.
          Pierre has kissed   Marie, and he also has   kissed   Lucie.
          'Pierre ate some pizza, and Jean also did.'

      b. ? C'est Pierre qui a   mangé de la   pizza, et   Jean aussi a   mangé de la
          ? It-is Pierre who has eaten   of the pizza, and Jean also   has eaten   of the
          pizza.
          pizza.
          '? It's Pierre who ate some pizza, and Jean also did.'

(77) a. Ce n'   est pas   Marie qui a   ri,   ce sont   Marie et
          It neg-not is   neg-not Marie who has laughed, it be-3pl-prst Marie and
          Jean.
          Jean.
          'It isn't Marie who laughed, it's Marie and Jean.'

      b. C'est en tout cas Marie qui a   lu   quelques livres de   Zola.
          It-is in any case Marie who has read some   books from Zola.
          'It is anyway Marie who read some books from Zola.'

59

The issue raised by the data in (73) though (77) concerns the nature of the exhaustive inference found associated with the French *c'est*-cleft. While it is undisputed that some exhaustive effect is present, the cleft behaves differently from exclusive sentences and canonicals.So, where does exhaustivity come from? Is it semantically encoded in the meaning of the cleft-sentence, or is it a pragmatic phenomenon arising from the use of the cleft in the language?

This issue cause much debate in the literature on focus and clefts, especially cross-linguistically for structures identified as similar to the French *c'est*-cleft like the English *it*-cleft and the Hungarian preverbal position (Horn, 1981; Kiss, 1998; Beaver and Onea, 2011). The apparent similarities between the cleft-sentences and the exclusive-sentences introduced above have led some researchers to argue for a 'semantic' account of exhaustivity, positing that these two sentence types have the same truth-conditions (Kiss, 1998; Clech-darbon *et al.*, 1999). In view of the differences, others have argued for a 'pragmatic' account, according to which the exhaustive inference is either implicated (Horn, 1981; Percus, 1997; Büring, 2010a). In recent years, a new research trend has emerged and scholars started looking at this 'old' issue with a 'new' tool, i.e. by turning to experimentally testing the claims they proposed (Onea and Beaver, 2008; Zimmermann and Drenhaus, 2009; Dufter, 2009). In the French literature, surprisingly, little is said about the nature of this exhaustive inference. To the best of my knowledge, Clech-darbon *et al.* (1999) and Doetjes *et al.* (2004) are the only two studies that propose an analysis, both arguing for a 'semantic' analysis of exhaustivity. Yet, the biggest problem for these studies is that empirical evidence suggests clefts are not invariantly exhaustive.

The goal of this chapter is to examine the nature of exhaustivity in the French *c'est*-

cleft. I start in section 3.2 by discussing the cross-linguistic literature on the semantics of clefts, reviewing the three major types of analyses proposed: entailment, presupposition and implicature. This literature review gives a cross-linguistic perspective on the issue and explains the shortcomings of some of these past analyses. In an effort to cope with these shortcomings and better reflect the behavior of the *c'est*-cleft, I present in section 3.3 an on-line forced-choice experiment designed to test the level of exhaustivity of the cleft compared to exclusive-sentences and canonical sentences by examining how speakers contradict the exhaustive inference. Results show that the *c'est*-cleft is significantly different from both sentence forms. Drawing from these results, in section **??**, I develop an alternative analysis where exhaustivity is treated as a scalar implicature. I conclude the chapter by positing a constraint on the cleft's interpretation, which demands that pivots be interpreted as full answers to the QUD.

3.2 Crosslinguistic background on Exhaustivity

Cross-linguistically, a few constructions have been studied that all seem prone to an exhaustive interpretation: the English *it*-clefts, their German and French counterparts, and preverbal focus in Hungarian (Kiss, 1998; Beaver and Onea, 2011). In the English literature, it has become almost formulaic to begin a paper on the semantics of clefts with the observations that "It is X that P" presupposes the existence of an X that Ps ($\exists x\, P(x)$) and implies that X in some way exhausts the set $\{x|P(x)\}$. But deriving an exhaustive statement such as "nobody else Ps" from the form "It is X that P" is trickier than imagined because the exhaustive inference does not pattern like standard presuppositions. In particular, exhaustivity does not project out of embedded clefts in the way existence does. If we take

61

the exhaustive inference to be represented as in (78), then this inference surely does not follow from (78a-78c).[1] Quite the contrary in fact since in (78a), the exhaustive inference "No one other than Marie is sick" is straightforwardly inconsistent with the conjunction of the existence presupposition "Someone is sick" and the assertion "Mary is not sick". And in (78b, 78c), it cannot be the case that exhaustivity is presupposed, for then the two presuppositions "Someone is sick" and "No one other than Marie is sick" would jointly entail that Mary is sick, which is precisely what is being questioned.

(78)   No one other than Marie is sick.

    a. It isn't Marie who's sick.

    b. Is it Marie who's sick?

    c. If it is Marie who's sick, we're all in trouble.

In the literature on the nature of the exhaustive inference, three main types of accounts exist: exhaustivity is either argued to be asserted (semantic account found in studies such as Kiss 1998), implicated or presupposed (pragmatic account found in studies such as Zimmermann and Drenhaus 2009). Let's review each type of account.

3.2.1   Exhaustivity is an entailment

Kiss (1998) claims that two distinct types of focus must be distinguished that have different semantic and syntactic properties: *identificational focus*, which involves the

---

[1]The literature on exhaustivity has posited different ways to represent the exhaustive inference. Here I choose to follow the representation assumed by Beaver, among others.

movement of an expression from its initial argument position to a dedicated focus position, and *information focus*, which remains in situ. Syntactically, she proposes that identificational focus has the structure in (79) where the dedicated focus position is situated in the left periphery of VP and the relative clause is a complement of the projection hosting the focus constituent.

(79)

```
                              IP
                         ╱         ╲
                      DP              I′
                      │            ╱      ╲
                      it        I+F_i       CP
                                │        ╱     ╲
                               was     DP        F′
                                       │      ╱     ╲
                                     John   t_i      CP
                                                   ╱    ╲
                                                that Mary kissed
```
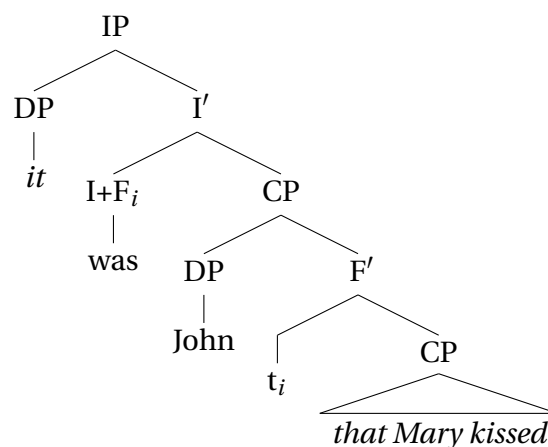
Semantically, the pronoun *it* is an expletive element inserted purely to satisfy the requirement for a structural subject in SpecIP (the EPP constraint). Moreover, Kiss argues that expressions of identificational focus must be interpreted as if under the scope of an exclusive particle like *only*. Under this view, English *it*-clefts and Hungarian preverbal focus position are considered to be instances of identificational focus: they are semantically exhaustive (i.e. the exhaustivity is part of the truth-conditions of the sentence) and they syntactically occupy a dedicated focus position. In contrast, expressions of informational focus do not require such an interpretation and are simply used to convey the novelty of the information in focus. English pitch-accent is an instance of informational focus. As a result, researchers have attributed any potentially exhaustive readings arising from such

prosodically focused expressions to pragmatic inferences rather than semantics.

Semantic accounts of the nature of exhaustivity are found crosslinguistically in studies like Szabolcsi (1981) for Hungarian preverbal focus, Atlas and Levinson (1981) for the English it-clefts, and Clech-darbon *et al.* (1999) for the French *c'est*-cleft. Atlas and Levinson (1981), for example, analyze exhaustivity in terms of an entailment, whereby a sentence of the form "It is X that P" amounts to something close to "X Ps and only X Ps". Applied to the example in (80), their analysis attempts to account for the intuitions that (80) entails but does not presuppose (80a, 80b, and 80c); whereas (81) presupposes (80b), entails the negation of (80a) and does not imply (80c) since entailments do not survive negation.

(80)   It was John that Mary kissed.

    a. Mary kissed John.

    b. Mary kissed someone.

    c. Mary kissed (exactly) one person.

(81)   It wasn't John that Mary kissed.

Under this analysis, exhaustivity is modeled as part of the truth-conditional meaning of the cleft by giving the cleft "It is X that P" the logical form in (82) and paraphrased as *The group of individuals that satisfy P is identical to X.*

(82) $\lambda x.(x = NP)(\gamma x P(x))$

In her 1998 paper, Kiss attributes the exhaustivity of the cleft to an operator of exhaustive identification present in the cleft. Some of the evidence she puts forward to sup-

port her claim contains examples of an apparent distributional restriction on the pivot. She argues that some expressions like universal quantifiers (*everybody, all*) and additives (*even, also*) are banned from appearing in pivot position because their truth-conditional meaning clashes with the truth-conditional exhaustivity of the cleft. Thus, as it is ungrammatical to combine (83).

(83)   * It was everybody/ all/ even/ also John that Mary invited to her birthday party.

However many scholars have given compelling evidence against a semantic account of exhaustivity (Horn, 1981; Vallduví, 1992; Delin and Oberlander, 1995; Wedgwood, 2005). Horn (1981) argues that this type of account fails to explain the inappropriateness of sentences in (84). Indeed, the continuations in (84a-84d) are infelicitous precisely because the cleft does not assert anything about exhaustiveness and is therefore inconsistent. To make these continuation felicitous, all is needed is the explicit indication of exhaustivity: each continuation becomes felicitous if an exclusive particle like *only* is added. If the cleft were indeed semantically exhaustive, this addition would not be necessary to produce felicity.

(84)   I know Mary ate a pizza ...

    a.   # but it wasn't a pizza that she ate!

    b.   # but was it a pizza that she ate?

    c.   # but I've just discovered that it was a pizza that she ate!

    d.   # if it was a pizza she ate, then all is well.

(85)   a.   but it wasn't only a pizza that she ate!

b. but was it only a pizza that she ate?

c. but I've just discovered that it was only a pizza that she ate!

d. if it was only a pizza she ate, then all is well.

In order to distinguish between truth conditional content and presupposition, von Fintel (2004) proposes a test that uses the locution *Wait a minute*. In this test, the speaker utters a sentence and the addressee follows up with a comment introduced by *wait a minute*, which is contradicting part of the speaker's meaning. This test can be applied to the context of clefts to see if the exhaustivity is *suspendable*, a property that presuppositions display. Consider the example in (86):

(86) a. Speaker: It's Paul who kissed Mary.

b. Addressee: Wait a minute! It's Paul and John who kissed Mary.

If the meaning expressed by the speaker can be correctly suspended by the follow-up "wait a minute"-sentence, then that meaning is taken to be a presupposition.

Finally, Dufter (2009) presents empirical evidence against Kiss' idea that certain expressions are banned from pivot position. According to Dufter, if exhaustivity were in fact semantically encoded, it would be hard to explain the following: (i) exclusives like "only/ne...que, uniquement, seulement" are found modifying clefted constituents without being redundant (87); (ii) additives like "also/même" occur in a cleft (88a-88b), whereas a paraphrase with "only" would not be felicitous (88c-88d); and (iii) universal quantifiers are allowed without clashing with the restriction otherwise impeded by an exclusive (89). These distributional patterns are studied in more details for the French cleft in corpus studies in Chapter 4.

(87) a. It is only on that basis that a balanced approach can be found.

    b. C'est uniquement sur cette base qu'une approche équilibrée peut-être envisagée.

(88) a. Moreover, it is also out of respect for your electorate that, as President-in-Office of the Council, I do not wish to set myself up as judge.

    b. D'ailleurs, c'est également par respect pour vos électeurs qu'en tant que président en exercice du Conseil, je ne désire pas m'ériger en juge.

    c. *Moreover, only also out of respect for your electorate that, as President-in-Office of the Council, I do not wish to set myself up as judge.

    d. *D'ailleurs, seulement également par respect pour vos électeurs qu'en tant que président en exercice du Conseil, je ne désire pas m'ériger en juge.

(89) a. In this case, it is everyone who is being discriminated against.

    b. À ce propos, ce sont tous les citoyens qui font l'objet d'une discrimination.

    c. *In this case, only everyone who is being discriminated against.

    d. *À ce propos, seulement tous les citoyens qui font l'objet d'une discrimination.

In the French literature, very few studies propose an analysis for exhaustivity in the *c'est*-cleft. To the best of my knowledge, there are only two - Clech-darbon *et al.* (1999) and Doetjes *et al.* (2004) - who both offer a semantic analysis of exhaustivity. Inspired by examples from Szabolcsi, Clech-darbon *et al.* couch their analysis in the intuition that, just like the pair of examples in (90), the pairs of examples in (91) have contradictory truth values. In other words, clefts behave similarly to exclusive sentences in that (90a) and (91b) do not entail the sentences in (90b) and (91b) respectively.

(90) a. Seuls Michel et   Jean-Pierre sont   sortis.
       Only Michel and Jean-Pierre be-3pl-prst gone-out.

       'Only Michel and Jean-Pierre went out.'

    b. Seul   Michel est   sorti.
       'Only Michel be-3sg-prst gone-out.'

       'Only Michel went out.'

(91) a. Ce sont   Michel et   Jean-Pierre qui sont   sortis.
       It   be-3pl-prst Michel and Jean-Pierre that be-3pl-prst gone-out.

       'It is Michel and Jean-Pierre who went out.'

    b. C'est Michel qui   est   sorti.
       It-is   Michel who be-3sg-prst gone-out.

       'It is Michel who went out.'

Yet this data is easily challenged by embedding the prejacent under a factual, in which case only the versions that include an exclusive are felicitous (92b-92c), demonstrating that exclusive sentences and cleft sentences cannot simply be treated as allosentences.

(92) a. # Je sais   que Michel est   sorti,   mais c'est Michel qui
       # I   know that Michel be-3sg-prst gone-out, but   it-is Michel who
       est   sorti.
       be-3sg-prst gone-out.

       # 'I know Michel went out, but it is Michel who went out.'

    b. Je sais   que Michel est   sorti,   mais seul Michel est
       I   know that Michel be-3sg-prst gone-out, but   only Michel be-3sg-prst
       sorti.
       gone-out.

       'I know Michel went out, but only Michel went out.'

c. Je sais   que Michel est   sorti,   mais c'est seulement Michel qui
   I   know that Michel be-3sg-prst gone-out, but   it-is only   Michel who
   est   sorti.
   be-3sg-prst gone-out.

   'I know Michel went out, but it is only Michel who went out.'

Clech-darbon *et al.*, similarly to Percus (1997), argue that exhaustivity directly arises from the semantics of the pronoun *ce* given in (93), where *ce* introduces a definite description (ιx(P(x)) and relates two constituents, the focused constituent Q and the predicate P.

(93) $[_{dp}$ ce] $\Rightarrow \lambda$Q[Q (ιx(P(x)))]

Doetjes *et al.* (2004) mainly follow the argument proposed by Clech-darbon *et al.*, claiming that an exhaustive reading arises in the cleft only in cases where the pivot is a referential expression which a definite description can identify. Thus, cases where the pivot is not referential are argued not to get an exhaustive reading. Doetjes *et al.* gives the example in (94) where the event is predicated over by the non-referential pivot 'avec plaisir' and the copular verb 'est' is not equational (in the sense that it doesn't uniquely identify the subject).

(94)   C'est avec plaisir   que je vous   invite à participer à ce
       It-is with pleasure that I you-direct-obj-pro invite to participate to the
       séminaire.
       seminar.

       'It's with pleasure that I'm inviting you to participate in the seminar.'

In the English literature, authors like Percus (1997) and Hedberg (2000) also argue that the exhaustive inference must be treated in parallel to definite descriptions. However,

69

in these accounts, exhaustivity is argued to be presupposed rather than asserted. These accounts are discussed next.

3.2.2   Exhaustivity is a presupposition

One widely assumed presupposition of the cleft is the *existential* presupposition, which is the presupposition that there exists an individual x that has the properties P (95).

(95)   'It was X that P'

   presupposes: ∃x P(x)

Authors like Delin (1992), Percus (1997) and Hedberg (2000) have further presupposed exhaustivity, exclusivity, maximality, uniqueness, uniformity or some combination of the above. The commonality of these accounts lies in the idea that clefts pattern like definite descriptions because of the referential status of the pronoun *it*. In other words, the exhaustivity effects are reduced to a uniqueness presupposition triggered by the lexical meaning of the covert determiner (Heim, 1991). Thus, a cleft of the form in (95) has the same presupposition as a sentence containing a definite description, and is structurally indistinguishable from a sentence of the form in (96).

(96)   'The one that P is X'

The concealed definite description in the cleft amounts to exhaustivity (97), but also uniqueness (98).

(97)   'no one else than X P'

(98)   'there is exactly one X that P'

This argument quickly runs into problems because plurals are in fact allowed in clefts, as shown in (99). Now, by analogy, 'The one that laughed is Mary and Paul' simply does not make sense.

(99)   It is Mary and Paul who laughed.

But crucially, if exhaustivity were indeed a presupposition, it would survive negation and questions. This is simply not the case: in (100), the exhaustive inference in (100a) does not follow from the negated statement (100b), nor from the question (100c).

(100)   It was Mary who kissed John.

    a.   No one other than Mary kissed John.

    b.   It was not Mary who kissed John.

    c.   Was it Mary who kissed John?

In a recent account, Büring (2010a) follows Percus (1997) arguing that clefts of the form 'It is X that P' and definite descriptions of the form 'the ones that P are X' have identical asserted and presuppositional contents, as illustrated in (101). The crucial difference is that Büring (2010a) posits the presupposition to be a conditional illustrated in (102).

(101)   a.   The ones that she invited are Fred and Sue.

    b.   It is Fred and Sue she invited.

    c.   Asserted: She invited Fred and Sue.

d. Presupposed: If she invited Fred and Sue, then she invited no one else.

(102) If X ∈ P, then {X} = max(P)

This idea is attractive because it allows for exhaustivity to disappear when the cleft is negated. In (103), it simply does not follow that Mary only invited Fred and Sue (nor that Mary invited exactly/at most one person). However, we get the desired result since the presupposition (103b) survives negation.

(103) It wasn't Fred and Sue that she invited.

a. Asserted: She did not invite Fred and Sue.

b. Presupposed: If she invited Fred and Sue, then she invited no one else.

3.2.3 Exhaustivity is a conventional implicature

Halvorsen (1989) used a sort of maximality criterion to generate exhaustivity inferences of the form 'nobody but', in positive declarative clefts. In his analysis, the cleft skeleton in 'It was X that P-ed' generates two conventional implicatures given below:

(104) $\exists x\ P(x)$

(105) $\exists x\ \forall y\ (P(y) \rightarrow y = x)$

The implicature in (105) is a boundedness criterion which requires that at most one individual (within a contextually inferrable set) satisfies the predicate. (105) taken together with the existential implicature (104) entails that exactly one individual satisfies the predicate. To handle clefts with plural NPs, Halvorsen offers a more general rule in which the

72

boundedness implicature is relaxed to the cardinality of the NP. Thus, a sentence like (106) would generate the implicatures in (106a) and (106b).

(106)   It was Paul and John that Mary invited.

    a. Mary invited someone.

    b. Mary invited at most two people.

Both Horn (1981) and Atlas and Levinson (1981) provide arguments against such a view. Horn points out that if the generalization of (105) to n-sized NPs is conventionally implicated by a cleft and its negated counterpart, then (107) should be infelicitous.

(107)   It wasn't Paul and John that Mary invited, it was Fred and David.

The maximality argument also runs into trouble for mass nouns, as illustrated in (108), where it is hard to conceptualize that the speaker is committed to believe that something with at most the cardinality of 'sugar' is the cause of the headaches.

(108)   I started catching on and figured out it was sugar that was now giving me headaches.

3.2.4   Exhaustivity is a conversational implicature

Horn (1981) argues that the exhaustivity effects observed with a cleft result from default pragmatic strengthening procedures, i.e. a generalized conversational implicature. This same implicature is generated by a number of other devices, including English prosodic focus. In general, Horn claims that any device which asserts P(a) and presupposes ∃x P(x) gives rise to the following conversational reasoning: if there were other contextually relevant xs that satisfy P, the speaker would have mentioned them. If the speaker

doesn't then it is because they aren't any. Horn (1981, 134) states this formally: "The utterance in context C of any sentence which entails Fα and conventionally implicates $\exists$ x F(x) will induce a generalized conversational implicature to the effect that ¬$\exists$x (x $\neq \alpha$ & Fx), where the variable x ranges over entities determined by C."

One feature that characterizes conventional implicatures is the fact that they are *nondetachable*, which means that any expression that carries the same coded content will carry the same implicature. In other words, nondetachability arises because the implicature is attached to the content of the utterance rather than to the form of the expression that triggers it. And indeed, the exhaustivity inference in (109c) does not seem to disappear in the clefted-sentence (109a) or in the canonical sentence with contrastive focus (109b).

(109)   a. It's a hat that Mary bought.

      b. Mary bought A HAT.

      c. Mary bought nothing else than a hat.

The only difficulty with this pragmatic account concerns another characteristic of conversational implicatures: *cancelability*. Horn illustrates such a problem with the example in (110), where it seems strange for the same speaker to say both parts of the utterance without sounding like he is contradicting himself. However, cancelability occurs when the correction is uttered by the addressee (111).

(110)   Speaker A: ?It was a pizza Mary ate; indeed, it was a pizza and a calzone.

(111)   a. Speaker: It was a pizza Mary ate.

b. Addressee: Indeed, it was a pizza and a calzone.

Generally speaking, Horn proposes that cancelability might not be a attributable to the difference between conversational vs. conventional implicatures, but rather to the effort involved in producing the utterance, such that the more effort is required, the less cancelable is the inference. Despite being attractive because it explains the fact that exhaustivity in clefts can sometimes be absent, what's missing from a pragmatic account is an explanation of what clefts do that prosodic focus does not. Indeed, if clefts are semantically equivalent to canonical sentences with contrastive focus, why would people ever use them? Are there any contexts in which it is not appropriate to use a cleft instead of a canonical sentence with prosodic focus and vice versa? To spoil the suspense, I will argue that the exhaustivity in the cleft is indeed an implicature, and more specifically a scalar implicature, whereby the cleft is used by a speaker to convey the total answer to the Question-Under-Discussion that he wishes to commit to, having a reason for not using a stronger term (i.e. an exclusive) on the same scale.

3.3 Experiment 1: How exhaustive is the French *c'est*-cleft?

A forced-choice task was developed to address Clech-darbon *et al.*'s claim that the exhaustivity in the *c'est*-cleft is truth-conditional, and more generally to contribute to the crosslinguistic debate on the nature of exhaustivity. This experiment tests the nature of exhaustivity by comparing the level of exhaustivity of three sentence forms: exclusive-sentences, canonical-sentences and *c'est*-clefts. Results show that the French cleft does not pattern either like an exclusive sentence or like a canonical sentence, thus challenging a semantic account of exhaustivity (contra Kiss, 1998; Clech-darbon *et al.*, 1999).

3.3.1 The core idea

Designing an experiment to test exhaustivity is a non trivial task. The main difficulty in designing such an experiment is that, following Grice's principles of rational communication, speakers will generally say no more or less than necessary and hearers will assume that speakers (S) have observed this principle. Therefore, hearers (H) will tend to conclude that all the information contained in the statement delivered by the speaker is relevant, and that no information is withheld. In other words, H will understand the information as complete (i.e. *exhaustive*), and does not expect S to have left out any relevant piece of information. To give a concrete example, if S says that Mary laughed, H will assume that, for the purpose of the present conversation, Mary is the only relevant person who laughed. But now compare "Marie a ri/Mary laughed" to "Seule Marie a ri/Only Mary laughed". S's choice in form explicitly conveys exhaustivity in the latter case. And it would be false to say that "Marie a ri" and "Seule Marie a ri" are true in the exact same situations, since "Marie a ri" is true even in a situation where someone else laughed too. The exhaustivity in the canonical sentence is not asserted, but is simply triggered by the way people use language to communicate content.

The design of the present experiment relies exactly on this idea. While exhaustivity might be inferred by H in many cases - based on the Maxim of Quantity (make your contribution as informative as is required, but not more informative than is required) - some statements simply cannot be interpreted otherwise because exhaustivity is part of the truth-conditional meaning of the sentence. Therefore, if H wants to add information assumed to be left out by S, or if H felt that there is an inherent incompatibility between the statement and the situation, H would have to overtly contradict the statement

made by S. I assume that the most natural way in which an at issue inference triggered by some utterance is contradicted in natural dialogue is by using the explicit contradiction marker "Non/No". This is illustrated in (112), where the statement made by S carries truth-conditional exhaustivity by virtue of the exclusive expression "Seule/Only" and H uses an overt contradiction marker to correct the faulty statement.

(112)   S: Seule Marie a   ri.
        S: Only Marie has laughed.
        'S: Only Mary laughed.'

        H: Non. Jean aussi a   ri.
        H: Non. Jean also   has laughed.

        'H: No. Jean also laughed.'

On the other hand, if hearer wants to add information assumed to be left out by speaker but that this information does not contradict what speaker said (i.e. that it does not contradict the truth-conditions of the statement), H should not be using the explicit contradiction marker "no". Instead, to show a lesser amount of conflict while still supplementing information, H should use a locution like "oui,et/yes,and" or "oui, mais/yes, but". Thus, any information supplemented to a sentence such as a canonical sentence, which is not truth-conditionally exhaustive, is expected to be introduced by "oui,et/yes,and" or "oui, mais/yes (113).

(113)   S: Marie a   ri.
        S: Marie has laughed.
        'S: Mary laughed.'

H: Oui, et   Jean aussi a   ri.
H: Yes, and Jean also   has laughed.

'H: Yes, and Jean also laughed.'

We shall note that, of course, there may be other factors affecting how hearers structure the form of a continuation sentence. One possible factor could be politeness: verbally disagreeing with another person directly threatens that person's positive face (Holtgraves, 1997), and therefore speakers may occasionally seek to perform disagreement more politely, using a lighter version of 'non', such as "oui (c'est vrai), mais", or maybe even "oui, et." In that sense, they deviate from maximum Gricean efficiency, and vary in the degree of politeness use to disagree. Even more extreme, in some cultures (for example in Japanese), direct confrontation is very rude. Instead, speakers are encouraged to agree wholeheartedly with their addressee before offering their point of view or correction. This being noted, for the purpose of the present experiment, I will assume that "non" will be predominantly selected to contradict truth-conditional meaning, and "oui, mais/oui, et" will be selected to contradict implicit or presupposed meaning.

3.3.2   Method

3.3.2.1   Participants

Twenty four undergraduates from the University of Toulouse Le Mirail participated in an online forced-choice task that lasted approximately 20 minutes. All participants were native monolingual speakers of French. The participants all had normal, uncorrected vision. Finally, it was made sure that subjects were naive, i.e, they were neither linguists nor students in linguistics.

3.3.2.2 Procedure and Materials

The experiment was conducted remotely over the internet. The subjects accessed the experiment using their web browser. The browser established an internet connection to the experimental server, which was running WebExp 2.1, an interactive software package for administering web-based psychological experiments. Before the actual experiment started, participants filled a short demographic questionnaire. The questionnaire included name, email address, age, sex and language region. The language region was defined as the city/region/province where the subject learned his first language. After the demographic questionnaire, a set of instructions in French was presented to the participants.

The instructions explained that the format of the experiment would be as follow: On each trial, participants will be presented with written stimuli containing a question-answer pair in the upper half of the screen and three continuation sentences in the bottom half of the screen. The instructions emphasized that participants needed to understand each item (the question, the answer and the continuation) as being uttered by three different people, thus reading a conversation between three French speakers: "Anne" asking the question, "Paul" answering and "Nicolas" supplementing. The question asked by Anne was included to ensure that subjects correctly identified the focus element. The answer given by Paul appeared in either one of three forms: exclusive (114a), canonical (114b) and cleft (114c). The continuation supplemented by Nicolas was introduced either by "Non", "Oui, mais" or "Oui, et" (114c-i through 114c-iii). The instructions then presented the task to participants by asking that they select the continuation they found the most appropriate for the question-answer pair given. Instructions emphasized that

79

there was no correct answer and that subjects should base their judgments on their first impressions, not spending too much time thinking about any one sentence.

(114) Qui est-ce-que le directeur a grondé?
       Who is-it-that the director has scolded ?
       'Who did the director scold?'

    a. Le directeur n' a grondé que la secrétaire.
       The director not has scolded only the secretary.
       'The director scolded only the secretary.'

    b. Le directeur a grondé la secrétaire.
       The director has scolded the secretary.
       'The director scolded the secretary.'

    c. C'est la secrétaire que le directeur a grondé.
       It-is the secretary that the director has scolded.
       'It's the secretary that the director scolded.'

      i. Non, le directeur a aussi grondé le cadre.
         No, the director has also scolded the executive.
         'No, the director also scolded the executive.'

      ii. Oui, mais le directeur a aussi grondé le cadre.
         Yes, but the director has also scolded the executive.
         'Yes, but the director also scolded the executive.'

      iii. Oui, et le directeur a aussi grondé le cadre.
         Yes, and the director has also scolded the executive.
         'Yes, and the director also scolded the executive.'

After reading the instructions, the introductory phase and practice phase were administered. Items were presented with the question at the top of the screen preceded by the name of the character uttering it (Anne). The answer (the stimulus) appeared below

the question and was also preceded by the name of the second character (Paul). Finally, the three possible continuations appeared on the lower half of the screen. The name of the third character (Nicolas) preceded the first continuation and was repeated prior to the next two. The subjects indicated their choice by clicking on the continuation they wanted to select. Presentation and response procedure in the experimental phase were the same as in the practice phase.

Introductory materials. The experiment included a set of introductory materials that were designed to familiarize subjects with the judgment task. The training set contained three items for which the appropriate continuation was pre-checked. The three introductory items all very clearly supported only one of the three continuations. The subjects were told that the pre-checked continuation was a suggestion and that they might participate even if they disagreed with the pre-checked choice.

Practice materials. The experiment included practice items. The goal of these items was to have the subjects freed from the bias of the introductory phase, and have them starting to think of the task on their own. The practice set consisted of two items that were representative of the test materials.

Fillers. The experiment included distractors created to prevent the development of specific expectations or strategies on the part of the subjects. These items consisted of 3 types of sentences; entailment (e), implicatures (i) and presuppositions (p). Twelve distractors were randomly ordered within the test material.

Experimental materials. Two variables were controlled for in the experimental stimuli: the form of the answer (exclusive, canonical or cleft-sentence), and the grammat-

ical function of the focused element (subject or object). This yielded a total of 6 conditions (3 answer forms X 2 grammatical functions). Within the experiment, each participant judged exactly 4 items per condition. So, each participant judged a total of 24 experimental items and 12 fillers (for a total of 36 items). When collapsing the results for grammatical function, and looking at the condition "form of answer", a total of 192 sentences per condition were judged by the 24 participants (i.e. 192 exclusives, 192 canonicals and 192 clefts).

3.3.3 Hypotheses

Assuming the truth of a sentence of the form "X Ps" (there exists an X who does P), the addressee has a few options concerning his role in conversation.

• Continuation with *Non*: Addressee can choose to contradict the proposition if it is interpreted as $\exists x[P(x) \& x = X]$. If the predicate P in fact holds of X but also Y, the speaker can conclude that there is some inherent incompatibility between the truth-conditions of the stimulus sentence and the situation.

• Continuation with *Oui,mais*: Addressee can alternatively choose to accept the stimulus sentence despite the fact that the predicate holds for more than one entity (x = {X, Y}). In that case, addressee will not signal his disagreement overtly and the speaker will interpret the continuation as an incongruity, but not as a contradiction.

• Continuation with *Oui,et*: Addressee can accept the stimuli sentence and offer a simple continuation. In that case, he does not interpret the stimulus as exhaustive and is able to provide a continuation that's not a contradiction or a correction.

The general prediction is that there will be an effect of SF (stimulus form) on the continu-

ation chosen. The specific predictions are as follow:

- When presented with an **exclusive sentence** (which conveys exhaustivity semantically via the lexical item *seulement/only*), subjects will interpret the proposed continuation as an overt contradiction and will choose the continuation form introduced by *non/no, ...*

- When presented with a **canonical sentence** (which does not contain semantic exhaustivity), subjects will interpret the proposed continuation as a simple addition of information, and will choose the continuation form introduced by *oui,et/yes,and ...*

- When presented with a **cleft sentence**, subjects will interpret the proposed continuation as an incongruity and will choose the continuation form introduced by *oui,mais/yes,but.*

The forced-choice task presented here compares the level of exhaustivity in a cleft, a canonical and an exclusive-sentence by comparing the counts of continuation type chosen by participants for each sentence form. If counts for the cleft are congruent with the ones for exclusive-sentences, I will conclude that exhaustivity in the cleft is similar to exhaustivity in exclusive-sentences, i.e. truth-conditional. If counts for the cleft differ significantly from the ones for exclusive-sentences, I will conclude that exhaustivity in the cleft is not truth-conditional. If the sentence form has no effect on the continuation type selected, the counts should not vary deadening on whether the stimulus is a cleft, a canonical or an exclusive.

3.3.4 Results

As predicted, participants did not randomly choose a continuation, but the form of the answer did affect their choice. A goodness-of-fit chi-square statistic was applied to the data and showed that the difference in distribution of responses across the three answer forms was highly significant ($\chi^2(4) = 100$, $p < 0.001$). On the other hand, the grammatical function of the focused item had no effect on the continuation chosen. More specifically, the results show that participants overtly contradict semantically exhaustive sentences by updating the conversation with a *non*-continuation, while consistently choosing *oui, mais* for cleft sentences: the distribution of sentences chosen after exclusives was statistically different from the distribution of continuations chosen after clefts ($\chi^2(2) = 311,9$, $p < 0.001$). The difference in the distribution of continuation between canonical and cleft sentences was also found to be statistically significant, although obviously much smaller ($\chi^2(2) = 20,81$, $p < 0.001$). Thus, participants reliably chose the *oui, mais* continuation when supplementing a cleft-sentence and chose *oui, et* when supplementing a canonical-sentence. The two figures below give an overview of how the responses were distributed. Figure 1 illustrates the distribution of continuation chosen per answer type and per grammatical function of the focus. However, since the grammatical function did not have an effect on the continuation chosen, I collapsed that condition in Figure 2, which represents the counts of responses by answer type.

Figure 1: Counts of responses by Grammatical function & Answer type.
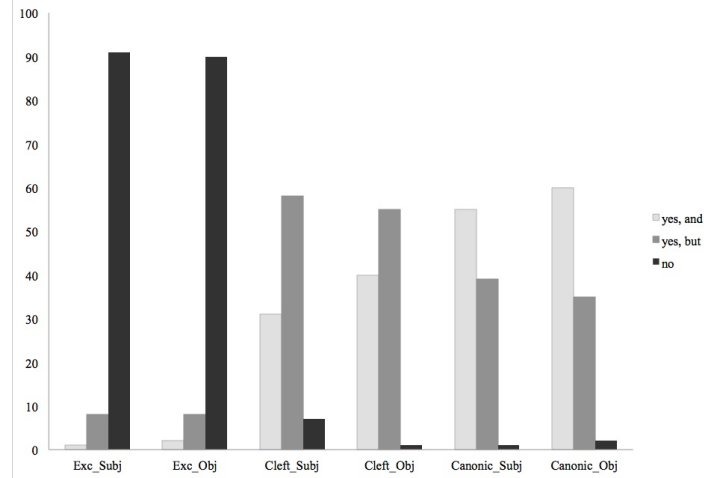

Figure 2: Counts of responses by Answer type.
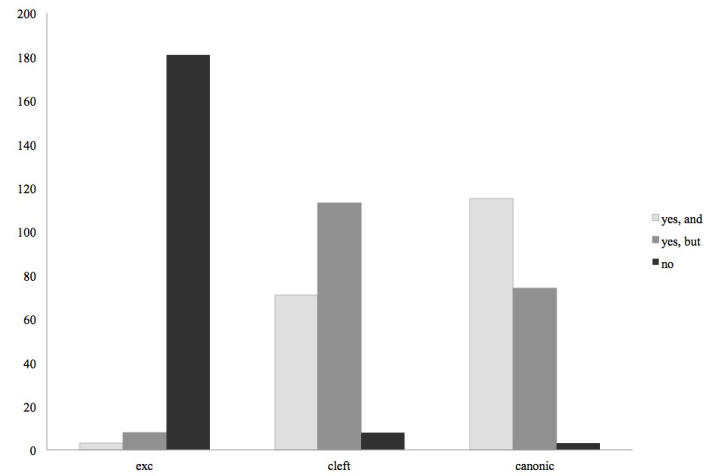
To further test the hypothesis that the continuation selected does vary with the form of the answer, I fit a logistic regression model to the data. Results are illustrated in Figure 3 where the y-axis represents the p-value.

Figure 3: Plot of the probabilities of continuations per answer form.
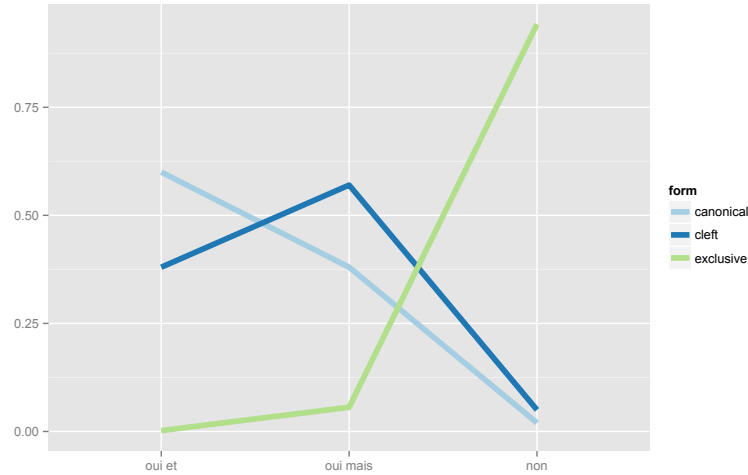
Concentrating on the results for the exclusive sentences, there is a clear pattern for overtly contradicting the proposition expressed in the answer given the situation. Participants assign an inherent incompatibility between the truth values of the answer and the continuation given by a third party: far more *non* continuations were chosen after exclusives than after clefts or even canonicals (181 out of 192 vs. 8/192 and 3/192, respectively). Within the exclusive condition, far more *non* were selected than other continuations, this difference also being statistically significant ($\chi^2(2) = 312$, $p < 0.001$). To test the hypothesis that the type of continuation chosen varies with the form of the answer, I fit a multinomial logistic regression to the data, we see that the probability of having a *non* occur after an exclusive sentence is equal to 0.94, whereas it is equal to 0.05 and 0.002 for *oui,mais* and *oui,et* respectively. These results are interpreted as demonstrating that the exclusive particle seulement/ne...que triggers a semantically exhaustive interpretation of the focus constituent.

Canonical sentences (SVO) appear to pattern the opposite way: subjects rarely choose to overtly contradict the change in focus-argument (only 3 *non* continuations were selected out of 192). Rather, subjects reliably pick the continuation introduced by *oui, et* (115 out of 192), which conveys the weakest disagreement of all possible continuations proposed in the experiment. The difference in distribution between the three continuation forms was statistically significant within this sentence form ($\chi^2(2) = 100$, p < 0.001). Results from a logistic regression model further supported that claim: the probability of having a *oui,et* following a canonical sentence is p=0.60, whereas the probability of having *oui,mais* is equal to 0.38, and *non* is equal to 0.02. I take this result as indicating that canonical sentences do not trigger a strong exhaustive interpretation, if such an interpretation is even present at all.

The results of special interest for this study are the ones concerning the *c'est*-cleft. Subjects do not directly accept the change of focus-argument as an addition to the answer presented (*oui, et*), but do not overtly contradict it either (*non*). They choose the intermediate option which is to select the *oui, mais* continuation (113 out of 192), conveying a medium degree of disagreement. The difference in distribution between the three types of continuations is statistically significant ($\chi^2(2) = 87.3$, p < 0.001). These results correlate with the prediction that cleft sentences can carry an exhaustive reading, but this reading is cancelable, therefore not part of the at-issue content of the sentence.

Finally, I tested whether there is a difference between the *oui,et* and the *oui, mais* continuations for the cleft and the canonical by creating prediction intervals for the probabilities and compared them. The results show that there is no overlap in the intervals, meaning that there is a difference in the continuation chosen in these cases. More specif-

ically, *oui,et* is always more likely with the canonical than with the cleft (the probability "oui et | cleft" falls into an interval [.31, .44] whereas the probability "oui et | canonical" falls in an interval [.53, .67]; these two intervals do not overlap). Concerning *oui,mais,* it is always more likely with the cleft than with the canonical (the probability "oui mais | cleft" falls into an interval [.51, .64] whereas the probability "oui mais | canonical" falls in an interval [.32, .44]; these two intervals do not overlap).

3.3.5 Discussion

From this experiment, I conclude that the *c'est*-cleft is associated with an exhaustive reading which is not part of the truth-conditions of the construction. As a consequence, *c'est*-clefts and the exclusive sentences do not have the same truth-conditions. Semantically, *c'est*-clefts are therefore identical to canonical sentences (as argued by Lambrecht 1994). What it means for an exclusive sentence to be semantically exhaustive is that both the *sentence meaning* (the truth conditional meaning) and the *speaker meaning* (what is meant) are aligned: the speaker intends to convey that he believes there exists no strictly stronger answer to the QUD and the sentence is indeed true iff it is the case that there are no strictly stronger answers. What it means for a cleft to not be semantically exhaustive is that it lacks alignment between sentence meaning and speaker meaning: the speaker intends to convey that he believes P but does not know that 'only P'.

	Exhaustivity is part of Sentence meaning	Exhaustivity is part of Speaker meaning
Canonical	-	-
Cleft	-	+
Exclusive	+	+

Table 3.1: Exhaustivity in different sentences

It is important to note that, as such, this experiment does not allow us to differentiate between a presupposed or an implicated nature of the exhaustive inference. With the current design, the experiment shows that the exhaustivity present in the cleft is suspendable and cancelable by the addressee without jeopardizing the truth-conditional meaning of the sentence. From the perspective of the speaker's mental state, uttering an exclusive sentence means that not only the speaker believes that there exists no strictly stronger answer to the QUD but that he wishes to assert it because he knows so. If the exhaustiveness effect is, in fact, presuppositional in nature, it would seem to be hardwired in the linguistic form of the cleft, possibly in form of a covert definite description (as argued in Percus, 1997). On the other hand, if it is a (generalized) conversational implicature (Horn, 1981), the exhaustiveness effect would not be coded in the structure of the cleft. Instead, it would be the result of a default pragmatic strengthening procedure that hearers apply, presumably in order to justify the use of a non-canonical, and uneconomical, cleft-structure by the speaker. It's worth noting that, from a cross-linguistic perspective, the findings in this experiment are in line with results of similar experimental studies on exhaustivity in the Hungarian preverbal focus position, and in the English *it*-clefts. Using analogous experimental design as the one described in this chapter, Onea and Beaver (2008) show that there is indeed an exhaustiveness effect observed with this position and construction. However, this exhaustivity is not as strong as the one observed with the truth-functional exclusive particle '*csak*/only', contra claims in the literature by Kiss (1998), among others.

Recently, Beaver (2011) proposed a new account of the nature exhaustivity in English *it*-clefts. Under his view, the cleft and an exclusive like 'only' are taken to make similar contributions to the utterance, but the exhaustivity in an exclusive-sentence is part of

89

the proffered/asserted content, whereas the exhaustivity associated with the cleft is not central to the speaker's assertion. Rather, the exhaustivity associated with a cleft is part of the non *at-issue* content of the utterance: it is understood as not being the main point of the utterance. As such, the exhaustive inference is considered as being part of a larger class of inferences that are argued to pattern in the same way, including classical presuppositions and Potts's conventional implicatures.

In this work, I choose to develop an account based on Stochastic Optimality Theory where exhaustivity is an inference that arises from a condition on the pivot position. I posit the constraint in (115) that will later be inserted into the full-fledged OT account developed in Chapter 6.

(115)   FOCCLEFT$_{exhaustive}$: A focus element with an exhaustive interpretation must be clefted.

This constraint requires a total answer to the question under discussion (QUD) be realized in a cleft. Exhaustivity arises in the cleft because the construction is used to mark that a proposition provides a full answer to what the speaker takes to be the relevant QUD. In other words, the speaker will choose to produce a cleft when she takes her answer to constitute the strongest answer possible out of the set of alternative answers instantiated by the QUD. When this constraint is ranked high, it ensures that there is no stronger alternative answer that can satisfy the QUD. If ranked low, this constraint allows non-exhaustive readings for the cleft.

Let me give a concrete example. The constraint FOCCLEFT$_{exh}$ will be responsible for the emergence of a canonical sentence instead of a cleft in an answer to the question

90

in (116). In that example, where the focus element is non-exhaustive, the canonical form (116a) is favored over the cleft version (116b), since this latter form violates EXH. Similarly, this constraint predicts that quantified focus will be realized in a canonical sentence rather than in a cleft (117).

(116)   QUD: Qui est   venu à la   fête?
          QUD: Who be-3sg-prst came to the party?
          QUD: 'Who came to the party?'

     a.   Ma famille et   aussi les amis   de Jean sont   venus à la   fête.
        My family   and also   the friends of Jean be-3pl-prst came   to the party.
        'My family and also Jean's friends came to the party.'

     b.   # C'est ma famille et   aussi les amis   de Jean qui   sont   venus à la
        # It-is   my family   and also   the friends of Jean who be-3pl-prst came   to the
        fête.
        party.
        # 'It's my family and also Jean's friends who came to the party.'

(117)   QUD: Qui a   assisté   à la   conférence?
          QUD: Who have-3sg-prst participated in the conference?
          QUD: 'Who participated in the conference?'

     a.   Trente étudiants et   quelques professeurs ont   assisté   à la
        Thirty students   and some   professors   have-3pl-prst participated in the
        conférence.
        conference.
        'Thirty students and some professors participated in the conference.'

     b.   # C'est trente étudiants et   quelques professeurs qui   ont
        # It-is   thirty   students   and some   professors   who have-3pl-prst
        assisté   à la   conférence.
        participated in the conference.
        # 'It's thirty students and some professors who participated in the conference.'

In the next chapter, I start moving from examining the interpretation of the *c'est-*cleft to examining its production. I present a couple of corpus studies examining the distributional pattern of certain expressions, determiners and quantifiers in the French cleft. I propose other constraints governing the pivot position that explain the restrictions occurring on the distributional patterns.

Chapter 4

Distributional Patterns in the French *c'est*-cleft

This chapter is concerned with the distribution of noun phrases, specifically lexical subjects, in the pivot position of the cleft and in the preverbal position of a canonical sentence. I report on two corpus studies (one quantitative and one qualitative) designed to answer the following questions: (i) what is the nature of the NPs allowed in pivot position? (ii) what are the conditions that determine NPs distribution in canonicals vs. clefts?, and more generally (iii) what does the cleft add to a sentence that a canonical alone cannot do? The type of NPs investigated in both studies is motivated by two claims found in the previous literature. First, in the literature on French, scholars claim that indefinites and quantified NPs cannot felicitously occur as pivots (Moreau, 1976; Katz, 1997) because these expressions clash with the existential presupposition and the exhaustivity associated with the cleft. Second, in the cross linguistic literature, scholars who support a semantic account of *exhaustivity* claim that modified NPs (including quantifiers, additives and numerals) are banned from pivot position because these expressions clash with the semantic exhaustive operator associated with the cleft (Kiss, 1998; Gussenhoven, 2007). The two corpus studies presented in this chapter support the claim made in the French literature for the indefinites but challenge the claim for quantified NPs, and the claims made in the cross linguistic literature. Indeed, the results provide further evidence against a semantic account of *exhaustivity* in the cleft, as they demonstrate that quanti-

fied NPs among others do in fact occur as pivots. The qualitative study analyzes the surrounding context in more details in order to get a better understanding of the conditions under which these modified NP occur. It is shown that these example arise in a special pragmatic context, one where there is *contrast* or *correction*.

4.1 Introduction

One of the most striking property of the French *c'est*-cleft is the apparent categorical freedom of the pivot, and the variety of grammatical functions it can have with respect to the element found in the relative clause. This freedom is illustrated in (118), where the pivot can equally be a NP, a PP, an AdvP, an infinitival clause or a finite clause. However, this freedom is only apparent, and a few studies show that there are in fact certain restrictions on the type of element that can occur in pivot position.

(118) a. C'est [les filles]$_{NP}$ qui ont   remporté la   première place.
           It-is the girls   who have.3pl won   the first   place.
           'It's the girls won first place.'

       b. C'est [dans cette rue]$_{PP}$ que j'ai   appris à faire du vélo.
           It-is in   this street that I-have.1sg learned to do   the bike.
           'It's in this street that I learned to ride a bike.'

       c. C'est [à mes parents]$_{PP}$ que je dois mon amour pour les voyages.
           It-is to my parents   that I owe my love   for   the travels.
           'It's to my parents that I owe my love for traveling.'

       d. C'est [à l'   automne 2003]$_{AdvP}$ que les premiers signes de la   recrudescence
           It-is in the fall   2003   that the first   signs of the upsurge
           ont   été   observés en Mauritanie, puis au Maroc.
           have.3pl been observed in Mauritania, then in Morocco.

94

'It is in the Fall 2003 that the first signs of the upsurge were observed in Mauritania and then in Morocco.

e. C'est [faire de la plongée]*infinitivalclause* que je préfère faire en vacances.
   It-is make of the dive that I prefer make on holidays.
   'Diving is what I prefer on holidays'

f. C'est [quand elle aura fini ses études]*finiteclause* que ses parents
   It-is when she will-have finish her studies that her parents
   se sentiront vraiment libres.
   refl.pro.themselves will-feel truly free.
   'Her parents will truly feel free when she finishes her studies.'

4.1.1 Background on the distribution of NPs in the French literature

In her 2009 paper, Carter-Thomas reports on data taken from the French daily newspaper *Le Monde*, and finds that out of 36 clefts collected on an ad hoc basis, 19 pivots were adjuncts (prepositional phrases or adverbials), 15 were subjects and only two were objects.[1] In her 1997 thesis, Katz makes observations about the distribution of pivots, arguing that NPs and PPs constitute the vast majority of pivots. She attributes the fact that NPs are the most common pivots to a more general constraint on the subject position: French subjects can only receive pitch accent if they exist outside of sentence-initial position.[2] In other words, the preverbal subject position in a canonical sentence is restricted to elements that are not foci.[3] Within the NP category, Katz (1997, 193) argues that certain

---

[1]Unfortunately, the study does not focus on this difference, but instead explains that a cleft is used in every case to structure the discourse and provide disambiguation as to what the focus of the sentence is.

[2]However, this claim is not undisputed, since Claire *et al.* (2004) and Beyssade *et al.* (2011), among other phonologists, argue that a focus subject can be marked via phrasing by an L% boundary tone on the right edge of the focus domain.

[3]Katz's claim closely follows the claims made by Lambrecht in his various publications.

nominals are banned from the pivot position based on a combination of semantic and pragmatic factors, namely indefinite pronouns, indefinite articles and quantifiers.

• **Indefinites:**

First, Katz notes that the referent of a pivot must *exist*: the pivot must refer to an entity that exists even if the referent is unindentifiable to either or both of the participants in the conversation. Therefore, a sentence like (119) is plainly infelicitous: the existential presupposition carried by the cleft '∃ x such that x is killed' clashes with the semantics of the indefinite pronoun *personne* which entails 'there exists no such x'. On the other hand, a canonical sentence is acceptable because it lacks the presupposition.

(119)   a.   *C'est personne qui a   tué   la   baronne   Mazette.
             It-is   no-one   who has killed the baroness Mazette.
             * 'It is no one killed the baroness Mazette.'

        b.   PERSONNE n'   a   tué   la   baronne   Mazette.
             No-one   not has killed the baroness Mazette.
             'NO ONE killed the baroness Mazette.'

Similarly, Katz argues that (120) is pragmatically ill-formed because the existential pre-supposition '∃ x such that x bothered Jules' causes (120a) to be redundant, while this is not the case in a canonical sentence. It is important to note that cases where the clefted pronoun *c'* is referential are completely well-formed, more commonly found and even preferred to their canonical counterparts, as illustrated in (121). In example (121a), the element *ami* is introduced in the left dislocation and referred to by the pivot *quelqu'un*. (121b) is pragmatically odd because French seems to disallow lexical subjects due to cognitive constraints according to which one cannot, within a single clause, both introduce a referent and predicate something about it (Lambrecht 1987, 1994).

(120) a. *C'est quelqu'un qui a   ennuyé Jules.
       It-is   someone   who has bothered Jules.

       * 'It is someone who bothered Jules.'

       b. QUELQU'UN a   ennuyé Jules.
       Someone   has bothered Jules.

       'SOMEONE bothered Jules.'

(121) a. Un ami,   c'est quelqu'un qui   connaît tout sur toi.
       A   friend, it-is someone   who knows   all   on you.

       'A friend, it is someone who knows everything about you.'

       b. ? UN AMI est quelqu'un qui   connaît tout sur toi.
       A friend   is someone   who knows   all   on you.

       ? 'A FRIEND is someone who knows everything about you.'

Katz (1997) also takes a look at indefinite pivots, and claims that they must be referential. For example, the pivot 'un idiot' in (122) must refer to a particular individual. The cleft is not used to attribute the property of idiocy to a given person, but instead the cleft identifies the pivot as having that property. In other words, the cleft is identificational rather than predicational. With definite articles, on the other hand, the pivot can refer either to certain cases that are identifiable to both speaker and addressee, or refer to the general case.

(122) *C'est un idiot qui a   écrit   ce devoir.
       It-is   an idiot who has written this homework.
       * 'It's an idiot who wrote this homework.'

(123) C'est les   chats qui aiment les   pelotes de laine.
       It-is the.pl cats who love.3pl the.pl balls   of wool.
       'CATS love balls of wool.'

- **Quantifiers:**

    Finally, Katz takes a look at clefted quantified NPs. She argues that universal quantifiers are acceptable only when the cleft is what she calls a *corrective cleft* (the cleft has the discursive function of correcting a faulty assumption in previous discourse) but not when the cleft is *identificational*, e.g. when it serves to identify a missing argument from a presupposition and inserting it into the assertion as the focus, as illustrated in (124a). Thus in (124b), the cleft is interpreted as correcting a previous statement where the speaker wrongly assumed that they do not cook crepes everyday.

(124) a. Paul: Quand est-ce-que vous faites   des   crêpes ?
           Paul: When do   you   make.2pl some crepes?
           'When are you making crepes?'

           Marie: *C'est tous les jours qu' on   fait   des   crêpes.
           Marie: *It-is all   the days that we.3sg make some crepes.
           'We make crepes every day.'

       b. Paul: C'est bien que vous   fassiez   des   crêpes le dimanche matin.
           Paul: It-is nice that you.2pl make.2pl some crepes the sunday   morning.
           'It's nice that you make crepes on sunday morning.'

           Marie: C'est tous les jours qu' on   fait   des   crêpes (pas juste les
           Marie: It-is all   the days that we.3sg make some crepes (not just   the
           dimanches).
           sundays).
           'We make crepes everyday (not just on Sundays).

    I argue that these apparent restrictions on pivots can be understood in terms of a competition between the two positions: the preverbal subject position and the pivot positions. The expressions that are claimed to be "bad" pivots will be more likely realized

in preverbal subject position, and vice versa, "bad" subjects will be realized in pivot position. The examples above seem to illustrate that two of the properties for a "good" pivot are *existence* and *referentiality*. When these properties are not met, the element is realized in preverbal position. Similarly, according to the apparent restriction on quantified NPs, a "good" pivot should be interpreted as *exhaustive*, otherwise be realized in subject position. If correct, these intuitions should be supported by empirical data drawn from the quantitative corpus study presented below: indefinite pronouns, indefinite articles and quantified NPs shall be realized in situ.

4.1.2 Background on the distribution of NPs in the cross linguistic literature

In the cross linguistics literature, the distribution of NPs occurring in the pivot position of an English *it*-cleft, or in preverbal focus position in Hungarian has been studied in relation to *exhaustivity*. Scholars arguing for a semantic account of exhaustivity, such as Kiss (1998), predict that expressions that do not allow restrictive modification (such as quantifiers, numerals and additives) are categorically banned from occurring in the positions associated with exhaustivity. For example, Kiss argues that the Hungarian preverbal position is associated with a semantically exhaustive operator, which prevents [-restrictive] expressions to felicitously occur, as illustrated in (125) taken from Kiss (1998, 252). Correspondingly, the English *it*-cleft glosses are equally bad.

(125)   a.   *Mari minden KAPALOT nézett ki   magának.
              Mary every   hat.acc   picked out herself.dat
              *'It was EVERY HAT that Mary picked for herself.'

        b.   *Mari EGY KAPALOT IS   nézett ki   magának.
              Mary a   hat.acc   also picked for herself.

'?It was ALSO A HAT that Mary picked for herself.'

c. *Mari még EGY KAPALOT is   nézett ki magának.
   Mary even a   hat.acc also picked for herself.

   *'It was EVEN A HAT that Mary picked for herself.'

d. *Mari valamit   nezett ki maganak.
   Mary something.acc picked out herself.dat

   *'It was SOMETHING that Mary picked for herself.'

Similarly for French, under a semantic account of exhaustivity in the *c'est*-cleft
(Clech-darbon *et al.*, 1999), the corresponding expressions are expected to be disallowed
in pivot position.

(126) a. *C'est CHAQUE CHAPEAU que Marie a   choisi pour elle.
          It-is each   hat   that Marie has.3sg picked for   her.

          *It's EVERY HAT that Marie picked for herself.

       b. ?C'est AUSSI UN CHAPEAU que Marie a   choisi pour elle.
          It-is   also a hat   that Marie has.3sg picked for   her.

          '?It's ALSO A HAT that Marie picked for herself.'

       c. *C'est MÊME UN CHAPEAU que Marie a   choisi pour elle.
          It-is   even a hat   that Marie has.3sg picked for   her.

          '*It's EVEN A HAT that Marie picked for herself.'

However, in a recent study, Dufter (2009) presents naturally occurring data for a
range of languages, including English and French, that directly challenge the intuitions
proposed by Kiss. Indeed, Dufter shows that, for example, universal quantifiers and addi-
tives do occur in pivot position, as illustrated in (127) taken from Dufter (2009, 97). This

data is taken to demonstrate that the pivot position of a cleft is not as restricted as previously argued, that the pivot position is not semantically exhaustive, and that the motivation for the use of a cleft cannot solely be exhaustivity.

(127)   Context: *[I see that we could spend another hour playing cat and mouse on this subject. But I am not going to take the place of the nationally elected Members of Parliament who, I am confident, will be able to interpret the texts as they have been agreed. They will also be able, as will both you and your colleague here present, to read the text and thus convince their electorate.]*

   a. Fr: D'ailleurs, c'est également par respect pour vos électeurs qu'en tant que président en exercice du Conseil, je ne désire pas m'ériger en juge.[4]

   b. En: Moreover, it is also out of respect for your electorate that, as President-in-Office of the Council, I do not wish to set myself up as judge.

4.1.3   Goal of the chapter

The overall goal of this chapter is to explore the distribution of NPs in *c'est*-clefts as opposed to canonicals based on the apparent restrictions discussed in the past literature. In analyzing corpus data, I aim to expand previous studies by providing empirical evidence supporting or challenging past claims. I also seek to answer the following questions: (i) what is the nature of the NPs allowed in pivot position? (ii) what are the conditions that determine NPs distribution in canonicals vs. clefts?, and more generally (iii) what does the cleft add to a sentence that a canonical alone cannot do? The chap-

---

[4]EP 97Ð09Ð17, Speaker 146, Original in French.

ter is structured as follow: First, I present a quantitative corpus study that examines the interaction of a group of expressions with the French cleft. The motivation for choosing the expressions examined comes from claims made in past studies: first, the claim made by Kiss (1998), in relation to the exhaustive inference, that quantifiers and additives are banned from appearing in the English cleft or in the Hungarian preverbal position, and second, the claim that the French *c'est*-cleft bans further expressions with a clashing semantics such as indefinite pronouns like *personne/nobody*, indefinite articles and quantifiers (Katz, 1997). I report on the ratio of the number of canonical tokens to the number of cleft tokens found in the EUROPARL corpus, and I observe a clear effect for exhaustivity in the cleft (the expressions examined occur to a much lower extent in a cleft than in a canonical). The results are in line with the forced-choice task discussed in the previous chapter, where the cleft was found to be more exhaustive than canonicals. The results also challenge a semantic account of exhaustivity as in Kiss (1998), as expressions such as quantifiers are found felicitously occurring in pivot position. Finally, the quantitative study shows that the restrictions on the pivot position discussed by Katz (1997) are empirically substantiated. In the second part of the chapter, I present a qualitative analysis of data drawn from the EUROPARL corpus and Google searches. This qualitative study goes more in depth into analyzing the conditions under which these expressions occur in the French cleft. It explains why in certain context, clefts are preferred over canonicals. By analyzing the surrounding discourse, I am able to determine that the NPs under study - quantified NPs, numerals and additives - are allowed in the cleft under a specific pragmatic condition, that is if the sentence conveys a contrast or a correction, and more specifically if the speaker believes the hearer has divergent expectations about the

focused constituent, or in other words, if the speaker and hearer's expectations about the focus differ.

4.2 Quantitative study

Most data in previous studies on the exhaustive inference and on the distribution of pivots in the *c'est*-cleft is drawn from judgments on constructed examples (Moreau, 1976). In this study, I supplement such evidence with quantitative data. I report on results from the distribution of subjects in pivot position and in canonical preverbal position taken from the EUROPARL corpus.

4.2.1 Methodology

The EUROPARL corpus is a parallel corpus extracted from the proceedings of the European Parliament (Koehn, 2005). In version three, the corpus has data ranging from 1996 to 2006. All eleven languages that have had an official status in the European Union since 1996 are included. Five of them belong to Germanic (English, Dutch, German, Danish, and Swedish), four to Romance (French, Italian, Spanish, and Portuguese), and two to neither branch (Finnish and Greek). The corpus contains 1,904,613 untagged sentences and 55,088,177 words of French spoken during the proceedings of the European Parliament (spoken originally in French in the proceedings). The corpus is useful because it allows to search for data in a particular original language and get a sentence-aligned translation of the data in the other ten languages. For the purpose of the current study, I used the Europarl interface with CQP query to formulate my searches, and I made sure that the original language for the data was consistently set to French in order to collect original

data and avoid translations.[5] This interface allows a multilingual search by formulation of regular expressions over attributes, and provides standard operators such as disjunctions (|), a wildcard ([ ]), the Kleene star (*) and search limit terms such as "within s", to specify that no sentence-final punctuation mark may intervene within the sequence of expressions in the query string. Thus, I searched for subject NPs in the cleft and the canonical using the formulations in (128) and (129) respectively.

(128)   Cleft: "C|Ce" "est|sont" [pos="DET:ART"] [pos="NOM"] "qui"

(129)   Canonical: [pos="DET:ART"] [pos="NOM"] [pos="VER"]

I ensured that the elements I searched for were all **subjects** in the cleft's pivot position (as opposed to direct or indirect objects) by having the "qui" relative pronoun in the search string.[6] I also ensured that the cleft was not embedded by having an uppercase C for the pronoun c'. I ensured that the searches returned a grammatical subject in preverbal canonical position by having an uppercase letter on the first word.

A couple of important notes should be made at this point. First, because the genre of the corpus is quite formal and homogenous, generalizations drawn from this study can only apply to similar genres (formal, oral speech). The data gathered in the next section is taken from a different source (Google searches) and will therefore allow me to make more general predictions about distributional patterns of expressions in *c'est*-clefts across language styles and genres. Yet, we will see that the sporadic occurrence of quantified, addi-

---

[5]http://opus.lingfil.uu.se/cwb/Europarl/frames-cqp.html

[6]In French, *qui* functions as the subject of the subordinate clause, whereas *que* functions as the direct object. Therefore, one can determine the grammatical function of the clefted element by looking at which pronoun is used to introduce the subordinate clause.

tive and exclusive pivots is not restricted to formal language, but appears homogeneously across genres. Second, in this study, I only consider the alternation between *c'est*-clefts and canonical sentences. One could wonder if there is a risk in looking at correlations only across two sentence forms when other non-canonical word orders exist in French. For example, one could wonder if a better competition should also include the existential construction *il y a* (a.k.a the *have*-cleft), arguing that if a "bad" subject cannot be realized in preverbal canonical position, is there an option between the pivot position of a *c'est*-cleft and also of a *have*-cleft. However, this risk is in fact minimal; these two types of clefts never compete because they mark different focus type, the *have*-cleft is never felicitous as an answer to a narrow focus type (i.e. where the focus is on an argument) whereas the *c'est*-cleft is preferred, and the *c'est*-cleft is rarely used for answering a broad focus question whereas *have*-clefts have that presentational function.

In the present study, the searches varied by the type of NP investigated (they were all standard lexical subject NPs headed by an indefinite, a definite, a quantifier, an additive or an exclusive). This factor "NP type" is predicted to have an effect on the dependent variable "sentence form", that is to say whether the NP occurred in a cleft or in a canonical sentence. Finally, the way I conducted my searches was to search for the frequency of tokens in one sentence form and in the other (i.e. the number of matches found for the cleft and the canonical form given a NP type). There are different ways one can study the relative frequency of NP types in canonical and cleft sentences. One way to report on the results would be to calculate the proportion of a certain NP type in the cleft out of the total number of tokens of that NP type in both sentence forms. Another way is the one used in Beaver *et al.* (2005). In that study, the authors examined the NP type occurring

alternately in an existential construction (*there is X that P*) and in a canonical sentence, and analyzed the ratio of the number of canonical tokens on the number of existential tokens. In the present study, I follow this method: First, I ran a search on the quoted strings in (128) and (129). To give a concrete example, for a search involving the universal quantifier "all/tout", I ran searches on the quoted strings in (130) for cleft tokens and (131) for canonical tokens.

(130) Cleft: "C|Ce" "est|sont" "tout|tous|toute|toutes" [pos="DET:ART"] [pos="NOM"] "qui"

(131) Canonical: "Tout|Tous|Toute|Toutes" [pos="DET:ART"] [pos="NOM"] [pos="VER"]

Figure 3: Example of Corpus results

Then, the raw results were manually sorted and a sample of the hits found for each NP types was submitted to three French native speakers for ratings. I provided the three raters

with a pdf file of up to 100 hits to judge (when the number of hits was lower than 100, they manually checked all the results. When the number of hits was higher than 100, I randomly sampled 100 hits from the results). For cleft sentences, the three raters were asked to give the sentence a 1 if it could be felicitously reformulated as a canonical sentence, a 0 if it couldn't and an X if they were not sure. Similarly, they were asked to rate a canonical sentence as a 1 if it had a felicitous cleft counterpart, a 0 if it couldn't and as an X if they did not know. After collecting the results from the three raters, I looked at the error rate (the sentences rated 0) and the unknown rate (the sentences rated X). If these rates were higher than 50%, I sampled another 100 hits and asked the annotators to rate these new sentences. The purpose of these ratings was to establish that only the hits for which syntax and interpretation was actually sought were selected. More specifically, concerning syntax, the goal was to exclude cases that did not syntactically conform to the template "C'est X qui Ps" for clefts and "X Ps" for canonicals. For example, the items that only superficially resembled a cleft (cleft lookalikes) such as restrictive sentences which have no focus-marking function were discarded.[7] Clefts where the *c'* is used as a resumptive pronoun that refers back to a preceding proposition or referent instead of a 'dummy', as in (132) taken from Reichle (pted), were also discarded.

(132) Si on donne le   prix à   une Musulmane et   à un Juif, c' est le   symbole que ce
       If we give   the price to a   Muslim   and to a Jew, it is the symbol   that it
       n'   est pas une guerre de religion.
       not is not a   war   of religion.
       'If we give the prize to a Muslim and a Jew, it's the symbol that it's not a war of
       religion.'

---

[7] cf. back to section 2.3.3 for a discussion of these structures.

Concerning interpretation, the goal of the ratings was also to exclude cases where the expression modifying the NP takes scope over the predicate rather than the NP itself. This pattern was extremely common with additives in pivot position, where the interpretation of the cleft sentence was not the desired "It is X that Ps. And it is also Y that Ps" but instead was "It is X that Ps. And it is also X that Qs", as illustrated in (133).

(133) Pierre Soulier est l'   employé le   plus ancien du Terrain   d' Aventure. C'
      Pierre Soulier is   the employee the most ancient of playground of adventure. It
      est lui,   année après année, qui réalise la   plupart des   constructions d'
      is him, year   after year,   who execute the most   of-the.pl constructions of
      importance. [. . .] C' est aussi Pierre qui   taille les   saules   et   autres arbres
      importance. [. . .] It is   also   Pierre who cuts the.pl willows and others trees
      et   arbustes du   Terrain.
      and shrubs   of-the playground.
      'Pierre Soulier is the the oldest one at the Adventure playground. He is the one, year
      after year, who builds most of the important constructions. [. . .] It's also Pierre who
      cuts the willows and other trees and shrubs around the playground.'

Once all ratings were collected, I checked for discrepancies between the annotators, and also made sure my intuitions were consistent with the ratings. Then, I calculated the ratio of the canonical frequency to the cleft frequency. For undefined cases where there was no hit, the ratio was treated as tending to infinity. To give a concrete example of how the ratio were calculated, in the search for the quantifier "tout", I first ran the search on Opus and gave the annotators a sample of the results. Then, I analyzed the ratings from the three annotators. This specific search returned a total of 2496 canonical sentences starting with "Tout|Tous|Toute|Toutes" for only 5 clefts. The hundred of canonical hits judged by the annotators were vastly acceptable (rated 1).

4.2.2 Hypotheses

If, following Kiss (1998), exhaustivity is indeed semantically encoded within the cleft, then expressions which semantics defy exhaustivity (universal quantifiers, additives, etc.) should not felicitously occur in pivot position associated with an NP which they modify. Therefore, there should be a zero occurrence rate of these expressions in the searches ran in the corpus. Moreover, following Katz (1997), the *c'est*-cleft being associated with an existential presupposition, this should prevent indefinites without a referent to occur as pivots. Thus, overall, the type of expression modifying an NP should have an effect on whether it is realized in a cleft or in a canonical sentence.

4.2.3 Results

The frequencies of each expression found in both sentence forms (cleft and canonical) are reported on in Table 4.1. From these frequencies, I then calculated the ration canonical/cleft, which is reported on in Figure 4, on a logarithmic scale where a null has a value of 1. The figure reads as follows: the higher the value, the more likely the constituent is to appear in a canonical sentence (in preverbal subject position). The closer the value is to 1, the more likely the constituent is to appear in a cleft sentence (in pivot position).

Expression	Canonical	Cleft
Indefinites, *un/une/des*	715	458
Only, *seul(ement)*	586	105
Definites, *le/ la/ les*	2097	146
Each, *chaque*	32	2
All, *tout/ toute/ tous/ toutes*	2496	104
Demonstratives, *ce/ cet/ cette/ ces*	25000	613
At most N, *tout au plus*	89	2
A few, *quelques*	79	1
Several, *plusieurs*	948	3
Many, *beaucoup*	357	0
Most, *la plupart de(s)*	636	0
Some, *certains/ certaines*	1824	1
Few, *peu*	38	0
At least N, *tout au moins*	173	0
No N, *aucun*	87	0
Personne, *nobody*	943	0

Table 4.1: Frequencies of expressions per sentence form

As expected under Katz's observations, empirical data shows the indefinite pronoun *personne/noone* is never found in a cleft since the ratio canonical/cleft tends to infinity. Indefinite articles, such as *un, une, des*, are the most frequent DPs in a cleft. When looking at the raw counts, the results come out to 715 and 458 realized in a canonical and a cleft respectively. This result is in line with the general preference for languages to have definite subjects over indefinite ones (Keenan, 1976). If we compare this result to the literature on the distribution of *existential* pivots - NPs occurring in pivot position of an existential sentence like '*There is/Il y a* a hole in the wall' - we observe similar trends: indefinites will be preferred in pivot position and definites will be preferred in sentence initial position (Mikkelsen, 2002). However, taking a close look at the interpretation of the indefinite DPs in *c'est*-cleft pivot position, all of these DPs fall in the category described

Figure 4: Ratio of French canonicals/clefts in Europarl corpus
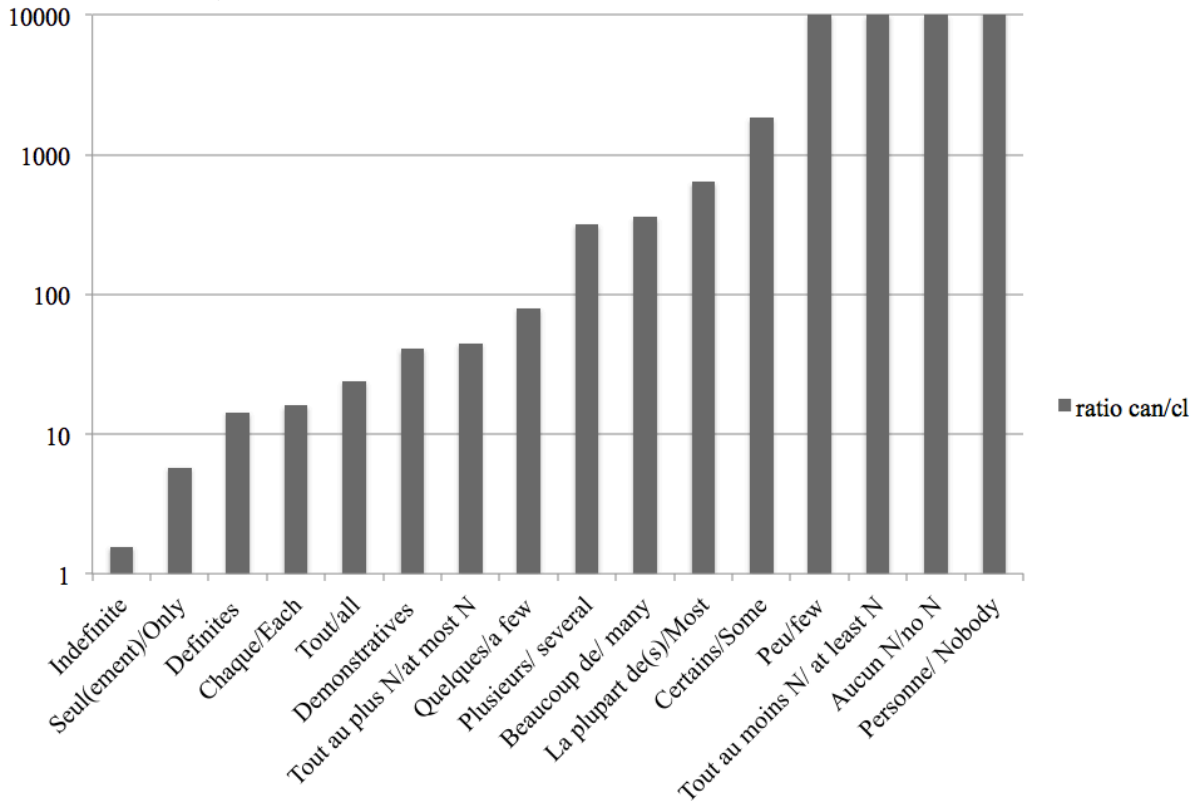
by Katz (1997) where they have a definite referent. When aligning the English translation of the corpus with the original French text, the vast majority of them include the demonstrative pronoun "this|these", and not the dummy pronoun "it" as illustrated in example (134).

(134) [Je dois dire que je regrette vivement la   suppression de la   notion de Conseil
       [I   must say that I regret   vividly   the suppression of the notion of Council
       législatif.]   C'est une décision qui ne   favorisera   pas   la séparation
       Legislative.] It-is a   decision that neg.not will-promote neg.not the separation
       des pouvoirs au sein   de l'   Union européenne.
       of   powers   in within of the Union european.
       'I must say I deeply regret the abolition of the concept of the Legislative Council. **This**

111

is a decision that will not promote the separation of powers within the European

Union.'

Definite DPs, on the other hand, are less *clefty* (i.e. less likely to occur in *c'est*-cleft pivot position) than indefinites. Looking at the ratio, a DP is x times as *clefty* as another DP if its canonical/cleft ratio is 1/x as big. Here, indefinites are 10 times as clefty as definites. Quantifiers are by and large found in canonical sentences. The ratio 'beaucoup/a lot', 'plusieurs/several' and 'peu/few' are all well over 100. Yet, they emerge in pivot position with a non-zero frequency contra Kiss (1998). In those cases, they do not always carry a correction as discussed by Katz (1997). One example which includes the most *clefty* quantifier 'tout' is given in (135).

(135) [Monsieur le   Président de la   Commission, le   Conseil européen va
       [Mister   the president of the Commission, the Council european go.3sg
       se   réunir alors   que l' Allemagne, l'   Italie, le   Portugal sont
       refl.pro.3.sg meet   whereas that the Germany,   the Italie, the Portugal be.3pl
       entrés   en récession et   que la   France est   en passe de les
       entered in recession and   that the France be.3sg in pass   of dir.obj.pro.3pl
       suivre.] Au-delà,   c'est toute la   zone euro qui connaît une dépression
       follow.] Above-all, it-is all   the zone euro that knows   a   depression
       économique mais aussi social.
       economical but   also social.
       'Mister President of the Commission, the European Council will meet while Ger-

       many, Italy, Portugal are falling into recession and while France is poised to follow.

       In addition, the entire euro area is experiencing economic but also social depres-

       sion.'

In this example, the interpretation is not corrective, but rather contrastive. The speaker

uses the cleft to mark a contrast between the list of countries that he cites in the previous discourse and the countries for which the predicate 'experiencing economic but also social depression' holds. Generally speaking, every case that involved a quantifier in pivot position carried this pragmatic inference of contrast/clash between the speakers' expectations. No quantified pivot was found in a neutral/informational context, such as for example, an answer to an explicit QUD.

4.2.4 Observations

Observation 1: No categorical ban on pivots

The results provide no evidence for a categorical ban on what can occur in the *c'est*-cleft pivot position. This supports the idea that there is no categorical exhaustive effect associated with the cleft, in other words that the cleft is semantically exhaustive. Indeed, the NPs claimed to be excluded in the past literature were found with a non-zero frequency in the searches I ran (the only expression that returned a zero occurrence rate in the cleft was "la plupart/most"). However, this result doesn't imply that there is no exhaustive effect at all, but simply that exhaustivity should not be understood as part of the truth-conditions of a cleft sentence (contra Kiss, 1998).

Observation 2: Strong non-categorical exhaustive effect

While the French *c'est*-cleft is not semantically exhaustive, it is still associated with strong exhaustive effects. These effects do affect the type of NPs occurring in pivot position, as the results show that most quantified NPs are favored in canonical sentences. Yet, there exist major differences within quantifiers (see Observation 3 below).

113

Observation 3: Not all quantifiers behave the same

Some quantifiers occur more frequently, and are less pragmatically odd than others when occurring as pivots in clefts. Following the argument that a quantifier that clashes with the presupposition and inference associated with the cleft will appear in canonical position, some quantifiers are predicted to interact better with the cleft than others. Quantifiers that have a scalar implicature like "quelques" or "certains" will be less felicitous in a cleft than quantifiers that maximally express quantification like 'tout' or numerals. Indeed, quantifiers with a scalar implicature imply that there exists a stronger alternative but that the speaker didn't choose to utter it. In sentences where the element constitutes a full answer to the question under discussion, thus no stronger alternative exist, the cleft is acceptable ... maybe even required. This explains why, for example, "tout" is more clefty than "some".

Observation 4: Strong NPs are more clefty

An interesting comparison to make is with another pivot position known to have restrictions: the existential pivot position. Existentials are sentences of the type "There is a cat on the mat". Cross linguistically, the existential pivot position has been claimed to be associated with a definiteness effect: weak NPs are more likely to occur in an existential and strong NPs in a canonical sentence. In our data, the cleft doesn't seem to pattern at all like the existential construction since strong NPs (definites, each, all, …) have lower Can/Cleft ratio than weak NPs (no, some, at least N). This result is also interesting when looking at the studies claiming that the cleft involves a definite description (Percus, 1997).

If the pronoun *c'* was indeed a definite triggering a definite description, we would expect to see definiteness effects in the pivot position.

Observation 5: No categorical ban on lexical subjects

The ratios challenge the claim mainly started by Lambrecht that French disallows lexical subjects in sentence-initial position. This is not to say that French does not prefer subjects to appear in clefts, but it shows that claiming lexical subjects never appear is too strong of a claim.[8] What this study shows is that an element whose semantics clashes with the existential presupposition and/or the exhaustive inference associated with the cleft will appear in canonical position.

4.2.5 Intermediate discussion

This corpus study supplements the forced-choice task presented in Chapter 3 by providing further empirical evidence against a semantic account of exhaustivity in the cleft: Contra Kiss (1998), universal quantifiers and additives do occur in pivot position, even if there occurrence rates are much smaller than their occurrence rate in canonical sentences. This study also provides evidence for the first research question motivating this chapter: What is the nature of NPs occurring in pivot position? Indeed, there is are clear preferences for semantic types of expressions modifying NPs, and the results can be summarized in terms of a prominent scale illustrating which expressions are more likely to occur in a cleft, that is to say which expressions constitute "good" pivots.

---

[8]The production experiment reported on in Chapter 5 will show that speakers tend to cleft lexical subjects 90% of the time. However, one shortcoming of that experiment is that it did not include modified NPs.

(136) *Pivothood scale*: Referential indefinites > exclusives > definite descriptions > strong NPs > non-referential indefinites.

According to this scale, the most important characteristic for a "good" pivot is to be referential. The data shows, following Katz's (1998) idea, that if an indefinite occurs as a clefted element, it will necessarily have a referent in the preceding discourse. This requirement is imposed on the element by the existential presupposition associated with the cleft, which would otherwise fail. Secondly, a "good" pivot is more likely to associate with an exclusive element than with a non-exclusive one, which falls in line with the claim that the cleft is pragmatically exhaustive. This preference is translated as another requirement for a "good" pivot to be exhaustive, or more specifically to constitute the total answer that the speaker wants to commit to. Finally, we already know that a "good" pivot is one that's a focus: the cleft is a construction used to mark prominence on an element that's pragmatically important. Put differently, we can say that a "bad" pivot is one that's a topic: a clefted element should not be marking information that's not answering a question-under-discussion.

(137) *Pivothood features*: Good pivots are referential, focal and pragmatically exhaustive.

4.3 Qualitative analysis

As introduced earlier, scholars such as Szabolcsi (1981), Kiss (1998) and Gussenhoven (2007) argue, for English, that all occurrences of clefting can be understood in terms of quantificational semantics. By virtue of clefting, an exhaustive reading is enforced on

the clefted constituent, and a sentence like 'It is John who baked a cake' can be considered semantically equivalent to 'Only John baked a cake'. Under this view, many expressions such as universal quantifiers, *also*-phrases and *even*-phrases are predicted to be banned from pivot position because "the semantic role of these operators is incompatible with that of identificational focus" (Kiss, 1998, 252).[9]

The present section reports on a qualitative analysis of corpus data from three different sources: two French corpora, EUROPARL and PFC[10], and Google hits. The goal is to show that the cleft is not always exhaustive (i.e. exhaustivity is not truth-conditional) by showing that expressions such as universal quantifiers do in fact occur in the cleft without causing infelicity. The data presented also sheds light on what the cleft adds to a sentence that a canonical alone does not do: For each expression presented in each sub-section, I analyze the context in which the cleft occurs and compare it to its constructed canonical variant.

Here is how I proceed: I ran a search for the expressions investigated in the quantitative study presented in the preceding section. In all my searches, I looked at the broader context (+ 3/4 sentences) to ensure the examples did not instantiate a special reading whereby the second proposition is a correction or an elaboration on the type of predicate that holds of the focus element, thus giving a kind of 'list' reading of the type 'It is X that Y. It is also/even X that Q.' I seek to exclude these cases because the expression does not quantify over the focused element occurring in the cleft, but takes scope over the predi-

---

[9]Kiss 1998 argues that English marks identificational focus via clefting.

[10]The PFC corpus is a collection of contemporary spoken French. It includes data from 450 speakers representing 75 geographical areas. The data contains recordings of speakers in various speech situations (guided conversation and spontaneous conversation).

cate in the relative clause. This is especially relevant for *also*-phrases, because, as noted by Kiss (1998, 252) herself, "*also*-phrase appears to be acceptable precisely in a context where it can be understood to identify a member of a relevant set in addition to one or more members identified previously as such for which the predicate holds, with the rest of the set still excluded."

4.3.0.1 Exclusives

The first expressions examined are expressions with a restrictive semantics such as 'seulement/only'. These expressions are problematic for semantic accounts of exhaustivity because, if exhaustivity is indeed part of the truth-conditional meaning of the cleft, one may wonder why another restrictive expression (e.g. *seulement*) is needed. Moreover, the combination of an exclusive like *only* with any other exclusive is infelicitous: **only uniquely his daughter came*, **only solely his daughter came*, **only exclusively his daughter came*. And yet, these exclusives are found in clefts without a sense of redundancy. Analyzing the extract in (138) taken from a blog discussing the story-line of various animes,[11] I argue that the cleft felicitously combines with an exclusive because it serves another function than exhaustivity; it marks the pivot element as pragmatically important.

(138) Contexte: [Ce   robot puissant et   colossal a   été   caché après la
       Context: [This robot powerful and huge   have.3sg been hidden after the
       guerre jusqu'à ce qu' il   soit   découvert par Shotaro, fils de son
       was   until   it that he be.subjunctive.3sg discovered by Shotaro, son of the
       inventeur, professeur Kaneda.]
       inventor,   professor Kaneda.]

---

[11]http://www.focusonanimation.com/tag/focus-on/

'This huge and powerful robot had been hidden after the war until it was discovered by Shotaro, the son of his inventor, professor Kaneda.'

a. À partir de ce jour, les vies et les destins de T28 et de Shotaro
At starting of this say, the lives and the destiny.pl of T28 and of Shotaro
sont liées, et c'est SEULEMENT SHOTARO qui peut piloter T28 avec
be.3pl bounded, and it-is only Shotaro who can.3sg pilot T28 with
un dispositif mobile de commande.
a device mobile of command.
*'Since that day, the lives and destinies of T28 and Shotaro are connected, and it
is ONLY SHOTARO who can pilot T28 with a mobile device control.'*

b. *À partir de ce jour, les vies et les destins de T28 et de Shotaro
* At starting of this say, the lives and the destiny.pl of T28 and of Shotaro
sont liées, et SEUL SHOTARO peut piloter T28 avec un dispositif
be.3pl bounded, and only Shotaro can.3sg pilot T28 with a device
mobile de commande.
mobile of command.
' Since that day, the lives and destiny of T28 and Shotaro are connected, and
ONLY SHOTARO can pilot T28 with a mobile device control. '*

Both (138a) and (138b) are truth-conditionally equivalent: they convey the meaning that no one else than Shotaro can pilot the robot T28. Yet, in the blog, the cleft is used and the constructed canonical equivalent is pragmatically odd. What is being discussed in this passage is the fact that someone can pilot the robot T28. We can think of it in terms of an implicit question-under-discussion (QUD) of the form 'Who can pilot T28?'. In the preceding context, another character is mentioned, Kaneda, the inventor of the robot, who constitutes a potential alternative answer to that implicit QUD. Because he is the inventor of the robot, 'Kaneda' could be thought of by the reader as the most likely character to be able to pilot the robot. Yet, he is not. In fact, the only person who can is Shotaro.

119

The cleft, unlike the canonical sentence, conveys that sense of contrast (both with that other character, and more generally with all other potential characters in the story), but also conveys the idea that Shotaro is unexpected. The use of a cleft allows the blogger to bring the element Shotaro into a prominent position, which indicates to the reader that this element must be understood as pragmatically important given the context. Despite being already present in the discourse and therefore being recoverable and given, Shotaro needs to be signaled as unpredictable or surprising given the context. Clefting serves exactly that purpose. The canonical version here is pragmatically infelicitous because it does not have the effect of surprise present with the cleft.

4.3.0.2 Quantifiers

Under semantic accounts such as Kiss (1998), universal quantifiers are predicted not to occur in identificational focus position, a.k.a as clefted constituents in English. The reasoning behind this argument is that clefted quantifiers would defy restrictive quantification since it would be difficult to combine a quantifier with an overtly exclusive expression: ?*it is only all*/?*ce sont seulement toutes*, ?*it is only every*/ ?*ce sont seulement quelques*. Kiss (1998, 252) explains that "under Kenesei (1986) framework, the semantic operation performed by the identificational focus is characterized as *exclusion by identification*. If the relevant set on which a universal quantifier operates is that specified by its restrictor, a universal quantifier performs identification without exclusion."

Examples such as (139) and (140) constitute a challenge for semantic accounts of exhaustivity. Indeed, in these examples, the quantification takes scope over the clefted element. A canonical version of the cleft is formulated in (139b):

(139) Contexte: [Il faut   donc   se   réjouir que le   Conseil des
Context: [It need.3sg therefore refl.pro.3sg rejoice that the.sg Council of-the.pl
Ministres ait   accepté l'   idée que les états membres et
ministers have.subjunctive.3sg accepted the.sg idea that the states members and
les institutions de l'   Union devaient,   ensemble, passer des   mots à
the institutions of the Union should.past.3pl, together,   pass   from words to
l'   action, et   rendre vraiment efficace le combat contre le   crime
the.sg action, and render truly   efficient the fight   against the crime
organisé. La Commission a   la   lourde responsabilité de la   mise
organized. The Commission have.3sg the heavy responsibility of the.3sg put
en application politique de 20 recommandations sur les 30 que compte le
in application political of 20 recommendations on the 30 that counts the
programme d' action.]
program   of action.]

'It is therefore gratifying that the Council of Ministers is encouraging both the Member States and the institutions of the Union to jointly put their weight behind the text of the plans of action and make the fight against organized crime truly effective. The Commission bears an important responsibility for ensuring that 20 of the 30 recommendations in the plan become a political reality.'

a. Le débat d' aujourd'hui montre que ce sont TOUTES LES
   The debate of today   shows   that it be.3pl ALL   THE.PL
   INSTITUTIONS DE L'   UNION qui doivent   assumer leur propre part de
   INSTITUTIONS OF THE UNION that must.3pl assume   their own   part of
   responsabilité dans ce   combat.
   responsibilities in   this fight.

   'Today's debate shows that it is ALL OF THE UNION'S INSTITUTIONS *that must accept their respective responsibilities in this fight'.*

b. Le débat d' aujourd'hui montre que TOUTES LES   INSTITUTIONS DE L'
   The debate of today   shows   that ALL   THE.PL INSTITUTIONS OF THE
   UNION doivent   assumer leur propre part de responsabilité   dans ce
   UNION must.3pl assume   their own   part of responsibilities in   this

121

combat.

fight.

'Today's debate shows that ALL OF THE UNION'S INSTITUTIONS *must accept their*

own responsibilities in this fight'.

(*Europarl corpus*)

In (139), the speaker conveys that, in order to fight against organized crime, a sin-
gle institution accepting its responsibilities is not enough. Rather, the whole set of insti-
tutions engaged in the fight must do so. Despite being grammatically correct, as well as
pragmatically correct, the canonical version does not highlight the element "toutes les in-
stitutions". The use of the cleft in (139a) allows the speaker to signal that the expectations
of the addressees (i.e. that one institution might be enough) are not compatible with his
belief (i.e. that more than one is required). Clefting serves the dual function of (1) allowing
the speaker to signal a discordance between his state of belief and the addressee's, and (2)
foregrounding the element of discord. In the clefted version, the most relevant piece of
information given the context (*toutes les institutions*) is brought to a prominent position.
This element is the most important in the situation both because it constitutes the answer
to the implicit QUD "how many institutions must take responsibility for the fight against
crime to be efficient?" and because it constitutes the element of discord. The discourse
goal is not to decide what actions should the institutions take in order to fight organized
crime (the answer being 'to take responsibilities' and predicated in the relative clause),
but rather to define the set of institutions that need to engage in order for the fight to be
efficient. An exhaustive reading is therefore not available in the clefted version or in the
canonical version. There is no singled out element that uniquely holds of the predicate,

but rather a set containing all the elements for which the predicate is important to hold.

A similar situation is found in example (140) below. The extract is taken from a website where critics write about newly released books and movies.[12]

(140) Contexte: [Enfin une commémoration qui ne se limite pas à
Context: [Finally a commemoration that neg.not refl. pro.3sg neg.not to
quelques heures d' émotion planétaire.]
some hours of emotion planetary.]
'Finally a commemoration which does not limit itself to a few hours of planetary

emotion.'

a. À travers les aventures tragiques et truculentes de ce Gavroche d'
At across the adventures tragical and brutal of this Gavroche of
Hiroshima, ce sont DES PANS ENTIERS DE L' HISTOIRE NIPPONNE
Hiroshima, it be.3pl SOME SWATHES WHOLE OF THE.SG HISTORY JAPANESE
qui ressurgissent.
that resurface.
'Through the tragical and brutal adventures of this Gavroche from Hiroshima,

it is WHOLE SWATHES OF JAPANESE HISTORY *that are resurfacing.'*

b. ?À travers les aventures tragiques et truculentes de ce gavroche d'
At across the adventures tragical and brutal of this Gavroche of
Hiroshima, DES PANS ENTIERS DE L' HISTOIRE NIPPONNE
Hiroshima, SOME SWATHES WHOLE OF THE.SG HISTORY JAPANESE
ressurgissent.
resurface.
'Through the tragical and brutal adventures of this Gavroche from Hiroshima,

WHOLE SWATHES OF JAPANESE HISTORY *are resurfacing.'*

---

[12]http://www.makassar-diffusion.com/?path=/cat/manga//VERTIGE

What is being discussed here is why this film is better than other commemorative movies about the nuclear bombing of Hiroshima during World War II. This information is conveyed in the proposition in bold. This piece of information is the most pragmatically important because it provides a full answer to the implicit QUD "what makes this movie different from previous movies dealing with the same theme?". The use of a cleft allows the writer (1) to signal this piece of information as being the answer to the QUD, and answer which is of essential value in the rating of the movie and (2) to contrast the perspective taken by previous movies, which are limited to depicting the emotion felt world-wide at the time of the bombing, and the way the story is told in this movie, which enables "whole swathes of japanese history" to come out. Both of these discourse goals are lost with the canonical version, which is in turn pragmatically infelicitous. From an informational point of view, the element clefted here is neither given nor recoverable from the discourse. Moreover, an exhaustive reading is not conveyed by the clefted version in (140a): it is not equivalent to 'only whole swathes of japanese history are resurfacing'.

4.3.0.3 Additives

Another piece of data challenging for exhaustive semantics view of clefting is found when additives are found in pivot position. The problem with additives is that, if exhaustivity were indeed part of the truth-conditions of a cleft-sentence, the sentence would have to be synonymous to "*only also". Note that one must be careful when searching for clefted additives because the meaning instantiated can often be "It is only X that P. It is also only X that Q", where the additive and the exhaustive marking do not enter in contradiction. The challenge arises when the additive takes scope over the clefted element, not

the predicate. Yet, we do find examples of this type. Consider (141a) and its manipulated canonical version (141b):

(141) (Europarl corpus, EP 97Ð09Ð17, Speaker 146, French original):

Contexte: [Je suis très inquiet et je déplore la réponse donnée par le
Context: [I am very worried and I deplore the response given by the
Conseil. Pour reprendre l' expression de Mme Sandbaek en ce qui concerne
Council. To retake the expression of Mrs Sandbaek in it that concerns
le chat, je vois bien qu' on pourrait encore jouer au chat et à la
the cat, I see clearly that we.3sg could still play at cat and at the.sg
souris pendant une heure dans ce domaine. Mais, je ne me
mouse during an hour in this domain. But, I neg.not refl.pro.1sg
substituerai pas aux députés et parlementaires élus
will-substitute.1sg neg.not to-the.pl deputies and MPs elected
des parlements nationaux qui sauront, je leur
from-the.pl parliament national who will-know.3pl, I direct-obj.pro.3pl
fais confiance pour cela, interpréter les textes tels qu' ils ont
make.1sg trust for that, interpret the.pl texts as that they have.3pl
été arrêtés. Ils sauront également, comme vous allez le faire
been decided. They will-know.3pl also, as you go.2pl it make
ainsi que votre collègue ici présente, faire la lecture du texte et
as-well-as that your colleague here present, make the reading of-the text and
donc convaincre aussi leurs électeurs.]
therefore convince also their voters.]

'I am very concerned and unhappy about this answer from the Council. To go back to Mrs Sandbaek's expression concerning a cat, I see that we could spend another hour playing cat and mouse on this subject. But I am not going to take the place of the nationally elected Members of Parliament who, I am confident, will be able to interpret the texts as they have been agreed. They will also be able, as will both you and your colleague here present, to read the text and thus convince their electorate.'

125

a. D'ailleurs, c'est ÉGALEMENT PAR RESPECT POUR VOS ÉLECTEURS qu' en tant
   Indeed,   it-is ALSO   BY RESPECT FOR   YOUR VOTERS   that in case
   que président en exercise du   Conseil, je ne   désire   pas
   of president in exercise of-the Council, I neg.not wish.1sg neg.not
   m'   ériger en juge.
   refl.pro.1sg set-up in judge.

   Moreover, it is ALSO OUT OF RESPECT FOR YOUR ELECTORATE THAT*, as President-*

   in-Office of the Council, I do not wish to set myself up as judge and thus replace

   the national Members of Parliament and their different governments.

b. ?D'ailleurs, en tant que président en exercise du   Conseil, je ne
   ?Indeed,   in case of   president in exercise of-the Council, I neg.not
   désire   pas   m'   ériger en juge ÉGALEMENT PAR RESPECT POUR
   wish.1sg neg.not refl.pro.1sg set-up in judge ALSO   BY   RESPECT FOR
   VOS ÉLECTEURS.
   YOUR VOTERS.

   '?Moreover, as a President-in-Office of the Council, I do not wish to set myself up

   as a judge OUT OF RESPECT FOR YOUR ELECTORATE*.'*

The meaning conveyed by both sentences is that the president in office of the

Council does not wish to set himself up as a judge and tries to rationalize his decision by

giving his reasons, one of which is the respect for the electorate. In the previous discourse

context, the speaker starts by expressing his disappointment about an answer given by the

Council, which is the main reason why he does not wish to be a judge. In the proposition

under study, he adds to the discourse another reason for making this decision, hence the

use of the additive expression "également". The QUD here is therefore something like 'Why

does the speaker not want to set himself up as a judge?'. Yet, the manipulated canonical

version does not constitute a natural answer to that QUD and is judged as pragmatically

infelicitous by French native speakers. Why? Because it relegates the important piece of information (the additional reason "par respect pour vos électeurs") to the less prominent sentence final position. In that case, the emphasis on that element disappears. By using a cleft, the speaker wants to indicate to the addressee that the reason 'par respect pour vos électeurs' is the most pragmatically important element in the sentence. The speaker may also want to convey that, despite the tendency to overlook the feelings of the electorate in politics, he is willing to decline the authoritative position of judge because of his concern for the electorate. The speaker therefore assumes such a reason is unexpected in the mind of the addressee and signals it as such by realizing it in a cleft. Clefting also marks that piece of information as the full answer the speaker is willing to commit to. This effect simply does not appear in the canonical version. Furthermore, the clefted example (141a) is not to be interpreted exhaustively. It is not the case that "par respect pour vos électeurs" constitutes the unique reason why the speaker does not wish to step in as a judge. Rather, this reason is added to a previous reason already mentioned in the discourse. The cleft here is used to fully answer the QUD and indicate the pragmatic importance of the clefted element because of its unexpected status in the addressee's mind.

4.3.0.4   Numerals

The next example addresses the argument made by Kiss (2009) about the reading available for numerical expressions. She argues that the reading "*at least n, n or more*" is not available, and the exhaustive operator present in the focus position forces the specific reading "*exactly n*". Yet, in the example (142a), the element clefted isn't interpreted as "*exactly n*", but is instead understood as a vague number: $n \leq x \geq m$. In (142), the cleft's

function is again to signal contrast between the number of menhirs that were standing at the very beginning of the 20th century versus the number of menhirs that are standing today. The main point of the sentence is not convey the uniqueness of the answer.

(142) a. Grâce aux collectivités territoriales, ce patrimoine est en train de
Thanks to-the.pl communities territorial, this heritage is in process of
sortir petit à petit de l' ombre. En effet, tout au début
coming-out little by little of the shadow. In effect, all at-the.sg beginning
du XXe siècle, c'était TOUT AU PLUS ENTRE TROIS ET SEPT
of-the.sg XXth century, it-was ALL AT MOST BETWEEN THREE AND SEVEN
MENHIRS qui étaient encore debout. Aujourd'hui, ce sont AU
MENHIRS that be.imparfait.3pl still up. Today, it be.3pl AT
MOINS 80 MENHIRS qui ont retrouvé la verticale.
LEAST 80 MENHIRS that have.3pl found the vertical.
'Thanks to the territorial communities [...] this heritage is slowly coming out

of the dark. Indeed, in the very beginning of the 20th century, it was AT MOST

BETWEEN 3 AND 7 MENHIRS *that were still standing. Today, it is* AT LEAST 80

MENHIRS *that are in a vertical position again.'*

(*Wikipedia*)

b. *En effet, tout au début du XXe siècle, TOUT AU PLUS
In effect, all at-the.sg beginning of-the.sg XXth century, ALL AT MOST
ENTRE TROIS ET SEPT MENHIRS étaient encore debout.
BETWEEN THREE AND SEVEN MENHIRS be.imparfait.3pl still up.
Aujourd'hui, AU MOINS 80 MENHIRS ont retrouvé la verticale.
Today, AT LEAST 80 MENHIRS have.3pl found the vertical.
*'*Indeed, in the very beginning of the 20th century,* AT MOST BETWEEN 3 AND

7 MENHIRS *were still standing. Today,* AT LEAST 80 MENHIRS *are in a vertical*

position again.'

4.3.0.5 *En tout cas*

In his 2006 thesis, Spector discusses the expression "en tout cas/in any case" in relation with the exhaustive reading. He argues that the addition of this expression into a sentence blocks a possible exhaustive reading. Instead, the speaker wishes to convey that he may be poorly informed and that his answer may not be full (i.e. maximally true). His knowledge of the world only allows him to commit to the answer given but he does not wish to commit to whether his answer fully resolves the QUD. Let's give an explicit example. Consider (143):[13]

(143) a. J' aime bien ce clip, qu' on   dirait   bricolé avec de
        I   like well this video, that we.3sg say.conditional.3sg made-up with some
        vieilles images. D'ailleurs, je ne   sais pas   si c' est de la   récup
        old   images. Besides,   I neg.not know neg.not if it is of the.sg recycling
        ou s' ils   ont   vraiment tourné des   images. C' est EN TOUT CAS MARIE-EVE
        or if they really shot   the   images. It   is IN ANY CASE MARIE-EVE
        MARIE-EVE LAPOINTE qui a   réalisé ce   clip, avec l'   aide de
        LAPOINTE who   has directed this video, with the help of Vincent
        Vincent   Chalifour.
        Chalifour.

   *'I kind of like this video, which seems to be made up of old images. Besides, I
   don't know if it is recycled or if they have really shot these images. It is in any case
   Marie-Eve Lapointe who directed this video, with the help of Vincent Chalifour.'*

   b. *MARIE-EVE LAPOINTE EN TOUT CAS a   réalisé ce clip, avec l'   aide de
      MARIE-EVE LAPOINTE IN ANY   CASE has directed this video, with the help of
      Vincent Chalifour.
      Vincent Chalifour.

---

[13]http://www.mathieugruel.fr/2010/09/06/de-la-musique-en-images-1/l

*'*MARIE-EVE LAPOINTE IN ANY CASE directed this video, with the help of Vincent Chalifour.'*

In this passage, the speaker's goal is not to exhaustively identify the person who directed the video: the phrase "en tout cas" but also the addition of a second person working on the video "with the help of Vincent Chalifour" prevent exhaustivity. Yet, the cleft is chosen over a canonical, which is again pragmatically infelicitous. One reason here could be that French disallows prosodic marking on sentence initial position. In this case, the cleft would be prosodically motivated, occurring in order for the NP 'Marie-Eve' to fulfill rightward prosodic constraints of the language. But in reality, the cleft serves here again a discourse function that a canonical alone cannot do. It allows the speaker to bring the attention of the reader onto a new piece of information that he wants to convey as pragmatically important.

4.3.0.6 Exclusive-sentences

Finally, yet another convincing example against a semantic account comes from the comparison of the negated version of a cleft-sentence and a *seulement*-sentence. If it is the case that exhaustivity is truth-functional, then *seulement*-sentences and clefts should have the same truth-conditions, i.e be synonymous. Yet, when one negates these two types of sentences, they are very clearly not synonymous (Horn, 1981). Consider the pair of examples in (144) taken from a gardening blog, where (144a) is the original version, (144b) is the manipulated *seulement*-sentence version.[14]

---

[14]http://forums.jardinage.net/viewtopic.php

(144) Contexte: (Speaker A) Mon tout dernier zygopetalum, le   Blue Bird, pourquoi
       Context: (Speaker A) My   very last   zygopetalum, the Blue Bird, why
       la   nouvelle pousse est toute en train   de brunir? Il me   fait
       the new   sprout is all   in process of fading? It refl.pro.1sg make.3sg
       exactement la   même chose que   mon zygo précédent. Y   a   quelque chose
       exactly   the same thing than my   zygo before.   There has some   thing
       qui cloche dans la   culture de cette orchidée chez moi.   C'est l'   une
       that wrong in   the culture of this   orchid   at   mine. It-is the.sg one
       des   orchidées que je trouve la   plus belle   et j' arrive   pas,
       of-the.pl orchids   that I find   the.sg most beautiful and I manage neg.not,
       mais vraiment pas   à la   faire survivre. Pour mettre en
       but   really   neg.not to dire.obj.pro.3sg make survive. To   put   in
       contexte, le   zygo est juste à côté de moi, sur la   table d' ordi.   Il reçoit
       context, the zygo is just to side of me, on the table of computer. It receives
       des   rayons du soleil entre   8h et   9h à ce temps-ci de l'   année. Il est
       some ray   of sun   between 8   and 9   at this time   of the year.   It is
       environ   à 6 pieds de la   fenêtre.
       approximately to 6 feet   of the.sg window.

       'My very last zygopetalum, the Blue Bird, why is the new sprout all fading? It's doing

       exactly what happened with my previous zygo. There is something wrong with the

       culture of this orchid at my house. It's one of the orchid that I find the most beautiful

       and I cannot, like really not, make it survive. To put back into context, the zygo is

       just right next to me on the coffee table. At this time of the year, it gets some sun from

       8am to 9am. It's approximately 6 feet away from the window.

   a. Réponse: (Speaker B) Les zygopetalums ne   sont pas   des   plantes
       Response: (Speaker B) The zygopetalums neg.not are   neg.not some plants
       d' ombre profonde. Mais JE DOUTE QUE   CE SOIT   LE
       of shade profound. But   I   DOUBT THAT IT   BE.SUBJUNCTIVE.3SG THE
       MANQUE DE LUMIÈRE qui a   fait   en sorte qu' il perde
       LACK   OF LIGHT   that has make.3sg in sort   that it loose.subjunctive.3sg
       des   feuilles. Ils   sont sensibles à la   photopériode mais je ne
       some leaves.   They are sensitive to the photoperiod   but I neg.not

131

sais   pas   jusqu'à quel point.
know.1sg neg.not until-to what point.

'Zygopetalums are not deep shade plants. But I DOUBT THAT IT IS THE LACK OF

LIGHT *that caused it to lose its leaves. They are sensitive to photoperiod but I*

don't know to what extent.'

b. Les zygopetalums ne   sont pas   des   plantes d' ombre profonde.
The zygopetalums neg.not are   neg.not some plants   of shade profound.
Mais JE DOUTE QUE   SEUL LE   MANQUE DE LUMIÈRE a   fait   en sorte
But   I   DOUBT THAT ONLY THE LACK   OF LIGHT   has make.3sg in sort
qu' il perde   des   feuilles.
that it loose.subjunctive.3sg some leaves.

'Zygopetalums are not deep shade plants. But I DOUBT THAT ONLY THE LACK OF

LIGHT *caused it to loose its leaves.'*

In this extract, Speaker A seeks information about the reasons why his orchid is

losing its leaves. The context given in (144) shows the question asked by A 'why is the

new stem getting brown?'. Both speakers share background knowledge that, when cul-

tivating orchids, the most important aspect for a successful grow is the location of the

orchid (does not like direct light and does not require too much light). Being aware of

this, speaker A gives more context to B a few sentences later by describing the amount

of light his orchid gets daily and how far from a window it is located, therefore assuming

that the variable 'light' must be the reason for its leaves loss and stem browning. Speaker

B answers (144a), and conveys doubt about the lack of light being actually involved in

leaves loss. While not openly stating it, her answers implicates that she believes some-

thing else is the cause (i.e. lack of humidity). A very different meaning is conveyed in the

manipulated example (144b) where the exclusive is present. In this version, the meaning

conveyed is that the lack of light is indeed responsible for the leaves loss. The speaker conveys doubt as to "lack of light" being the *unique* reason, but does not convey doubt as to whether "lack of light" *is* a reason. So, in (144a), the meaning expressed is that the lack of light is not responsible, while in (144b), the lack of light is (one of) the reason(s). In the first example, the cleft is used as a device to highlight the unpredictability of the reason 'lack of light'. Speaker B indicates to speaker A that she does not share his intuition about 'light' being responsible for the unsuccessful growth of the orchid. Thus, speaker B's proposition is contradicting speaker A's belief state. The clefted element "lack of light" is therefore unpredictable, or surprising to A. The cleft signals the element "lack of light" as the most important in the context because it is not the answer expected by speaker B. In the *seulement*-sentence (144b), the important meaning conveyed is that "lack of light" is not, in B's mind, the unique reason for which "leaves loss" holds. Therefore, the communicative goals of speaker B are different in the two sentences: in (144a), B wants to indicate to A that his expectation is not correct, while in (144b), B wants to indicate to A that his expectation is correct but that it is too restricted.

4.4   Discussion and Constraints

The goal of the chapter was to answer the following questions: (i) what is the nature of the NPs allowed in pivot position? (ii) what are the conditions that determine NPs distribution in canonicals vs. clefts?, and more generally (iii) what does the cleft add to a sentence that a canonical alone cannot do? The first question was discussed in the intermediate section 4.2.5 where I argued that there is no categorical ban on the NPs that can occur in pivot position, however there are clear tendencies regarding which NPs are more

likely to occur in a cleft vs. in a canonical. The tendencies for the NPs preferred in pivot position were NPs that are referential and that strengthen an exhaustive reading such as exclusives. I showed that quantified NPs did not all behave the same; strong NPs were often preferred to week NPs in pivot position. Yet, a more detailed analysis of corpus data revealed that some of the expressions investigated appeared in specific contexts, which sheds light on the second research question; the conditions determining NPs distribution in canonicals vs. clefts. I argued that the competition between the pivot and the subject position underlies the competition between the cleft and the canonical forms. However, the ungrammaticality of one form is not directly related to the grammaticality of the other. Indeed, I have presented data showing, for example, that the definite article could occur felicitously both in the pivot or in the subject position, and similarly, quantifiers can occur in both sentences without causing ungrammaticality. The conditions under which an element to constitute a "good" pivot are that the element must be referential, a focus, pragmatically exhaustive and/or express a contrast or a correction. Finally, I answered the more general question about what a cleft adds that a canonical sentence alone cannot do. By arguing that the cleft is not semantically exhaustive, I ruled out *exhaustivity* as the invariant function of the cleft. Therefore, even though there are strong exhaustive effects associated with the cleft, these effects are not the only thing setting the cleft apart from the canonical. Examining corpus data qualitatively, taking a close look at the context in which a cleft is uttered and the meaning that the speaker wishes to convey by using that form rather than another, is where I pulled out the answer. Along the lines of what is proposed in Hupet and Tilmant (1990), a cleft is used if there is a need to *express a discrepancy between the speaker and the addressee's beliefs and/or expectations*, in other

words if there is a need to express a contrast or a correction. However, contrary to Hupet and Tilmant (1990), I argue that this discrepancy does not need to be explicitly stated in the previous discourse. It can also be inferred by the speaker himself. That is to say, the speaker can make the assumption that the addressee disagrees with his beliefs, even if this assumption reveals to be false later on. The cleft highlights the fact that X Ps, and that this relationship between the pivot and the predicate is surprising, unexpected or disputed. Because the preverbal subject position lacks this function, if an element is not marked for contrast or correction, or if the speaker does not wish to convey a discrepancy between beliefs, then the focus element will be realized in as a preverbal subject. This principle can be translated into a constraint like (145).

(145)   FOCCLEFT$_{contrastive}$: A focus element with a contrastive interpretation must be clefted.

In the qualitative study, I looked at non-exhaustive examples involving expressions like quantifiers and additives. For all of the examples discussed, I showed that an exhaustive reading was not available but that all these examples denoted a contrast. In order to explain this fact, the constraint FOCCLEFT$_{contrast}$ must be ranked higher than the constraint FOCCLEFT$_{exh}$ posited in Chapter 3, which demands that pivots are interpreted as total answers to the QUD and preverbal subjects are not. This will be demonstrated in the full OT analysis developed in Chapter 6.

Chapter 5

Focus Realization: Evidence from a production experiment

The goal of the chapter is to provide empirical evidence on how the information structural category of *focus* is realized in French. More specifically, I show how different grammatical (subject vs. non-subject) and pragmatic (informational vs. corrective) types of focus are realized. The chapter contributes to a growing trend in modern semantics to investigate phenomena by experimentally testing the claims associated with them. First, I report on a semi-spontaneous production experiment which shows that (1) a strict one-to-one mapping between focus and its realization is not tenable in French (as also observed in other languages like English and German by Zimmerman and Onea 2011), (2) there is a significant relationship between a marked focus realization and a semantically stronger interpretation, and (3) French displays an asymmetry between subject and non-subject foci.[1] Second, I discuss special cases from naturally occurring data (taken from different sources like Google hits and corpora) where subjects and non-subjects are realized differently from their default position: subjects occur in situ and non-subjects occur in a cleft. All results presented challenge Lambrecht's claims that *c'est*-clefts are always used to mark focus on arguments and that there is a one-to-one relationship between the grammatical type of focus and the way it is realized.

---

[1]Interestingly, this phenomenon is not specific to French but also occurs in Western African languages studied by Zimmermann.

5.1 Introduction

The fact that the phenomenon of *focus* is found universally in the world's languages is mirrored by the vast amount of literature on the expression of this category. English is widely described, along with many well represented languages such as Spanish and German, but there is also a growing body of literature on less-known and endangered languages such as languages from Africa (Zimmermann, 2006; Zimmerman *et al.*, 2010). While all these languages have grammatical means to mark focus, these means often differ from one language to the next, and sometimes even within a same language. English, a head-initial SVO language where stress is rightmost, is one example of languages that use phonological means to indicate focus. The focused element must receive prosodic prominence signaled by a pitch accent (Jackendoff, 1972). In (146) for example, the narrow object focus "a cookie" must receive main stress, otherwise leading to incompatibility in discourse and thus to infelicity.

(146)   What did Mary eat?

    a.   Mary ate A COOKIE.

    b.   # MARY ate a cookie.

Italian and Spanish also are head-initial SVO languages where stress naturally falls rightmost. When marking non-subject elements, they behave similarly to English; the focus element occurs by default in the rightmost position where main stress is assigned.

(147)   a.   Spanish: Maria comio UN PASTEL.

    b.   Italian: Maria UN BLAH.

Yet when marking focus on subjects, these two languages differ from what English does. While English displays prosodic flexibility by allowing main stress to be moved leftward, Spanish and Italian require main stress to occur rightmost. Hence,

stress is constrained to appear rightmost even if the focus occurs sentence initial (148). Similarly in Hungarian, both syntactic reordering and pitch accent play a role in focus marking: elements found in preverbal position receive pitch accent and that pitch accent often disambiguates cases where the DP or PP is composed of more than one constituent (149).[2] Finally, many languages from West Africa signal focus via syntactic reordering and/or morphological markers such as copulas, functional heads or affixes (see West Chadic example (150) from Zimmerman *et al.* 2010).

(148)   Spanish. Who ate a cookie?

   a. # MI HIJA   comió una galleta.
      # MY DAUGHTER ate   a   cookie.

   b. Comió una galleta MI HIJA.
      Ate   a   cookie MY DAUGHTER.
      'MY DAUGHTER ate a cookie.'

(149)   Hungarian. Q: What kind of car did Peter buy?

   A: Péter egy PIROS autót vett.
   A: Peter a   RED   car   bought.

   'Peter bought a RED car.'

(150)   West Chadic. Q: What is he chewing?

---

[2]from Zimmerman and Onea (2011)

A: Tí bà wúm Á   KWÁLÍNGÁLÁ
A: 3sg IPF chew FM COLANUT.

'He is chewing COLANUT.'

5.2 The debate

In French, the realization of focus is still controversial. Scholars differ on three major issues: (1) the realization of subject-focus, (2) the factors motivating the use of syntactic reordering as a focus marking strategy and (3) the exact nature of the prosodic correlates of focus marking.

Concerning the first issue, Lambrecht (1994) argues that, unlike English or German, spoken French must resort to clefting in order to signal focus on lexical arguments. Canonical sentences are prohibited, especially if headed by lexical focus subjects.

(151) How did you get here?

English: PAUL drove me.

German: PAUL fuhr mich.

French: C'est PAUL qui m'a conduit.

Yet, many studies such as Delais-Roussarie and Post (2008) and Claire *et al.* (2004) provide constructed examples arguing that a subject can be realized in situ in a declarative sentence:

(152) Q: Qui a   acheté une mandoline?
       Q: Who has bought a   mandolin?
       'Who bought a mandolin?'

A: MALLARMÉ a   acheté une mandoline.
A: MALLARMÉ has bought a   mandolin.

'MALLARMÉ bought a mandolin.'

In her dissertation, Hamlaoui (2009) aims to reconcile these diverging views by arguing that the two marking strategies - in situ and clefting - belong to two distinct grammars; the first is associated with standard French and the latter with demotic French. Indeed, studies that argue for an in situ realization of subject-focus are mostly based on constructed examples, while studies that argue for clefting rely on corpus data.

The debate regarding subject-focus marking leads to the second conflict introduced above, namely the factors motivating the use the cleft. Syntactic approaches argue for **pragmatically** motivated focus marking, whereby the focus element moves into a cognitively prominent position because grammar bans it from appearing elsewhere. This ban is attributed to an infelicitous mapping between syntax and information structure: salient parts of the discourse or new discourse referents (i.e. foci) constitute 'bad' grammatical subjects and must be realized in an alternative sentence form (Lambrecht, 1987). The obligatory rearrangement of focused material into a cleft guarantees the formation of a prosodically unmarked structure where main stress does not occur sentence-initial. In the case of subject focus, this non-canonical sentence also prevents the subject from being interpreted as a topic (Lambrecht, 1994). On the contrary, most phonologists argue for **prosodically** motivated focus marking, where the focus element is directly merged into the position where grammar assigns main stress. A subject-focus is realized in a cleft because the bi-partite structure of the cleft creates two independent Intonational Phrases which allow the focused subjects to receive main stress on the right edge of the first IP. A

non-subject focus, on the other hand, receives main stress in its canonical position.

Finally, a third issue arises. Amongst phonologists, the nature of the prosodic correlates of focus marking is not settled. While some argue that focus is marked via phrasing and tones, others claim that focus is marked via pitch movements (see. Jun and Fougeron 2000 vs. Féry 2001) .

The present chapter is structured as follow: First, I discuss the syntactic and prosodic strategies argued to mark focus in French. Then, I present a production experiment designed to test how speakers of French realize focus on various grammatical constituents and in different contexts. I report on the results that provide empirical support for a subject/non-subject asymmetry, but also for a certain amount of free variation. I complement the results by discussing a case where the canonical sentence is preferred over a cleft for subject focus when the focus is quantified. I conclude by relating the results found for French to cross-linguistic data, especially data discussed by Zimmerman for Chadic languages.

5.3   Focus marking strategies in French

In the literature on French focus marking, phonologists always acknowledge clefting as a prominent way to signal focal information in the language (Féry, 2001). And most syntacticians agree that a focus element bears some kind of prosodic prominence (Lambrecht, 1994). However, syntacticians, led by Lambrecht's work, tend to argue that prosody can never be used as the primary strategy to mark focus, while phonologists demonstrate that syntactic rearranging does not always have to occur (Claire *et al.*, 2004).

5.3.1 Syntactic strategy

Syntactic focus marking refers to the fact that a focus constituent is reordered relative to the other elements in the sentence. The work of Lambrecht argues that focus in French is primarily marked via syntactic means, proposing three different focus structures: argument-focus, predicate-focus and sentence-focus. Under this view, there exists a strict one-to-one relationship between the grammatical function and the grammatical realization of focus, and each realization is described below.

Argument-focus (AF) is said to identify or specify an argument in a presupposed open proposition, to have a focus-presupposition articulation, and to be realized in a *c'est*-cleft.[3]

(153) **Argument focus**

Qui est-ce-qui a vendu sa voiture?

'Who sold his car?'

  a. #Jean (a   vendu sa voiture).
     John have sold   his car.

  b. C' est Jean (qui a   vendu sa voiture).
    It is   John who have sold   his car.
    'JOHN sold his car.'

Despite being grammatically correct, the canonical sentence in (153a) is considered unavailable to French speakers. Lambrecht (1994) and Katz (1997) claim that, in colloquial French, a lexical focus subject cannot appear in situ (i.e. sentence initial) but must

---

[3]According to Lambrecht (2001, 485), this type of focus is also referred to in the literature as "specificational", "identificational" or "contrastive". The focus structure AF is also referred to as "focus-presupposition".

be realized in a *c'est*-cleft, where it can also be accented. Thus, there is no alternation be-

tween cleft and SVO sentences because no variability is available. The post focus element

doesn't bear an accent because it doesn't add information to the open proposition. The

information structure of argument-focus according to Lambrecht (1994) is represented

below:

(154) **Subject focus**

Sentence: C'est Jean qui a vendu sa voiture. *'It is John who sold his car'*

Presupposition: 'x's car has been sold'

Assertion: 'x = Jean'

Focus: 'Jean'

Focus domain: NP

(155) **Object focus**

Sentence: C'est sa voiture que Jean a vendue. *'It is his car that John sold'*

Presupposition: 'Jean sold his x '

Assertion: 'x = voiture'

Focus: 'voiture'

Focus domain: NP

Predicate-focus (PF) is said to predicate a property relative to a given topic. Accord-

ing to Lambrecht (2000, 615), this structure "expresses a pragmatically structured propo-

sition in which the subject is a topic (hence within the presupposition) and in which the

predicate expresses new information about this topic. The focus domain is the predicate

phrase (or part of it)." The arguments 'Jean' and 'voiture' are presupposed; they are part

143

of the knowledge shared between the interlocutors and are active by being mentioned in the preceding discourse.

(156) **Predicate focus**

Qu'est-ce-que Jean a fait de sa voiture?

'What did Jean do with his car?'

a. Il l' a   VENDUE.
He it have sold.

b. Sa voiture, il l' a   VENDUE.
His car,   he it have sold.
'He SOLD his car.'

The information structure of (156) is represented in (157). The focus of a PF structure bears prosodic prominence (i.e. the verb is accented). The object *voiture* is marked as a topic, hence being excluded from the focus domain. It is realized either as an unaccented pronoun *l'*, or as an unaccented pronoun combined with a lexical NP that can appear either in a left or a right dislocation.

(157) *Sentence*: (Sa voiture), il l'a vendue.

Presupposition: 'the speaker's car is a topic for comment x'

Assertion: 'x = vendue'

Focus: 'vendue'

Focus domain: VP

Finally, a *sentence-focus* structure (SF) is argued to be used for an event-reporting or presentational sentence type where the focus domain extends over both the subject

144

and the predicate.[4] Lambrecht (1994) claims sentence-focus is invariably realized either in an *avoir*-cleft introduced by a personal pronoun coindexed with the speaker (158a), or in an existential sentence (*Il y a* X *qui* Y) introduced by the dummy pronoun 'il' and the locative 'y' (158b).

(158) **Sentence focus**

Qu'est-ce-qui s'est passé?

'What happened?'

a. J' ai   mes élèves   qui sont   en grève.
   I have my   students who be-3pl in strike.
   'My students are on strike.'

b. (Il) y   a   mes élèves   qui sont   en grève.
   (It) there have-3sg my   students who be-3pl in strike.
   'My students are on strike.'

In this focus structure, no pragmatic presupposition is formally evoked. The element in the relative clause is not presupposed but is part of the focus domain, therefore asserted and accented via a pitch accent.

(159) *Sentence*: J'ai mes élèves qui sont en grève.

Presupposition: —

Assertion: 'x = mes élèves sont en grève'

Focus: 'mes élèves sont en grève'

Focus domain: S

---

[4]SF is often called "broad focus" or "all-new" in the cross-linguistic literature on focus.

The table in (5.1) summarizes the three focus categories proposed by Lambrecht (2000, 615).

	Arguments in Focus	Predicate in Focus
Predicate Focus	–	+
Argument Focus	+	–
Sentence Focus	+	+

Table 5.1: Focus Categories from Lambrecht (1994)

Under this view, focus realization is induced by the pragmatic structure of the sentence. In other words, the structure in which the focus is realized depends primarily on the pragmatic structuring of propositions into presupposed and non-presupposed elements. But these structures are not undisputed. It is interesting to see that scholars have considerably differing assumptions concerning the acceptability of a structure in a given context. In most syntactic studies (Lambrecht, 1994, 2000), it is assumed that a narrowly focused subject must obligatorily occur in a cleft with the structure *C'est [S]$_f$ qui V O*. Yet, some scholars have recently argued that in some cases a subject is preferred sentence-initially, undermining the idea that French has a categorical ban on canonical sentences with a lexical subject (Beyssade *et al.*, 2010). One case mentioned by Beyssade *et al.* (2010) is the semantic type of the NP; the plurality of the focus in (160) appears to hinder the use of a cleft. The canonical alternative is therefore produced. In her 2001 study, Féry also argues that French can realize focus subjects in situ via special phrasing and tones: the subject focus projects its own Phonological Phrase delimited by a final boundary tone (161).[5] Finally, other cases where an in situ focused subject is felicitous include quanti-

---

[5]A drawback of this study lies in its methodology. Because the author was specifically testing for prosodic correlates of focus marking in French, the stimuli used to elicit response were written cards. Therefore,

fied subjects as discussed in Chapter 4 such as (162).[6]

(160) [Waiter arriving at table with several drinks]

    A: Qui a   commandé les cafés?
    A: Who has ordered   the coffee?

    A: 'Who ordered coffee?'

  a. B1: # C'est PIERRE ET MARIE (qui ont   commandé les cafés).
      B1: # It-is Pierre and Marie (who have-3pl ordered   the coffee).

      B1: # It's PIERRE AND MARIE (who ordered coffee).

  b. B2: PIERRE ET MARIE (ont commandé les   cafés).
      B2: Pierre   and Marie   (have-3pl ordered the coffee).

      B2: PIERRE AND MARIE (ordered the coffee).

(161) a. Q: Qui peint le   garage en noir?
      Q: Who paints the garage in black?

      Q: Who is painting the garage black?

  b. A: LE   GARÇON peint le   garage en noir.
      A: THE BOY   paints the garage in black.

      'A: THE BOY is painting the garage black.'

(162) Pour le secteur qui nous occupe dans l' immédiat, celui   de la pêche
      For   the domain that us   occupies in   the immediate, the-one of the fishing,
      il nous paraît absolument nécessaire de veiller à la   bonne représentativité des
      it us   seem absolutely   necessary to ensure to the good   representativity of

---

participants were not spontaneously producing subject-focus structures, but were reading material. Participants were however instructed that they could change the sentence if they felt that there was a more appropriate way to state the answer. The results show that some participants did change canonicals for clefts (the author reports 20 changes into clefts out of 400 sentences produced, 13 cases being for subject foci.

    [6]Example taken from the Europarl Corpus.

divers acteurs au sein du comité consultatif. TOUS LES ÉTATS MEMBRES
different actors with in the committee advisory. ALL THE STATES MEMBERS
INTÉRESSÉS PAR L' ACTIVITÉ DE LA PÊCHE doivent être représentés, et
INTERESTED BY THE ACTIVITY OF THE FISHING must be represented, and
leurs professions aussi.
their professions also.

'Concerning the domain that matters to us, fishing, it seems absolutely necessary

to us that we ensure the fair representativity of the various players within the advi-

sory committee. ALL THE MEMBER STATES THAT ARE INTERESTED IN FISHING MUST

BE REPRESENTED, AND THEIR PROFESSIONS AS WELL.'

Lambrecht's AF structure is also challenged by many scholars who show that non-subjects do not need to be realized in a cleft (Hupet and Tilmant, 1990; Vion and Colas, 1995; Hamlaoui, 2008). Both Hupet and Tilmant (1990) and Vion and Colas (1995), relying on experimental data, demonstrate that the grammatical function of the focused element was not a crucial factor in its realization: The first study shows that non-subjects will remain in situ if the proposition expresses a weak contrast (where the degree of conviction of the speaker is low and doesn't clash with the hearer's beliefs) as illustrated in (163a). More generally, the second study shows that contrastive non-subjects are more often marked via focal accents (163b) than via clefting (163c).

(163) Context: You live in a communal house. It's Monday. You invited Luc, a friend, to come have dinner next Monday. One of your roommates sees you and says:

 a. A1: *(Weak contrast)* I thought Luc was going to come for dinner tonight, is that right?

 A2: *(Strong contrast)* I can't believe Luc is coming for dinner tonight. He was

here just a few days ago.

b. B1: (Non,) il vient   la   semaine prochaine.
   B1: (No,)   he comes the next   week.

   '(No,) he's coming next week.'

c. B2: (Non,) c'est la   semaine prochaine qu' il   vient.
   B2: (No,)   It-is the next   week   that he comes.

   '(No,) it's next week that he's coming.'

Finally, another challenge for Lambrecht's account is that a few studies claim broad focus can also be marked via a *c'est*-cleft, in which case the information in the relative clause is semantically presupposed (Clech-darbon *et al.*, 1999; Delais-Roussarie *et al.*, 2002; Doetjes *et al.*, 2004). This piece of data (illustrated in example 164) is in direct opposition with the idea that there exists a strict mapping between focus and its realization, and with the idea that *c'est*-clefts only have a focus-ground articulation.[7]

(164) Q: Qu' est ce qui s'   est passé?
       Q: What is   it that refl.3sg is   happened?

       'What happened?'

       A: C'est le   petit   qui est tombé dans l'   escalier.
       A: It-is the small-one who is   fallen in   the stairs.

       'The little one fell down the stairs.'

5.3.2 Prosodic strategies

In the literature on French focus marking, the majority of accounts (whether syntactic or prosodic) agree on the fact that a focused element receives some kind of prosodic

---

[7]The first idea is also recently refuted in Zimmerman and Onea (2011).

prominence, either in pivot position if the element is clefted or in canonical position if the element remains in situ. It is also undisputed that French is a right-edge language: by default, main stress falls rightmost (i.e. sentence-final). Prosodic accounts of focus marking also unanimously argue that the use of a specific syntactic construction (i.e. a cleft) is not mandatory to express focus in French. However, there is presently no consensus on the nature of the prosodic realization of focus. Two perspectives exist: scholars claiming that focus marking is prosodic and is marked via phrasing and tones, and scholars claiming that focus is marked by pitch movements.

Phrasing and tones:

Scholars who argue in favor of phrasing being . The differences stem from the type of phrasing referred to when talking about focus, but also the type of tone occurring associated with the focus constituent. The literature on French phonology distinguishes two types of *phrasing*. A first one that reflects the syntactic and metrical organization of the utterance and is realized at the Phonological Phrase level (PhP), and a second that reflects the informational status of the content of the utterance and is realized at the level of Intonational Phrases (IPs). Within the literature specific to the prosody of focus, scholars have used both types of *phrasing* to account for focus marking. Two prominent accounts within the literature are Féry (2001) and Claire *et al.* (2004). They both propose an account of how the segmentation into phrases is determined by the semantics and pragmatics associated with the utterance. Yet, one major difference is that Féry argues focus marking occurs at the PhP level, while Claire *et al.* argue that it occurs at the IP level. Féry also differs from other scholars by assuming (1) a complete absence of lexical stress in French which means that there is no stress at the word level, and assuming that (2) *phrasing* takes

150

over some of the roles attributed to pitch accents in other studies. Under her view, focus induces its own Phonological Phrase (PhP) and is tonally delimited by boundary tones.

Féry *et al.* (2010) claim: "In French, there is no deaccentuation, but also no true accentuation. Important prosodic phrases, as for instance those containing new or focused referents, are more clearly phrased than others, with small breaks separating them from neighboring phrases, and larger boundary tones. Important words are often realized at places where they get tonal excursions, but not necessarily so..."
Phrasing and syntactic structure coincide as far as possible, being processed in parallel, following the Alignment Theory developed in McCarthy and Prince (1993) in (165):

(165)   Align(Xmax, Right, PhP, Right): The right edge of a Maximal Projection coincides with the right edge of a Phonological Phrase.

The information status of each constituent often determines which is phrased separately since Féry (2001) argues that focused constituents are realized in independent PhP (166). Non-focused constituents, on the other hand, are dephrased, the exception being that non-focused subjects typically remain in their own PhP. In the case of a sentence-focus, the author argues that each syntactic element forms its own PhP. In cases where a single word is focussed within a phrase, often to mark contrast, Féry observes an exaggerated phrasing on that specific word (166f).

(166)   Focus induced Phrasing:

   a.   Subject-focus: [Le garçon]$_{PhP}$ peint le garage en noir.

   b.   Direct Object-focus: [Le garçon]$_{PhP}$ peint [le garage]$_{PhP}$ en noir.

c. Indirect Object-focus: [Le garçon]$_{PhP}$ [peint le garage]$_{PhP}$ [en noir]$_{PhP}$.

d. Predicate-focus: [Le garçon]$_{PhP}$ [peint]$_{PhP}$ le garage en noir.

e. Sentence-focus: [Le garçon]$_{PhP}$ [peint le garage]$_{PhP}$ [en noir]$_{PhP}$.

f. Contrastive Word-focus: Jean conduit la [petite]$_{PhP}$ décapotable de Marie.

Based on experimental data, Féry finds some amount of variation in the realization of these structures correlated with the length and weight of the focused element. For example, in the case of sentence-focus, Féry observes that subject-verb-object can be realized within the same PhP if the constituents are very short (167). The more they increase in length and weight, the more these constituents will tend to build their own PhP.

(167)   Variation in Sentence-focus Phrasing

Q: *Que se passe-t-il?*

a. [S][VdO][iO]: [Le petit garçon]$_{PhP}$ [peint le garage]$_{PhP}$ [en noir]$_{PhP}$.

b. [S][V][dO]: [Les marins]$_{PhP}$ [ont réparé]$_{PhP}$ [le grand mât]$_{PhP}$ [hier matin]$_{PhP}$.

c. [SVdO]: [La fille peint le garage]$_{PhP}$.

Similarly for predicate-focus, verbs can be realized in their own phrase or with the following object depending on their length and weight, and depending on whether there is a indirect object or not (168).

(168)   Variation in Predicate-focus Phrasing

Q: *Que fait Marie/ le garçon?*

a. [S][V][dO]: [Marie]$_{PhP}$ [a réparé]$_{PhP}$ [le garage]$_{PhP}$.

b. [S][VdO][iO]: [Le garçon]$_{PhP}$ [peint le garage]$_{PhP}$ [en noir]$_{PhP}$.

c. [S][V][dO][iO]: [Le garçon]$_{PhP}$ [peint]$_{PhP}$ [le garage]$_{PhP}$ [en noir]$_{PhP}$.

For subject-focus and objects-focus, however, Féry argues that the constituents always form their own PhP. In that case, the predicate can be either phrased or dephrased depending on its weight. For Subject-focus, the rest of the sentence was always dephrased (169).

(169) Variation in Argument-focus Phrasing

Q: *Que fait Marie/ le garçon?*

a. Subject: [Le garçon]$_{PhP}$ peint le garage en noir.

b. Object: [Les marins]$_{PhP}$ [ont réparé]$_{PhP}$ [le grand mât]$_{PhP}$.

[Marie]$_{PhP}$ peint [le garage]$_{PhP}$ en noir.

[Marie]$_{PhP}$ [peint le garage]$_{PhP}$ [en noir]$_{PhP}$.

Unlike the predictions made in Lambrecht's account, this prosodic account predicts variation within focus realization. The grammatical function of the focus plays an important role but is not the one and only decisive factor in phrasing placement. There are in fact focus-independent rules of phrasing placement, such as for example, the weight and length of the element. Thus, a single phrasing pattern is not systematically mapped onto a single focus structure. These observations are interesting because they cast doubt on Lambrecht's claim about the cleft's usage: argument foci are always realized via clefting. Moreover, they correlate with cross-linguistic findings on focus under-determination (see Zimmerman and Onea 2011).

153

On the other hand, in their 2004 paper, Claire *et al.* argue that a focus constituent induces its own phrase at a higher level: the intonational level. The phonological phrase, under this account, is left undiscussed as it plays no role in focus marking. In preceding studies, the same scholars do discuss the well-formedness of PhP, but the PhP is not involved in focus marking. Claire *et al.* (2004) argue that focus occurs within its own IP and comes with a low boundary tone (L%) on the right edge of the IP. A copy of this tone always occurs sentence finally. Unlike Féry (2001), they do not observe variation.

(170) a. Subject-focus: [Mallarmé L%]$_{IP}$ [a acheté une mandoline L%]$_{IP}$.

b. Sentence-focus: [Mallarmé a acheté une mandoline L%]$_{IP}$.

c. Contrastive Word-focus: [Mallarmé a acheté dix-sept L%]$_{IP}$ mandoline L%]$_{IP}$.

Concerning boundary tones, Féry and Claire *et al.* agree: both argue that the focused constituent is delimited by a low boundary tone on its right-edge.

Pitch movement:

The second view taken by phonologists is couched in the autosegmental-metrical framework (Beckman and Pierrehumbert, 1986; Ladd, 1996; di Cristo, 1999; Rossi, 1999). Within this standpoint, accounts diverge to a lesser extent than in the previous standpoint. Here, the phonology of focus is assumed to be intonational. French resorts to special pitch accent movements to signal focus (David and Yoo, 2000; Jun and Fougeron, 2002; Beyssade *et al.*, 2011). All studies also assume a fixed position for stress at the phrase level: the final full syllable of a phonological phrase is realized with longer duration and

higher intensity.[8] In a series of experiments on object-focus, Beyssade *et al.* (2011) show that focused elements can be marked by two types of intonational marking: they host a nuclear pitch accent on their right edge and/or are intonationally highlighted (involving an initial accentuation). The empirical results show variation within the prosodic marking with the pitch accent varying in its position; either occurring at the right edge of the focused element or at the end of the utterance (left-most where default stress is assigned).

5.4   Experiment 2: A semi-spontaneous production task

In this study, I am specifically interested in examining the effect of syntactic and pragmatic factors on the realization of focus (i.e. grammatical function and contrastiveness, respectively) in a controlled environment that triggers semi-spontaneous speech. The experiment provides evidence for what type of variability is actually observed in the speaker's output. It answers whether French exhibits a strict one-to-one relationship between marked focus realization and semantically stronger focus interpretation, or whether a marked focus realization is due to discourse-semantic factors, i.e. pragmatic factors (for example, the speaker's expectation of low discourse expectability of the focused constituent for the hearer). Another way to put it is that the experiment directly addresses Kiss's claim that languages realize two semantically different type of focus (i.e. *identificational/corrective* and *informational*) in two distinct ways, via syntactic and prosodic means respectively. Finally, the experiment provides ample evidence for a subject/non-subject marking asymmetry in French.

---

[8]Jun and Fougeron (2002, 147) note that the domain of stress in French has changed over the course of its evolution from Latin from a lexical domain to a phrasal domain.

5.4.1 Predictions

Considering what has been observed in previous works, the predictions for the realization of focus in French are the following:

1. If French categorically bans prosodic marking from sentence-initial position, participants should produce 100% of answers to a "Qui" question with a *c'est*-cleft (thus replicating the results from Hupet and Tilmant 1990, and supporting the claim in Hamlaoui 2009).

2. If there exists a strict one-to-one relationship between focus and its grammatical realization, we should observe a systematically different focus realization for each grammatical focus (as proposed in Lambrecht 1994). Therefore, a single focus realization would be mapped onto a single informational structure. In other words, the same focus realization (e.g. *c'est*-clefting) could not be mapped onto more than one grammatical focus (e.g. argument, predicates or sentence).

3. If there is a strict one-to-one relationship between marked focus realization and semantically stronger focus interpretation, *c'est*-clefts should always be used to mark contrast/correction, even in cases where the focus is a non-subject, and canonicals should be reserved for informational focus (following Hupet and Tilmant 1986; Kiss 1998).

5.4.2 Method

5.4.2.1 Participants

Participants were 21 native French speakers who had lived in Austin, TX for less than a year. Their age ranged from 28 to 42 years. They all had normal or corrected-to-normal vision.

5.4.2.2 Procedure

In this methodology (adapted from Gabriel 2010)[9], participants were seated in front of a computer screen where they saw a series of five different pictures, each of them accompanied by a written description of what was happening in the picture. Then, the participants saw a series of questions associated with each picture, delivered in written form, one by one, below the corresponding picture. An example of one of the pictures with the description and a following question is presented in Figure (5.1).

The purpose of the questions was to help participants identify the correct focus element targeted in the answer. Participants were asked to answer each question orally. Before the experiment started, participants received the written instruction to avoid answering with single constituent fragments by reusing the elements given in the question as much as possible. They were however instructed to answer as naturally as possible. Finally, the instructions made it clear to the participants not to worry if they seemed to

---

[9]Four out the 5 pictures used in my production experiment were reproduced from Gabriel's experiment with the author's consent. For these 4 questions, the questions triggering subject, direct object and indirect object foci in neutral context were directly translated into French from Gabriel's experiment on Argentinian Spanish. My experiment differs in that it also included predicate and whole sentence focus. My experiment also differs from Gabriel's in that he did not study the role of context and therefore did not have a corrective context vs. neutral context.

Voici Marie qui achète un journal au kiosque.   2. Où est-ce-que Marie achète le journal?
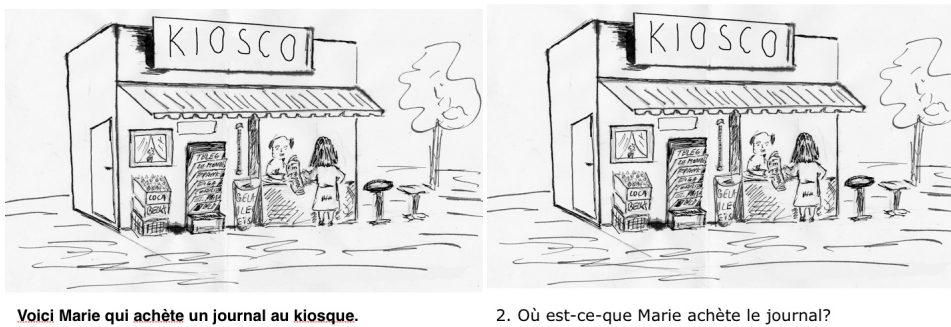
Figure 5.1: *Sample pictorial stimulus and a sample question.*

give similar answers to different questions (i.e. there was not necessarily one single correct wording for each question). The stimuli were presented as PowerPoint files on a Mac laptop. The participants's answers were recorded with the program Audacity.

5.4.2.3 Conditions

The design included two independent variables: grammatical function (GF) and context (CT). Each of these variables had different levels. For GF, I tested grammatical subjects, direct objects, indirect objects, double objects, predicate and whole sentence (six levels). Two contexts were tested: neutral and corrective (two levels).[10] Looking at the variables, the combination of the levels of the factors gives a total of 6*2 (twelve) conditions. However, only eleven tested were because the condition "whole sentence" x "corrective context" is not pragmatically appropriate: a correction can only been offered on a part of the sentence that's already activated in the discourse. It is therefore hard to imag-

---

[10]The latter context is labeled corrective rather than contrastive as often seen in the literature, because the stimuli found in that context involved sentences where the focus element was incorrectly identified and participants had to offer a correction according to what was really depicted.

ine the addressee offering a correction to a broad question-under-discussion "Qu'est-ce-qu'il se passe/What is going on?".

5.4.2.4 Material

A total of five lexicalizations were created for each condition, giving a total of 55 questions. The 55 questions were divided up over five different pictures. Therefore, each picture was associated with a question triggering each condition (i.e. eleven experimental stimuli per picture). Some examples are given below in Table 5.2.

	Neutral	Corrective
Subjects	Qui est-ce-qui achète un journal au kiosque? (*Who is buying a newspaper at the stand?*)	Regarde, on dirait Julie qui achète un journal au kiosque, non? (*Look, it seems that Julie is buying a newspaper at the stand, no?*)
Direct objects	Qu'est-ce-que Marie achète au kiosque? (*What is Marie buying at the stand?*)	Regarde, on dirait que Marie achète des cigarettes au kiosque, non? (*Look, it seems that Marie is buying cigarettes at the stand, no?*)
Indirect Objects	Où est-ce-que Marie achète un journal? (*Where does Marie buy a newspaper?*)	Regarde, on dirait que Marie achète un journal dans un supermarché, non? (*Look, it seems that Marie is buying un newspaper at the supermarket, no?*)
Double Objects	Qu'est-ce-que Marie achète et où? (*What does Marie buy and where?*)	Regarde, on dirait que Marie achète des cigarettes au supermarché, non? (*Look, it seems that Marie is buying cigarettes at the supermarket, no?*)
Predicate	Qu'est-ce-que fait Marie? (*What is Marie doing?*)	Regarde, on dirait que Marie vole le journal, non? (*Look, it seems that Marie is stealing the newspaper, no?*)
Sentence	Qu'est-ce-qu'il se passe? (*What is going on?*)	n/a

Table 5.2: Sample questions per grammatical function and context

The variable CT was controlled for by changing the type of question presented to

159

the participants. The description of the pictures remained unchanged across conditions. All questions in corrective contexts were of the form "Regarde, on dirait que X Y, non?". This form conveys that the speaker was making a wrong statement, and the participant was instructed to offer a correction. All questions in neutral contexts were presented in a non-clefted form, that is to say of the form in (171a) rather than in (171b), in order to avoid syntactic priming effects.[11] The form in (171a) was also preferred to the form in (171c) because the latter might have primed canonical answers and is also less natural.

(171)   a.   Qui   est-ce-qui   achète   un   journal   au   kiosque?
              Who   is-it-who   buys   a   newspaper   at-the   stand?

         b.   Qui   c'est qui/   C'est qui   qui   achète   un   journal   au   kiosque?
              Who   it-is   who/   It's   who   who   buys   a   newspaper   at-the   stand?

         c.   Qui   achète   un   journal   au   kiosque?
              Who   buys   a   newspaper   at-the   stand?

              'Who is buying a newspaper at the stand?'

The variable GF was also controlled for by changing the question presented to the participants. The open variable in the question corresponded to the focus element targeted in the answer. Thus, a question targeting a subject focus would be of the form "*Qui est-ce-qui?/Who it-it-who?*" and a question targeting an object focus would be of the form "*Qu' est-ce-que/qui?/What is it-that/who?*".

Along with the 55 experimental items, twenty distractors were created and divided

---

[11]It is interesting to note however that a pilot study where *question form* was included as a variable returned no significant effect. In other words, there was no clear effect of the form of the question (clefted or not) on the type of answer delivered by participants. Therefore, in the large scale experiment reported on in this work, I chose to only show *est-ce-que*-Questions, which seem to be the most natural and common way to ask a question in French.

up amongst the five pictures (i.e. four distractors per picture). The distractors did not follow the same general pattern as the experimental stimuli but asked a descriptive question about the picture. The experimental questions and distractors were then pseudo-randomized within each corresponding picture.

Each participant produced a total of 55 answers (five per condition) plus 20 distractors. Having 21 speakers taking part in the experiment, I collected a grand total of 1155 answers; 105 per condition for the 11 conditions. To illustrate, I collected 105 answers of subject-focus in neutral contexts, 105 answers of subject-focus in contrastive contexts, 105 answers of direct-object-focus in neutral contexts, etc. All five verbs allowed for a ditransitive construction. While the four first verbs of the verbs allowed for a exhaustive/non-exhaustive reading alternation, the last verb "gagner" entailed an exhaustive reading. All subjects were animate and were designated by one element proper nouns (i.e. Marie, Blanche-Neige, Pierre). Direct objects were both animate and inanimate.

Trials where participants did not correctly identify the focused element and thus produced a mismatched answer were excluded. The data were then transcribed and organized according to the strategy used in answering the question for the grammatical and pragmatic focus expressed.

5.4.3 Results

Figure (5.2) gives results collapsed over both context types. It concentrates on showing the type of focus marking strategy used depending on the grammatical function of the focus. Note that the bar for broad sentence on figure (5.2) is smaller due to the

161

fact that I collected sentences only for neutral contexts. Figure (5.3) shows results of the type of sentence given as an answer in a neutral (informational) context. Figure (5.4) on the other hand, shows results of the type of sentence given as an answer in a corrective context. All results are given in raw count numbers.
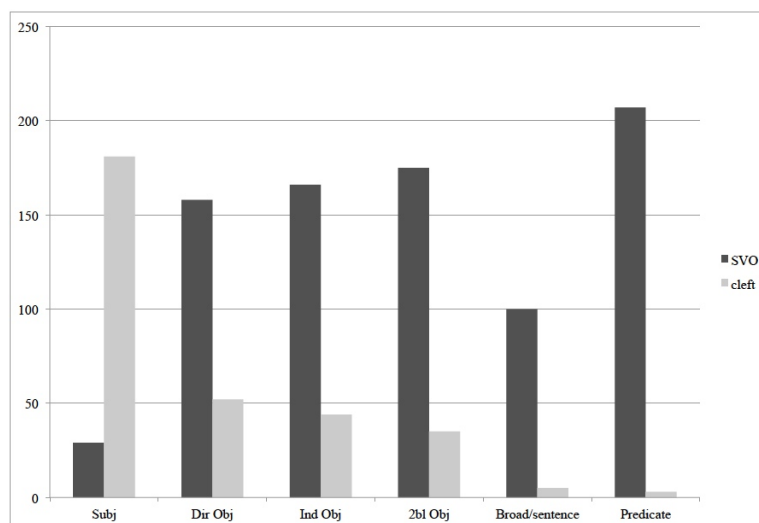

Figure 5.2: Total count per grammatical function collapsed over contexts

As predicted, the difference in distribution of focus strategy across the six grammatical functions is highly significant ($\chi 2(5)$ = 502.41, p < .0001). However, almost all this variation can be traced back to the differences between the subject condition and the other five conditions. There is indeed a clear subject/non-subject marking asymmetry: figure (5.2) shows that subjects are essentially realized via a marked structure (cleft), whereas non-subjects are not.

Hypothesis 1 predicted that focus subjects would be banned from occurring in situ. However, despite a non-significant rate of occurrence, participants did realize some subjects sentence initially. This result allows us to claim that French does not categor-
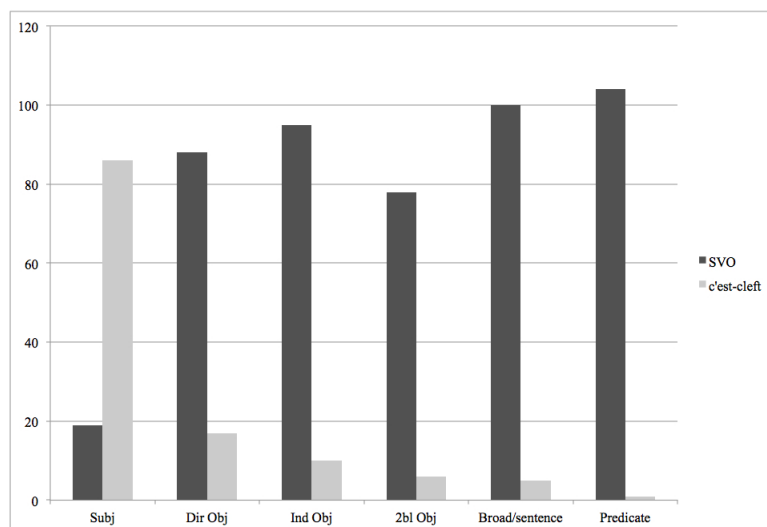
Figure 5.3: Total count per grammatical function in Neutral contexts
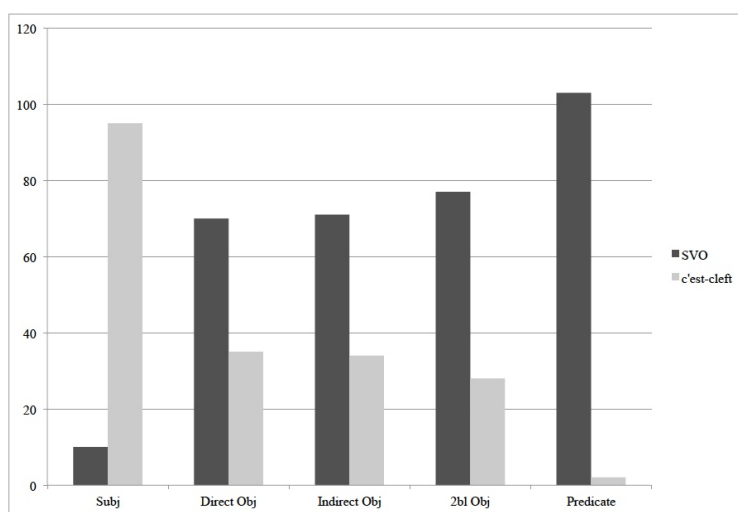
Figure 5.4: Total count per grammatical function in Contrastive contexts

ically ban subjects from sentence-initial position. In relation to the past literature, this result challenges the findings in Hupet and Tilmant (1990) where 0% of subjects occurred in a canonical form. It also challenges Lambrecht's claim that a cleft is a "compensatory

device" used because the preverbal position cannot receive main stress. This result, on the other hand, is consistent with accounts like Féry (2001) and Claire *et al.* (2004) that indicate French can mark focus on subjects in situ via prosody.

Regarding hypothesis 2, the results provide empirical evidence that there is no categorical one-to-one relationship between a given grammatical function and its focus realization: object foci and predicate foci are, for example, both realized via canonical sentences. Unlike the proposal of Lambrecht (1994), the results from this experiment do not exhibit three different focus structures associated with specific information-structural articulations. This does not preclude the existence of these structures, but it demonstrates that there is only limited systematicity in their occurrence with a grammatical category.

Finally, the results confirm hypothesis 3: there is a strong relationship between semantically stronger foci and the use of the cleft. Clefts are used both for corrective and neutral foci. However, a logistic regression model including contrastive/corrective context is significantly better than a reduced model (a model without the contrastive/corrective predictor), $\chi2(1) = 41.9$, p < 1e-10. An analysis of the full model (the model with the contrastive/corrective context predictor) showed that corrective/contrastive context significantly increased the log-odds of using a cleft (beta= 1.18, Wald = 6.18, p ≤ 0.0001). Across grammatical functions, the relative risk is between 1.2 and 3.2; in other words the probability of the cleft in the contrastive/corrective context is 1.2-3.2 times that of the neutral context. I also ran a logistic regression model that included interactions with *correctiveness.* Results show that such a model is not significantly better than a model without the interaction ($\chi2(4) = 4.8$, p = 0.31). Correctiveness has, however a slight effect on both indirect object (p=0.11) and double objects (Wald =1.71, p=0.084).

Focused subjects: Within the subject focus condition across contexts, the raw number count amounts to 181 clefts produced vs. 29 canonicals out of 210 sentences (86.2% vs. 13.8%). These numbers are in line with the widely assumed claim that lexical subjects must be marked via clefting (Lambrecht, 1994; Katz, 1997; Hamlaoui, 2008). The non-zero occurrence of subjects in canonicals is statistically insignificant in both contexts. Within neutral contexts, participants produced a total of 86 clefts and 19 canonicals (82% vs. 18%). In corrective contexts, participants produced a total of 95 clefts and 10 canonicals (90.5% vs. 9.5%). These numbers demonstrate that "context" plays no significant role in the realization of subject focus; the difference between the 86 vs. 93 clefts used is not significant. This result is in line with results found in past empirical studies by Hupet and Tilmant (1990) and Vion and Colas (1995). Indeed, these two studies showed that the grammatical function "subject" (or maybe rather the theta role "agent") dominates other factors such as the degree of conviction of the speaker (Hupet and Tilmant, 1990) and the number of alternatives considered by the addressee (Vion and Colas, 1995). In the case of subject-focus, clefting is pragmatically motivated: subjects occur in a marked position to avoid their default interpretation as topics (Lambrecht, 2000; Zimmerman *et al.*, 2010). Clefting is also prosodically motivated: subjects are merged into a position that can receive main stress (Hamlaoui, 2008). The following generalization accounts for the realization of subjects observed in the data:

(172) As a default, use a canonical for topical subjects and use a cleft to focus lexical subjects.

SUBJ=TOPIC: The grammatical subject of a sentence must be interpreted as a topic (it cannot be F-marked).

(173) As a default, main stress must be aligned with the right edge of an Intonational
Phrase.

ALIGN (I, R; HEAD(I), R) (HIR): Align the right boundary of every intonational
phrase with its head.

Across the 19 subject foci that were not clefted, I observed two consistent features:
the post focus sequence was deaccented and the subject carried some type of pitch move-
ment (either L or H) on the last syllable. This is illustrated in Figure (5.5) for the answer
triggered by the question "*Qui est-ce qui achète un journal au kiosque?/ Who bought the
newspaper at the stand?*". These features correlate with features observed by phonolo-
gists who studied the realization of subject focus in situ (Claire *et al.*, 2004), but also with
features found on clefted subject focus (Clech-darbon *et al.*, 1999; le Gac and Yoo, 2002).
This similarity is observed in the data, as seen when comparing the pattern of in situ vs.
clefted subjects in neutral contexts (Figure 5.5 vs. Figure 5.6).
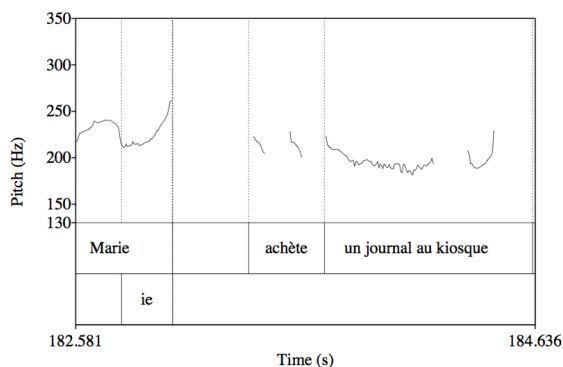
Figure 5.5: F0 curve of neutral subject realized in situ

I ran a preliminary analysis on the mean duration of the final vowel of the focused ele-
ment. The duration was 0.217 whereas the mean duration of the final vowel on non-foci

166
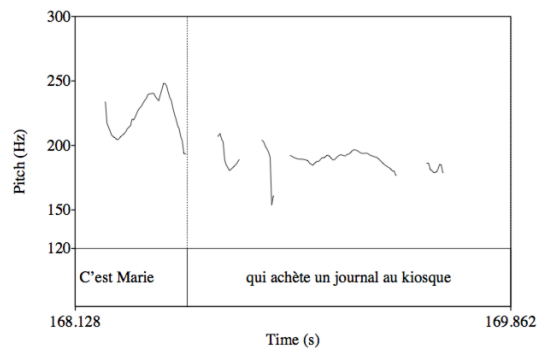
Figure 5.6: F0 curve of neutral subject realized in a cleft

was 0.111 and 0.102 for direct and indirect objects respectively. I also looked at how subject foci were *phrased* since some phonologists argue that French can mark focus on a subject by having that subject form its own intonational phrase (Féry, 2001; Claire *et al.*, 2004). A fair amount of variation was observed as not all subjects were realized in their own IP: Some participants had a very clear pause (> 0.3 seconds) between the focus subject and the rest of the sentence, others had no pause between S and V but produced a pause between SV and O. The mean duration of pauses between S and V is 0.232 and the mean duration between SV and O is 0.349. In cases where the S and the V were phrased together, the tonal pattern HL was still observed on the subject itself. Finally, some participants had no clear pause between constituents. The preliminary conclusion that can be drawn from these observations is that subject focus can be marked via some prosodic cues when remaining in situ, but is not necessarily always phrased separately.[12]

---

[12]Delais-Roussarie (1996) arguing that the length of the constituent directly influences the way it is phrased, the fact that all the grammatical subjects I used in the experiment are two syllable long may have very well influenced phrasing.

Focused objects (direct and indirect): In comparison with focused subjects that mostly appear in a syntactically marked structure (cleft), focused objects exhibit the inverse pattern, being realized in situ 77% of the time. Within the direct object condition across contexts, the raw numbers amount to 158 canonicals vs. 52 clefts out of 210 sentences (75.2% vs. 24.8%). This difference is statistically significant, showing that the grammatical function "object" is a good predictor for the focus strategy used: objects (either direct or indirect) are more likely to be realized in situ than in a cleft. This result is in line with previous studies (Hupet and Tilmant, 1990; Hamlaoui, 2008; Beyssade *et al.*, 2011) that show a strong preference for complements to be realized via prosody in situ. One interesting point that would need further investigation is that, out of the 52 clefts realized in the focused D.O condition, across contexts, only 9 of them occurred with an inanimate object (= 17.3%). In the experimental stimuli, there were four D.O; two were inanimate ('*journal*' and '*médaille*'), and two were animate ('*Tarzan* and '*le fils*'). Participants consistently preferred to cleft the animate D.O (43 tokens, or 82.7%).

Within the indirect object condition, the same trend is observed: overall, 166 sentences are canonicals vs. 44 clefts (79% vs. 21%). This tendency is easily explained by phonology: given that (indirect) objects appear canonically in the rightmost position of the clause where main stress is assigned in unmarked cases, (indirect) objects do not need to be moved in a different syntactic position to receive prosodic prominence.

In the cases of objects, as opposed to subjects, the context has an effect on the strategy used to mark focus. Direct objects occurring in a corrective context are more often clefted than ones occurring in a neutral context (16% of clefts in neutral contexts vs. 33% of clefts in corrective contexts), and indirect objects occurring in a corrective con-

168

texts are clefted significantly more often (9.5% of clefts in neutral contexts vs. 32.3% in corrective contexts). The use of a special pragmatic context triggers the use of a special focus marking. Raw counts are summarized in (5.3). However, when running a logistic regression, results turn out to be statistically significant only for indirect objects (p=0.11). For direct objects, the corrective context does not significantly increase the likelihood of using a cleft (p=0.54). This result parallels results found in Vion and Colas (1995). In their study, the researchers tested speakers' preferences on the marking of contrastive focus in French depending on its grammatical function. They found that when the focus was a non-subject, participants preferred focal accent marking over clefting.

	Neutral	Corrective
Direct objects	88 SVO - 17 clefts	70 SVO - 35 clefts
Indirect objects	95 SVO - 10 clefts	71 SVO - 34 clefts

Table 5.3: Raw counts of clefts vs. SVO usage for Direct and Indirect objects per context

One interesting piece of evidence stemming from this experiment is that, unlike in other Romance languages such as Spanish and Italian, the word order $SViO[dO]_F$ was not produced by participants. This structure is rarely discussed in previous accounts on French. To my knowledge, only Hamlaoui (2008) mentions that French can resort to heavy NP shift, yet she could not get any categorical judgments on the preferred word order (174a or 174c). In his study on Argentinian Spanish, Gabriel (2010) shows that under the direct object focus condition, 83% of sentences produced have the word order $SV[dO]_FiO$ (174b) whereas 17% display the direct object to the right of the indirect object. In (174d), the direct object is not dislocated, but appears in the same Intonational Phrase as the indirect object.

(174) What does Mary give to her brother?

    a. Marie donne UN JOURNAL   à son frère.
       Marie gives   a   newspaper to her brother.

    b. María le da   UN DIARIO   a su hermano.
       Marie it gives a   newspaper to her brother.

       'Marie gives A NEWSPAPER to her brother.'

    c. ? Marie donne à son frère   UN   JOURNAL.
       ? Marie gives   to her brother A NEWSPAPER.

    d. María le da   a su hermano UN DIARIO.
       Marie it gives to her brother   a   newspaper.

       'Marie gives to her brother A NEWSPAPER.'

Pulling from the results observed in the data, the following principles summarize the realization of complements:

(175) As a default, leave focal complements in situ.

(176) In special pragmatic contexts, where there is a contrast between the speaker and the hearer's expectations, complements occur in a cleft.

This last principle calls on a constraint that I already proposed in Chapter 4 when I analyzed the distribution of particles in pivot position, and repeated in (177).

(177) CONTRAST: Pivots are contrastive or corrective & in situ elements are not.

Prosodically speaking, I observed some variation on the realization of objects. For example, out of 88 canonical sentences realized in neutral context, I found that there was

a clear[13] marking in 76 cases (86.4%). In these cases, the object was not deaccented, but some accent was consistently observed on either syllable of the focused object. Moreover, the topical subject was also realized with an accent. A couple of representative examples of theses cases are illustrated by the F0 contour in the two figures below.
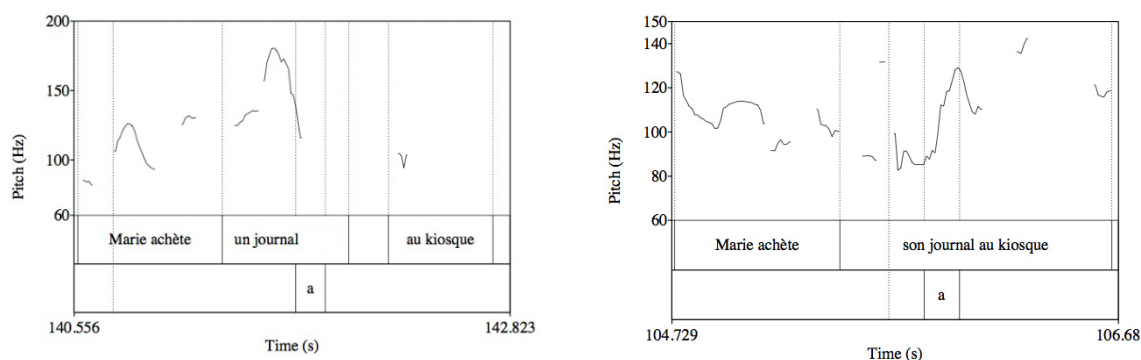

Figure 5.7: F0 curve of neutral direct object realized in situ with accent.

In the other 12 cases (13.6%), the object was realized with a descending intonation uncharacteristic of focus constituents. The subject still carried some kind of prosodic movement (see Figure 5.8 below). When the object was clefted, on the other hand, it was always accented (see Fig. 5.9).

---

[13]The criteria for considering an object clearly marked was to observe either a vowel lengthening or a pitch movement on the object. Such measures were instrumentally studied using Pratt to track f0 and intensity.
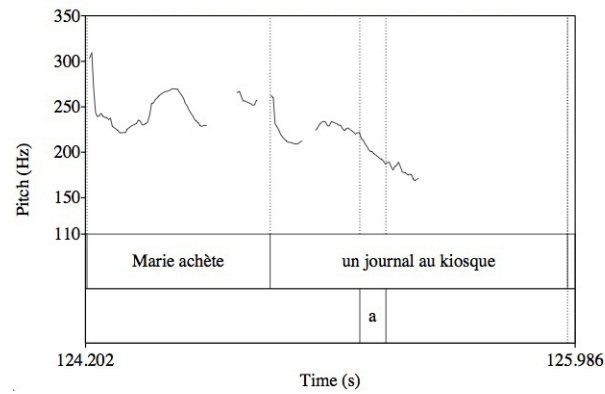
Figure 5.8: F0 curve of neutral object realized in situ without accent.
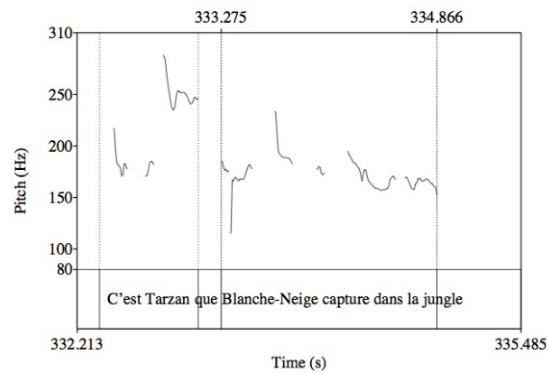
Figure 5.9: F0 curve of neutral object realized in a cleft with accent.

The indirect objects were also accented when focussed vs. deaccented when topics. Compare the two spectrograms in Fig.5.10, where the left one is focussed and the right one is not.
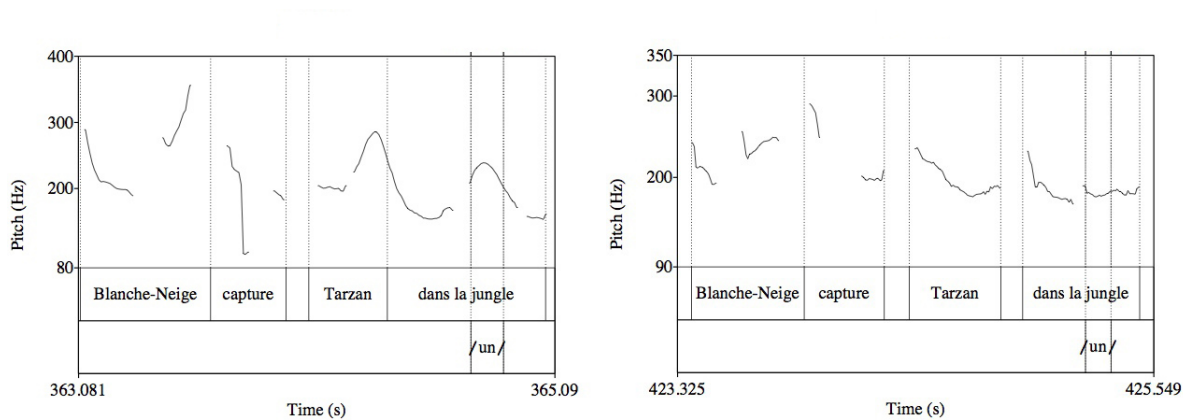
Figure 5.10: F0 curve of indirect object realized in situ with and without an accent.

Double focus: Another condition in the experiment was to test how speakers produced a sentence which answered a question with two open variables. This condition included focus on either the subject and direct object (178), or the direct object and the indirect object (179).

(178) Qui a   acheté quoi au   kiosque?
       Who has bought what at-the stand?
       'Who bought what at the stand?'

(179) Marie a   acheté quoi, où?
       Marie has bought what, where?
       'Marie bought what where?'

The results for this condition were quite straightforward since out of 168 sentences judged, twenty-two were clefts and 146 were canonicals. Looking at the twenty-two clefts, the majority of them occurred in a corrective context (16 vs. 6 in neutral context). When using a cleft, participants always clefted the focused element that was in higher syntactic

173

position. Thus, in a double focus question like (178) that triggered a subject focus and a D.O focus, the grammatical subject was clefted in the answer, and in a question like (179) that triggered a D.O focus and a I.O focus, the grammatical direct object was clefted.

Focused predicate: In the experiment, the results for predicate focus were categorical: canonical sentences were produced in 100% of responses.

Focused sentence: In the experiment, the results for broad focus were also very clear since they were realized in canonical sentences 100% of the time. This realization is simply explained by the fact that the focus domain corresponds to the right edge of the Intonational Phrase, which in turn corresponds with the right edge of the clause. Under this condition, French is similar to Spanish, English and Italian that all realize broad focus in a canonical sentence with the main accent occurring rightmost, at the right edge of the clause.

(180)   Q: What happened?

    a.   [Mary gave the newspaper to her brother]$_{IP}$

    b.   [Marie a donné le journal à son frère]$_{IP}$

    c.   [María dio el periódico a su hermano]$_{IP}$

    d.   [Maria diede il giornale a suo fratello]$_{IP}$

5.5   Discussion and Conclusion

The first point that I wish to discuss in relation to this chapter is the clear asymmetry between subject focus (SF) and non-subject focus (NSF) marking that emerged from the experimental results. Such an asymmetry is interesting because it is not exclusive to

French, but is in fact observed in various languages. Zimmerman *et al.* (2010) discuss such a phenomenon in West African languages, and find that there is not only a *structural asymmetry* where SF are marked differently from NSF, but also a *marking asymmetry* where NSF need not be marked whereas SF are. In the literature on French, the structural asymmetry between SF and NSF has been discussed by scholars such as Vion and Colas (1995) and Hamlaoui (2008) who showed that only SF is required to be marked via clefting. The present work provides experimental support to their claim as the results show that, by default, SF are realized in a cleft and NSF are realized in a canonical structure. My work provided further evidence showing that certain pragmatic uses of focus like *contrast* can inverse this default realization while still preserving the asymmetry (SF occurred in a canonical and NSF in a cleft). In line with Zimmermann's work, I conclude there is a *structural asymmetry* in French. Evidence for a *marking asymmetry* is a little bit more controversial. Zimmerman *et al.* (2010) argue that, in West African languages, NSF cannot or need not be marked syntactically, whereas SF must be. As a matter of fact, Zimmermann even argues that in Hausa, in situ NSF does not need to be marked prosodically either. French shows evidence for the former claim since NSF is not required to undergo any syntactic reordering to be interpreted as a focus. On the other hand, the preliminary prosodic analysis introduced in this chapter, but also the abundant body of work on the phonology of French focus marking seems to challenge Zimmerman's latter claim: NSF are marked via intonational movement in more than 80% of the time. It is also hard to believe that none of the prosodic correlated of focus, including phrasing, vowel length and intensity, would not be marked in any distinctive way when focused. In further work, my

175

goal is to provide a more detailed prosodic analysis for NSF.[14]

This asymmetry also makes it harder to classify French into a typology like the one proposed by Büring (2009). Borrowing Büring's terminology, French seems to pattern like other *Edge Languages* when it comes to SF marking: Focus is marked by non-standard constituent order, with the focus in left- or right-peripheral position. Similarly to other Romance languages like Spanish and Italian, French exhibits a strong preference for realizing SF in a designated syntactic position. For French, this position is the cleft pivot position, while for Spanish and Italian, this position is sentence final. Note that this position is not the only position available for focus in French - existential pivot position is another one - but this position is restricted to hosting focus constituents. Realizing focus in that position allows the language to minimize the material between the focus and the relevant prosodic phrase edge. When marking NSF, French behaves more like *Boundary Languages*: Focus is marked by insertion of a prosodic phrase boundary to the left or right of the focus. While French has a basic (S) (VO) or (S) (V) (O) prosodic pattern, this pattern can change in response to focus (among other factors, another one being weight and length of the constituent). Thus, Claire *et al.* (2004) show that the prosodic pattern for NSF differs from the one for sentence-focus by the presence of a low boundary tone L% positioned to the right of the object and by having the object realized in its own phrase. Finally, as mentioned by Büring for Boundary Languages, French shows that post-focal constituents are generally dephrased, therefore deleting post-focal PhPs.

---

[14]One drawback of the current data is that it includes a lot of words within the stimuli that do not comprise sibilants or voiced elements, which are easier to analyze when looking at a pitch track. I would also modify the stimuli to include focus elements that have more than two syllables in order to get a better analysis of intonational patterns on the focus word.

The second point I wish to consider is the variation and under-determination in focus discussed explored in Zimmerman and Onea (2011). In their paper, the authors argue that focus being a universal category at the level of information structure, a strict one-to-one relationship between focus and its grammatical realization is not expected. As a consequence, context resolution is required in many, if not all languages. One of the examples where underspecification is observed is in English, where a single prosodic pattern can be mapped onto various information-structural structures, illustrated in (181), where the single nuclear pitch accent occurs on '*bats*'.

(181)   A: Peter bought a book about BATS.

      Q1: What did Peter buy?   Triggering object-focus

      Q2: What did Peter do?   Triggering predicate-focus

      Q3: What happened?   Triggering sentence-focus

According to Zimmerman and Onea, all questions in Q1-Q3 can be appropriate QUDs triggering the answer in A, and which focus structure is expressed is subject to contextual resolution. In French, things do not seem as free, first because the language clearly makes use of clefting for subject-focus, but also because phrasing is theoretically changed in response to focussing. Looking only at main stress, the translation of (181)A in French in (182) bears the same under determinacy. However, phonologists would argue that, at the IP level, Q1 elicits an answer where the object is realized in a separate phrase, whereas Q2 elicits an answer where the verb will either form its own phrase or will be realized with the object. Finally, Q3 elicits an answer where there is a single IP which boundary is found sentence final. Overlooking this prosodic level seems to make wrong predictions.

(182) A: Pierre a acheté un livre sur les CHAUVE-SOURIS.

Q1: Qu'est-ce-que Pierre a acheté?   Triggering object-focus

Q2: Qu'est-ce-que Pierre a fait?   Triggering predicate-focus

Q3: Qu'est-ce-qu'il c'est passé?   Triggering sentence-focus

A further relevant experiment to run would be a perception experiment that would test whether or not speakers can consistently match intonational patterns with the correct focus structure. What I observed in the experimental results is a non-systematic one-to-one relationship between grammatical focus and its realization. Despite a fair amount of variation in the data, the preference for clefting SF and leaving NSF in situ can be considered as a relationship between focus and its realization. French is consistent in the fact that focus constituents must always carry prosodic prominence or occur in pivot position. However, focus is not always the only factor that plays a decisive role in the system of focus realization. Indeed, all along this work, I have talked about the importance of other pragmatic factors such as contrast.

The last point I want to discuss is related to these pragmatic factors. It concerns the one-to-one relationship between the grammatical realization of focus and particular types of focus - *informational* and *contrastive* focus - discussed in Lambrecht (1994), Kiss (1998), Kratzer and Selkirk (2007) and Zimmerman and Onea. I have already talked about the fact that some scholars argue these two types of focus are semantically different, and must be systematically realized in different ways. Yet, the experimental results in this work proves that a categorical one-to-one relationship between focus realization and pragmatically different uses of focus is not warranted. There is however, a preference for

using a cleft when a contrast occurs.[15] Importantly, clefting occurs in addition to prosodic marking on the focus constituent. Therefore, it is logical that clefting does something else that a prosodically marked canonical sentence alone cannot do. In this work, I have argued that the expression of pragmatic effects such as contrast, but also exhaustivity and correction, is preferred via clefting, superimposed on the prosodic pattern of the regular canonical sentence. When expressed in canonical sentences, these pragmatic effects require extra strategies from the language. Contrast and correction are marked via a special accent known as the *accent d'insistance* in the French literature, an accent on the first syllable of the word constituting the contrast or correction, which is characterized by being longer, more intense and higher than non-accented syllables. In the experiment, many participants repeated the wrong information before correcting it, along with particles expressing the contradiction, as illustrated in (183).

(183)   Q: Regarde, on dirait que Marie donne le journal à sa fille, non?
      Q: Look, one look-like that Marie gives the newspaper to her daughter, no?

      'Q: Look, it seems that Marie is giving the newspaper to her daughter, no?'

      A: Mais non, Marie donne le journal À SON FILS, pas à sa fille.
      A: But no, Marie gives the newspaper to her son, not to her daughter.

      'A: No, Marie is giving the newspaper to HER SON, not to her daughter.'

Summing up, special pragmatic uses of focus are frequently expressed by extra means in a language in addition to the regular focus marking strategy. One interesting

---

[15] *Contrast*, here, is defined following Hupet and Tilmant (1990) and Zimmerman and Onea (2011), in terms of the speaker's estimation of the hearer's expectations regarding likely and unlikely updates of the common ground.

case that should be analyzed in further work is the difference - if any - between subject-focus in pivot position in neutral/informational context vs. contrastive/corrective context. Indeed, since subjects are mainly clefted, I would expect to find a difference in prosody, or at least in realization somehow, between the special and the non-special pragmatic use of that focus.

The data reported on in this chapter show the following tendencies in focus realization in French. In the next chapter, I account for the data discussed so far in this work by proposing a model based on stochastic optimality theory. This model correctly predicts and explains why (i) subject foci are mostly realized in a *c'est*-cleft, (ii) non-subject foci are often realized in situ, and (iii) these generalizations are reversed in special contexts.

Chapter 6

A Stochastic OT analysis of focus marking in French

The goal of this chapter is to develop a predictive model that explains the alternation cleft/canonical sentence in the marking of focus in French. I propose to use a Stochastic Optimality Theory (StOT) framework to clarify the interactions of syntax, prosody and pragmatics in this alternation. The free variation observed in the production experiment motivates the use of such a framework, since by definition, StOT accounts for variation by allowing a ranking of constraints to be "perturbed" by a normally distributed "evaluation noise" (Boersma and Hayes, 2001).

6.1 Introduction

Cross-linguistically, the requirements on syntactic and phonological well formedness often lead to conflict. In the case of focus marking, this conflict is observed when the optimal syntactic position for focus does not match up with the optimal prosodic position for main stress. It is therefore common for languages to favor one requirement at the expense of the other. Yet, different languages favor different requirements. For example, in the case of subject focus, English displays canonical constituent order with the subject receiving main stress preverbally (184a), whereas Italian moves the subject constituent rightmost where main stress occurs (184b). French employs yet another strategy which is

to add syntactic material in order to merge the subject into a position where it can receive main stress (184c). Interestingly, while all these languages differ in how they mark subject focus, they all exhibit a canonical pattern where main stress occurs sentence finally when marking non-subject focus (185). An analysis of focus realization must account for this puzzle: constraints must interact in a way that will favor syntactic movement in some contexts and disfavor it in others.

(184) Context: Who laughed?

    a. JOHN laughed.

      # John LAUGHED.

    b. Ha riso GIANNI.

      # Gianni ha RISO.

    c. C'est JEAN qui a ri.

      # C'est Jean qui a RI.

(185) Context: What happened?

    a. John LAUGHED.

    b. Gianni ha RISO.

    c. Jean a RI.

6.2 The challenges

Looking more specifically at French, the empirical data collected in the production experiment (Chapter 5) and the data gathered from the two corpus studies present some challenge for previous accounts of focus marking. The first challenge concerns the

realization of subject focus. It is almost formulaic in studies on French focus to say that lexical subject focus cannot appear in preverbal position, i.e. in its canonical position, but must appear in a *c'est*-cleft (Lambrecht, 1994; Katz, 1997). And yet, phonologists including Jun and Fougeron (2000), Féry (2001), Claire *et al.* (2004) and Delais-Roussarie and Post (2008), have all argued that subject focus can be realized in situ via phrasing (forming its own intonational phase), or via pitch accent (a LH or HL tonal pattern at right edge of the phonological phrase formed by the focused element). There is a lack of consensus on how subject focus is realized in the language. The data shows that in fact, both accounts are substantiated. Clefting is used as the most frequent strategy to focus a subject (the production results demonstrate that a cleft is used 86% of the time), but no categorical ban on preverbal subject focus exists (free variation occurs 15% of the time). Furthermore, there seems to be cases where a preverbal subject is required, such as quantified subjects or modified subjects. Consider (186).

(186) Context:

Qui est-ce qui est allé à la manifestation le week-end dernier?
Who is-it   who has went at the demonstration the weekend last?

'Who went to the demonstration last weekend?'

  a. Beaucoup d' étudiants y   sont allés.
     Lots   of students there be.3pl went.
     'Lots of students went there.'

  b. ? Ce sont   beaucoup d' étudiants qui y   sont allés.
     ? It be.3pl lots   of students who there be.3pl went.
     '? It's a lot of students who went there'

c. Beaucoup d' étudiant.
   Lots   of students.

   'Lots of students.'

d. ? Ce sont   beaucoup d' étudiants.
   ? It   be.3pl lots   of students.

   '? It's a lot of students.'

A version that includes a clefted subject is pragmatically odd and is unlikely to occur as a direct answer to the QUD, but a canonical sentence is perfectly acceptable (186a). Even when considering an elliptical version of the answer, a non-clefted answer (186c) is more acceptable than a clefted one (186d). A fully predictive model will have to account for the free variation and the variation triggered in stronger interpretations.

The second challenge concerns non-subject focus. Here again, there is no clear consensus in the prior literature. Although Lambrecht's account of focus does not differentiate between subjects and non-subjects (arguments regardless of their grammatical function are argued to be clefted), recent studies have claimed that non-subjects are preferably realized in situ (Hamlaoui, 2008; Beyssade *et al.*, 2011). But, the empirical data collected in my production experiment reconciles the two sides since it shows that non-subjects are realized in situ more frequently than in a cleft, but that there was a fair amount of free variation. Moreover, results proved that the context could greatly impact that use of the cleft with non-subjects: in a corrective context, non-subjects were more likely to be clefted. These empirical observations must also be taken into account and explained in a predictive model.

The present chapter aims to provide a unified account for the non-random alternation cleft/canonical sentence in French which occurs when marking focus. The goal

is to explain both the strong preferences which constitute between 77 and 86% of the cases (subjects are clefted and non-subjects are not), but also for the variation observed and the special cases discussed constituting the remaining 14-23%. I choose to couch my account in a Stochastic Optimality Theory (StOT) framework, which is motivated by the variation observed in the cleft vs. canonical usage. The account is based on the interaction of several *floating constraints* on syntax, prosody and pragmatics that freely reorder to reflect the selection of different candidates from evaluation to evaluation. By positing a free ranking, my account will also give insights on what factors are the most important ones in governing the alternation cleft/canonical, and which ones play a lesser role in French. There are several advantages to analyzing the cleft/canonical alternation by using an OT framework. First and foremost, the account is unified and there is no need to posit the existence of various *c'est*-cleft constructions. Under my analysis, the differences observed are explained by the type of constraints that are being violated. Second, using this framework helps accounting for the free variation actually observed in the data which was mostly disregarded in previous studies (Katz, 1997; Hamlaoui, 2008). Using a Stochastic OT framework also predicts the variation observed across speakers. Finally, my analysis accounts for a wider range of data and interpretations of the *c'est*-cleft than previously studied.

6.3   Theoretical background

Optimality Theory (Prince and Smolensky, 1993) relies on the idea that the grammar of individual languages derives from a hierarchical ranking of universal, yet violable constraints on representational well-formedness of output form(s) given an input form.

The constraints are universal, i.e they are maximally general and capture the most elementary constraints on speech itself. Because of the universal nature of these constraints, they will be violated in individual languages depending on the way they are ranked, with some violations being more serious than others. In any given language, the set of constraints is therefore strictly ranked, and C_3 can never come to dominate C_2 or C_1. The optimal output is the candidate that violates the least of the highly ranked constraints compared to its competitors.

(187)   Strictly ranked constraints under Standard OT for language X:

$C_1 >> C_2 >> C_3$

The major drawback of this model is that, while it allows for crosslinguistic variation, it fails to account for variation within a single language. Indeed, it predicts that a candidate will emerge as the surface form. But empirical data show that variation does occur and complicates the picture. Boersma and Hayes (2001) propose a stochastic version of OT that remedies this problem by altering the Standard view in two major ways. First, instead of adopting an ordinal ranking as in (187), they adopt a ranking along a continuous scale (Figure 6.1).[1]

The second alteration comes into play at the stage of the evaluation of candidates for output forms. In StOT, every time a candidate is evaluated, the position of each constraint is "perturbed" by a random positive or negative value. Therefore, at the time of evaluation,

---
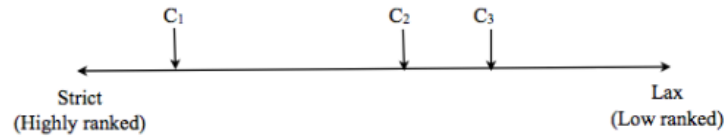
[1]adapted from Boersma and Hayes 2001, 47

Figure 6.1: Continuous scale.

the constraints are associated with a range of values rather than a single point. The value used at the time of evaluation is called the *selection point*. The value associated more permanently with the constraint is called its *ranking value*. Thus, each candidate has a range of selection points associated with its ranking value. These ranges are interpreted as normal probability distributions. If the ranges of, for example, C_1 and C_2 do not overlap, the more highly ranked constraint will strictly dominate the lower ranked constraint (Figure 6.2).
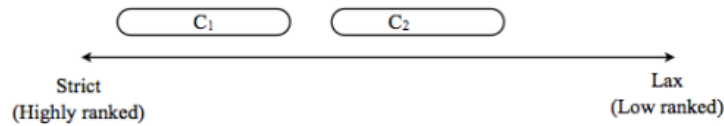
Figure 6.2: Continuous scale with non overlapping constraints.

If the constraints have overlapping selection points, they can be ranked freely. At the time of evaluation, the selection point can be selected within the ranges of both constraints. In a case where the candidates evaluated are in the upper part of C_2 and lower part of C_3, the resulting ranking will have C_2 dominating C_3 (Figure 3). As the distance between two crucially ranked constraints increases, their distributions overlap less and less, and categorical outputs arise.
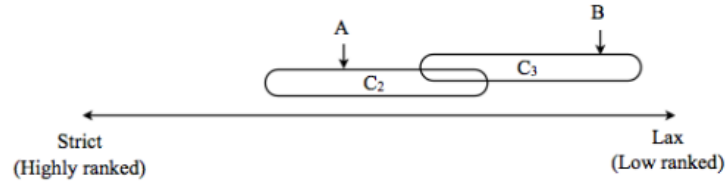
Figure 6.3: Continuous scale with C_2 overlapping C_3

In cases where the candidates evaluated are in the upper part of C_3 and the lower part of C_2, the opposite ranking will hold: C_3 will dominate C_2 (Figure 4). In that case, the two constraints are not too distant and variable outputs can be generated. Thus, the two opposite rankings in Figure (3) and (4) will lead to free variation in the sense that one underlying form can generate multiple outputs (instances of variability that do not arise because of differences in context or register). Now considering all three constraints, C_1, C_2 and C_3, at evaluation time, C_1 will outrank the two other constraints in 100% of the cases, and C_2 will outrank C_3 95% of the time , and C_3 will outrank C_2 5% of the time. Thus, in 95% of cases the ranking will be $C_1 > C_2 > C_3$, and will be $C_1 > C_3 > C_2$ in 5% of the cases.
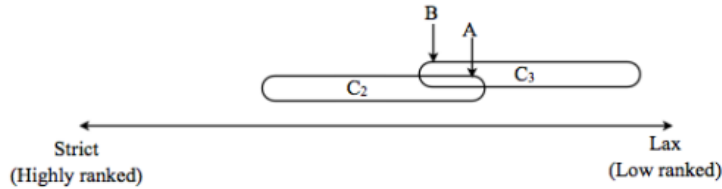

Figure 6.4: Continuous scale with C_3 overlapping C_2

188

6.4 Past OT accounts on focus marking

Up until a decade ago, OT was mainly reserved for phonology and its application to syntax and pragmatics was relatively sparse. Yet, OT proves to be a very useful tool to account for focus marking because this phenomenon is universal (all languages mark focus) but its realizations differ greatly from language to language.[2]

Concerning Romance languages, a major step is taken by Zubizarreta (1994, 1998) who argues for a correlation between focus and prosody. The author claims that, in Romance languages, main stress naturally falls sentence-finally, and that the position of focus must correspond to the position of main stress. Therefore, in most Romance languages, focused elements occur rightmost rather than in their canonical positions as illustrated for Spanish in (188) taken from Gabriel (2010).

(188) a. *Subject focus on "María"*: Compró el diario MARÍA.

b. *Object focus on "un diario"*: María le dio a su hermano UN DIARIO.

Translating these requirements into constraints, Zubizarreta argues that the position of the nuclear accent in an utterance is the result of the following two rules: the "Nuclear Stress Rule" as first introduced by Chomsky and Halle (1968) for English, and the "Focus Prominence Rule".

(189) a. NUCLEAR STRESS RULE (NSR): Given two sister categories C_i and C_j, if C_i and C_j are selectionally ordered, the one lower in the selectional ordering is more prominent.

---

[2]For an overview of how different languages mark focus, I turn the reader to Büring (2010b).

b. FOCUS PROMINENCE RULE (FPR): Given two sister categories C_i (marked [+F]) and C_j (marked [-F]), C_i is more prominent than C_j.

In Spanish and Italian, these rules seem to be rarely violated and prosody directly affects syntax: requirements on the position of stress outrank requirements on syntactic well-formedness since focused elements are generally realized rightmost in order to receive stress. In other words, prosodic well-formedness outranks syntax; the language will resort to moving the element outside of its canonical position in order to fulfill prosodic require-ments. On the other hand, focused elements in French behave differently and a subject focus can never be realized rightmost because the sentence-initial position can never be left empty. Therefore, while (188a) repeated below in (190a) is completely acceptable in Spanish, the corresponding example (190b) in French is ungrammatical.[3]

(190) a. Spanish - *Subject focus on "María"*:

Compró el diario MARÍA.

b. French - *Subject focus on "Marie"*:

∗a acheté un journal, MARIE.

Many subsequent studies have reanalyzed Zubizarreta's (1998) insight by couch-ing their analysis in Optimality Theory and positing a ranking of constraints on syntax

---

[3]In order to get the subject NP outside of preverbal canonical position and in rightmost position, French would resort to a right dislocation of the type "Elle a acheé un journal, MARIE". However, this sentence form is not a possible foci marking strategy in French because right dislocated elements are in fact often destressed and mark topics instead of foci. I refer the reader to de Cat (2007) for more information on the function of French dislocations.

and prosody. Concerning prosody, many studies propose to move away from a direct correspondence between main stress and focus, and instead discuss *alignment* to prosodic and morphological edges. For example, Selkirk (1996) and Truckenbrodt (1999) propose a set of constraints which requires the alignment of left or right edges of XPs with edges of phonological phrases φ (191).

(191)   Generalized Alignment constraints:

    a. ALIGN-XP,R: ALIGN (XP, R; φ, R) Align the right edge of XP to the right edge of φ.

    b. ALIGN-XP,L: ALIGN (XP, L; φ, L) Align the left edge of XP to the left edge of φ.

Truckenbrodt (1999, 228) ensures that all lexical items in a lexical projection are contained within a single phonological phrase by positing the constraint WRAP in (192). He ensures that each phonological phrase bears a phrasal stress by positing the constraint STRESSXP. This latter constraint is useful especially when the XP contains more than one element.

(192)   WRAP-XP: Each XP is contained inside a phonological phrase.

(193)   STRESS-XP: Each XP must contain a *phrasal stress*.

For example in Spanish, in (194), each lexical item is contained within a phonological phrase where it must be prominent. At the intonational level, the constraint on alignment requires the main stress to occur rightmost.

(194)   Prosodic Phrasing in Spanish:

(*)   Intonational Phrase (IP)

(*   ) (*) (*) (*)   Phonological/Accentual Phrase (PhP)

(*   ) (*) (*) (*)   Prosodic Word (PW)

Juan   le dió   un libro   a María.

'*John gave a book to Mary.*'

In his 2005 paper, Samek-Lodovici proposes an OT analysis for various rightmost languages such as English, Italian and French. He refines and expands on (191) by proposing a set of constraints in (195) that account for the position of heads in different prosodic categories above the phonological word.

(195)   a.   ALIGN (P, R; HEAD(P), R): Align the right boundary of every phonological phrase with its head.

b.   ALIGN (I, R; HEAD(I), R): Align the right boundary of every intonational phrase with its head.

Crucially, alignment constraints of this type allow gradient violation since one violation is inferred for each phrase boundary that occurs between the head and the right edge. In Spanish, where each prosodic word forms its own phonological phrase dominated by a single intonational phrase, and foci typically occur rightmost where main stress is assigned, Büring and Gutiérrez-Bravo (2001) assume that a constraint like (195b) must be undominated. In English, where a subject-focus is realized in canonical position, this constraint yields two violations because two potential head positions separate the head from the right edge of the IP boundary (196).

(196) ALIGN (I, R; HEAD(I), R) violated twice:

Context: Who kissed Mary?

(*   ) IP level

(*   ) (*   ) (*   ) PhP level

JOHN   kissed   Mary.

Thus, this constraint must be ranked low to allow the optimal candidate to be of the form SVO. Syntactic constraints in (197) will be responsible for enforcing canonical sentences as the norm (enforce faithfulness to SVO word order) and infering violations for each departure from that norm.

(197) Syntactic constraints:

    a. EPP: Clauses have subjects.[4]

    b. STAY: No traces.

    c. FAITHSYN(TAX): Do not insert structure not present in the input.[5]

The constraint EPP is violated by sentences that lack an overt preverbal subject. The constraint STAY, which demands that constituents do not move from their default position, will be highly ranked in languages like English that favor a canonical word order, but will be ranked lower in languages like Spanish and Italian that move narrow foci rightward. A similar exigency will arise for EPP, since, as mentioned earlier, English requires clauses to have overt subject whereas Spanish and Italian do not. The constraint FAITHSYN will be

---

[4]Constraint cited from Grimshaw (1997) and Samek-Lodovici (2005a).
[5]Constraint cited from Grimshaw (1997, 376).

ranked low in a language that uses a lot of cleft constructions such as French in order to allow these constructions to emerge as favorite candidates.

Finally, all OT studies on focus marking have a constraint on the stress-focus correspondence. Indeed, one of the properties of a focused constituent, as noted by Jackendoff (1972), is that it is more prominent than a non-focused constituent. The constraint widely adopted in the literature is known as STRESSFOCUS (Samek-Lodovici, 2005a; Hamlaoui, 2009; Gabriel, 2010).[6]

(198)   STRESSFOCUS (SF): For any XP_F and YP in the focus domain of XP_F, XP_F is prosodically more prominent than YP.

In English, focus is realized by matching prosodic prominence with the position of the focused element even if that means moving prominence to a non-rightmost position. This behavior is explained by the ranking of a prosodic, a syntactic and a prominence constraint, as illustrated in Tableau 6.1 adapted from Zerbian (2006).

	focus = subject	STRESSFOCUS	EPP	ALIGN (I, R; HEAD(I), R)
a. ☞	JOHN laughed)$_{IP}$			*
b.	laughed$_i$ JOHN t_i)$_{IP}$		*!	
c.	John LAUGHED)$_{IP}$	*!		

Table 6.1: Subject focus in English

In this tableau, candidate (a) is the favorite output because it bears a pitch accent on the focused element in subject position. The position of the pitch accent on the preverbal

---

[6]This constraint corresponds to the FOCUSPROMINENCE constraint adopted by Büring and Gutiérrez-Bravo (2001) and Zimmermann (2006).
FOCUSPROMINENCE (FP): A focus constituent X must be realized on or next to X in a clause S iff S also contains non-focused material

subject violates the low ranked constraint ALIGN (I, R; HEAD(I), R) which demands that main stress be aligned rightmost. Candidate (b) shows deviation from the canonical word order by having the subject moved to sentence final position, thus violating EPP. However, the constraint on prosodic prominence position is not violated since the movement placed the focused element in the rightmost position where main stress is assigned by default. EPP being ranked higher than ALIGN (I, R; HEAD(I), R), the violation incurred to (b) is fatal. Finally, candidate (c) shows canonical word order with an overt subject in SpecIP, therefore respecting EPP, and prosodic prominence occurring rightmost, therefore respecting ALIGN (I, R; HEAD(I), R). However, this candidate is ruled out because it violates the highest ranked constraint STRESSFOCUS because it is not the Phonological Phrase that contains the focused constituent that projects the head of the Intonation Phrase. Put differently, the focused element does not correspond to the element marked for prominence.

The ranking of constraints for English differs from the one that explains focus marking in Italian and Spanish. Indeed, in these two languages, the requirement on the pitch accent occurring rightmost is more strict. Tableau 6.2 illustrates that in Italian, the position of the focused element must match the rightmost prominent position even if this means that the canonical word order SVO is not maintained.

	focus = subject	STRESSFOCUS	ALIGN (I, R; HEAD(I), R)	EPP
a. ☞	Ha riso GIANNI)$_{IP}$			*
b.	GIANNI ha riso)$_{IP}$		*!	
c.	Gianni ha RISO)$_{IP}$	*!		

Table 6.2: Subject focus in Italian

Candidate (a) is the optimal candidate because, despite violating EPP by having the subject occurring clause-final, it fulfills the alignment constraint of main stress being rightmost. Candidate (b) is ruled out because main stress is not rightmost even though it respects canonical word order with an overt subject. Similar to English though, a candidate like (c) which does not convey prominence on the correct element is the worst candidate.

6.5 Accounting for focus realization in French

Turning to French, there are, to my knowledge, only two studies that propose to use optimality theory to account for the cleft/canonical alternation in marking focus: Samek-Lodovici (2005a) and Hamlaoui (2008), the latter following most of the constraints proposed by the former, both being couched in standard OT. However, I believe these two accounts make some incorrect predictions, and generally fail to account for the variation actually observed in the data. The first major incorrect prediction is made in Samek-Lodovici (2005a), where the author argues that subject foci are always realized in situ via prosody. Surprisingly, the author does not even discuss clefting as a possible focus marking strategy for French. However, my work (and previous literature on French) shows that clefts are used in most cases to mark subject focus. A second problem arises from the treatment of direct objects, which are predicted to move in clause-final position. Despite not being infelicitous, this realization was not produced in experimental settings, and was not observed in naturally-occurring data. More work would need to be done to investigate the appropriateness of such movement.

Another limitation occurs in Hamlaoui (2008), where the strict ranking of the con-

straints predicts that subjects must always be clefted when answering a "Qui"-question. Her account predicts that subject foci can never be realized via prosody since the possibility for subjects to form their own IP is not considered. Examples where the subject contains a modifier or a quantifier are also left undiscussed. Moreover, object foci are predicted to always occur in situ except in contrastive contexts, where they are predicted to always occur in a cleft. This account seems slightly too categorical and the data I collected shows that, even though the grammatical function and the context have an effect on the use of the cleft, the cleft is not categorically associated with subjecthood and contrast.

6.5.1 The input

In OT-phonology, there is little debate about how the input should be structured. The input is taken to be the underlying phonological representation of a word, and defines the set of candidates that compete for optimality and illustrate faithfulness constraints from which the output should not deviate. Defining the input in OT-syntax and pragmatics proves to be more complicated, because, according to Heck *et al.* (2002), syntax is information preserving while phonology is not. Heck *et al.* (2002, 354) claim that none of the existing notions of input in optimality theoretic syntax is suited to fulfill the major task of defining the candidate set. Different proposals are made regarding how the input should be structured: scholars argue about whether it should not be structured, partially structured (Grimshaw, 1997) or highly structured (Legendre *et al.*, 1998). A *partially structured* input as proposed by Grimshaw (1997, 376) includes a lexical head plus its argument structure and an assignment of lexical heads to its arguments. It also specifies tense and

aspect. Thus, the most economic syntactic structure represents the optimal candidate. A *highly structured* input as proposed by Legendre *et al.* (1998) consists of syntactic structure and the logical form of the utterance. Thus, in accounting for wh-questions, the authors take the target scope positions of wh-phrases to be part of the input. In my model, I follow Grimshaw (1997) in assuming the input is syntactically structured with a verb and its argument. Therefore, the input consists of the canonically ordered elements of the sentence which are encoded for their grammatical roles. I further follow Legendre *et al.* (1998) in the sense that the input represents a specified logical form of the utterance. Regarding focus marking, the logical form of a sentence includes the informational status of the elements.

6.5.2 The candidates

In OT, the function GEN takes the input and generates a set of possible candidates. Thus, the assumptions made for the input have direct repercussions for the set of candidates taken into account. Here again, there are major differences between the candidates in OT-phonology and in OT-syntax. The *richness of base* in OT-phonology assumes that there are no language specific restrictions on the input, resulting in an infinite number of candidates or possible realizations of that input. But in OT-syntax, the input is structured in differently. Given the assumptions made concerning the input, the candidate set generated by GEN contains candidates that diverge from the input both by their syntactic structure and their logical form (i.e. the informational status of the element).

6.5.3 The constraints

The present account fits with past OT analyses of focus realization such as Büring and Gutiérrez-Bravo (2001), Samek-Lodovici (2005a), Zimmermann (2006), and Hamlaoui (2008) as I posit a set of constraints on focus prominence, syntactic faithfulness and prosodic markedness which derive the relevant empirical generalizations. Some of these constraints, by being re-ranked, can also account for cross-linguistic differences. The set of constraints employed in my account is given in (199-summaryofconstr3):

Set of constraints:

(199) Interface constraint

STRESSFOCUS: A focused constituent must receive highest prominence in its focus domain. (*One violation if the main stress at the IP level does not correspond to the Focus marked nodes*)

(200) Prosodic constraints

a. ALIGN (P, R; HEAD(P), R) (ALIGNHIR): Align the right boundary of every intonational phrase with its head. (*One violation inferred for each possible head position occurring between the right edge of the IP and the actual head*)

b. *IP: Avoid IP boundaries. (*One violation for each IP boundary added*)

(201) Syntactic constraints

a. EPP: Highest A-specifier must be filled.[7] (*One violation for each Highest A-specifier position left empty*)

---

[7]The exact wording of this constraint is taken from Zimmermann (2006).

b. FAITHSYNTAX: Do not add syntactic material not present in the input. (*One violation for each sentence that does not follow the canonical order*)

c. STAY: Avoid traces. (*One violation for each trace*)

(202) Pragmatic constraints

a. *SUBJECTFOCUS: The grammatical subject of a sentence must not contain a focus. (*One violation for each subject which contains a focus*)

b. FOCCLEFT: A focus element must be clefted.

   i. FOCCLEFT$_{exhaustivity}$: A focus with an exhaustive reading must be clefted.

   ii. FOCCLEFT$_{contrast}$: A focus with a contrastive reading must be clefted.

6.5.3.1 The StressFocus constraint

The first and most important constraint to posit concerns the relationship between context, focus and prominence. In the literature on focus, there is a general consensus that the focus of a sentence must correspond to the open variable instantiated by the wh- in the preceding (explicit) QUD, and that in turn this element must receive highest prominence. In other words, the focus in the input selected by the preceding QUD must correspond to the focus in the output. Let's first consider an instance of a *narrow* focus, i.e. where the focus domain is a single element.

(203) Context: Who is baking a cake ?

a. JEAN cuisine un gâteau.
   Jean bakes   a   cake.
   'JEAN *is baking a cake.*'

200

b. * Jean cuisine UN GÂTEAU.

   * *'Jean is baking* A CAKE.'

The QUD in (203) instantiates an open variable x and demands answers of the type "*x* is baking a cake". The example (203b) is infelicitous because it is incoherent with the preceding QUD since it gives prominence to an element that's not providing a value for the open variable *x*. Instead, (203b) is giving prominence to an element already identified in the QUD. (203b) would be appropriate if the QUD was "What is Jean baking?", where appropriate answers must be of the form "Jean is baking *y*". The requirement for the mapping of focus to prominence is ensured by positing the constraint SF:

(204) STRESSFOCUS (SF): A focused constituent must receive highest prominence in its focus domain.

Similarly for broader focus domains (often triggered by a less specific QUD like "What happened?" or "What did you do?"), SF will rule out candidate (205a), which is infelicitous because it falls out of the focused VP, but will favor candidates (205b-205d) because they correctly answer the QUD. Generally speaking, SF will penalize candidates that fail to mark the appropriate focus element more prominently than a non-focus, therefore prohibiting non-identity between the focus in the input and the focus in the output, and prohibiting inconsistencies in the discourse.

(205) QUD: Qu'est-ce que Marie a fait?/ *What did Marie do*?

   a. *MARIE a   mangé du   chocolat.
      MARIE has eaten   some chocolate.
      'MARIE ate some chocolate.'

b. Marie A MANGÉ du chocolat.
Marie HAS EATEN some chocolate.

'Marie ATE some chocolate.'

c. Marie a mangé DU CHOCOLAT.
Marie has eaten SOME CHOCOLATE.

'Marie ate SOME CHOCOLATE.'

d. Marie, elle A MANGÉ DU CHOCOLAT.
Marie, she HAS EATEN SOME CHOCOLATE.

'Marie, she ATE SOME CHOCOLATE.'

The SF constraint is widely used in the literature on focus marking to account for the fact that focus constituents are more prominent than non-focus ones. In that sense, the constraint in (204) is similar to the Focus Prominence constraint proposed by Zubizarreta (1998), the Focus constraint proposed in (Truckenbrodt, 1995, 11), and the requirement in Schwarschild (1999) and Büring (2001b) that focus be accented. This constraint is undominated as it is the most important one in the hierarchy: no matter how focus is realized it is essential that it'd be signaled on the correct/relevant element.

Yet, SF has its limitations and cannot, for example, govern where exactly prominence falls when the focus domain is wider than a single element. SF also doesn't make any predictions on what sentence form is preferred between a canonical sentence bearing main stress or a cleft sentence. Indeed, in French, prominence can be achieved either via prosodic means, but also syntactic ones like clefting or dislocating. For example, in the case of a predicate-focus sentence, SF fails to predict whether candidate (205d) is more appropriate than candidate (205c). Therefore, we must consider further constraints on prosodic that govern the well-formedness of phonological phrases, and on syntax that

govern the structural well-formedness of the sentences.

6.5.3.2 Prosodic constraints

Constraints on prosody govern the parsing of syntactic constituents into prosodic constituents. Taken together, the prosodic constraints I employ here must satisfy two desiderata for French:

(i) phrasal accent is found on the rightmost metrical syllable of a phonological phrase.

(ii) main accent is found aligned with the right edge of an intonational phrase.

For constraints on well-formedness at the phonological level (PhP), I refer the reader to Delais-Roussarie (1996). Here, I follow her work by assuming that (i) prosodic words are grouped into PhP which only contains one head, the head being the most prominent element of the category, and (ii) prosodic prominence is required fall on the rightmost metrical syllable of a PhP. Let us turn to analyzing the intonational phrase (IP) level. Cross-linguistically, assignment of main stress is found at the IP level. In French, IP are strictly right-headed, which means that main stress is required to occur rightmost. This requirement is translated into the phonological constraint ALIGNHIR in (206), which imposes that the right boundary of every IP be aligned with its head.

(206)   ALIGNHIR: Align the right boundary of every intonational phrase with its head.
        (*One violation inferred for each possible head position occurring between the right edge of the IP*)

To illustrate how this constraint works, consider the examples in (207), (208) and (209).

(207)   Satisfying ALIGNHIR:

       (*) IP level

       (*) (*) PhP level

       (S) (V O)

(208)   Violating ALIGNHIR once:

       (*   ) IP level

       (*) (*) PhP level

       (S) (V O)

(209)   Violating ALIGNHIR twice:

       (*   ) IP level

       (*) (*) (*) PhP level

       (S) (V) (O)

The constraint ALIGNHIR is fulfilled in (207) because the right boundary of the IP is aligned with the position of main stress on the object. There are no possible heads between the right edge of the IP and the position of main stress. ALIGNHIR is violated once in (208) since main stress occurs on the grammatical subject and there is one φ boundary (or one possible head position) between mains stress and the right edge of the IP. With the PhP phrasing structure in (209), ALIGNHIR is violated twice because each PhP must contain a head, which means that there is not one, but two possible head positions between main stress and the right edge of the IP. Let us quickly note that because the well-formedness of the PhP is independent of the alignment at the IP level, a PhP structure of the form $(S)_{php} \, (V)_{php} \, (O)_{php}$ would also satisfy ALIGNHIR as long as main stress (marked with a *) at the IP level was found rightmost. An example where all three constituents, S, V and O

would be phrased within the same PhP like (S V O)$_{php}$, with the main stress falling on the S, would violate ALIGNHIR once since the accent would not be aligned rightward.

Concerning prosody, one more constraint is needed in order to prevent the language from creating intonational phrases unless needed to fulfill ALIGNHIR. Such a constraint is expressed in (210), and resembles the structure constraint posited by Delais-Roussarie in (??) for the PhP level.

(210)   *IP: Avoid IP boundaries.

     (*One violation for each IP boundary added*)

The constraint *IP is going to conflict with the constraint ALIGNHIR since, in French, main stress must occur rightward to the detriment of creating more IPs. The set of candidates considered for this conflict includes sentences that differ only by the number of IP they contain. So, everything else being equal and modifying simply the number of IPs, the tableau in (6.3) demonstrates that ALIGNHIR is ranked above *IP.

		X bought a purse focus = subject	ALIGNHIR	*IP
a1.	☞	(MaRIE)$_{IP}$ (a acheté un sac)$_{IP}$		**
b1.		(MaRIE a acheté un sac)$_{IP}$	*!	*
c1.		(MaRIE a acheté)$_{IP}$ (un sac)$_{IP}$	*!	**

Table 6.3: ALIGNHIR > *IP (with canonicals).

Similarly, this ranking holds when a cleft is produced: the cleft creates an additional IP boundary so that the clefted element can receive main stress rightward, and therefore fulfill ALIGNHIR. This is illustrated in tableau (6.4) below. The production study has shown that the preferred output for a subject focus is candidate a2., but that candidate a1. is also a possibility. The preference for a cleft will become apparent when constraints on the syntax will be added to the hierarchy.

	X bought a purse focus = subject	ALIGNHIR	*IP
a2. ☞	(C'est MaRIE)$_{IP}$ (qui a acheté un sac)$_{IP}$		**
b2.	(C'est MaRIE qui a acheté un sac)$_{IP}$	*!	*
c2.	(C'est MaRIE qui a acheté)$_{IP}$ (un sac)$_{IP}$	*!	**

Table 6.4: ALIGNHIR > *IP (with clefts).

The constraint ALIGNHIR has its limitations. Indeed, alone, this constraint does not make predictions regarding whether the *c'est*-cleft is a better candidate than the canonical sentence, that is whether (a1) is a better candidate than (a2). When a grammatical subject is in focus, whether it is realized in situ via phrasing (i.e. it is contained within its own independent IP) or whether it is realized in the pivot position of a *c'est*-cleft, the constraint ALIGNHIR is satisfied since two IPs are created in both cases (as illustrated for candidates (a1) and (a2) in the tableaux above). Therefore, constraints on the syntactic structure of the sentence must be added.

6.5.3.3 Syntactic constraints

Syntactic constraints will govern the mapping of constituents into syntactic positions. In my analysis, I follow Hamlaoui (2009) in assuming that canonical sentences have

the structure in (211), and clefts have the structure in (212):

(211) $[_{TP}$ J'i $[_{T}$ ai $[_{VP}$ t$_i$ $[_{V}$ lu $[_{DP}$ un $[_{NP}$ roman]]]]]]]

(212) $[_{TP}$ $[_{TP}$ C'est $[_{VP}$ $[_{DP}$ Ella]]] $[_{CP}$ Op qui <Op> a mangé $[_{DP}$ un biscuit]]]

As previously mentioned, one interesting facts about Romance language is that while Spanish and Italian allow a grammatical subject to be moved rightward, hence leaving the sentence-initial subject position empty, French disallows such a movement. A markedness constraint like EPP in (213) must therefore be employed to account for the fact that French cannot leave the subject position empty.

(213) EPP: Highest A-specifier must be filled.
       (*One violation for every highest A-specifier position left empty*)

According to the structures assumed above for canonical sentences, the constraint EPP will conflict with a constraint which prevents traces such as STAY.

(214) STAY: No traces.
       (*One violation for every trace*)

In canonical position, the subject is generated inside the VP and then moves to Spec-TP, leaving a trace behind. Although this movement violates STAY, it is required in order to have an overt grammatical subject in Spec-TP. This conflict is represented in tableau 6.5 and leads to the ranking EPP > STAY:

X ate a cookie. focus = subject	EPP	STAY
a. ☞ ($_{TP}$Les enFANTS) ($_{VP}$ **t** ont mangé un biscuit)		*
b.　($_{TP}$($_{VP}$ Les enFANTS ont mangé un biscuit))	*!	

Table 6.5: EPP > STAY

Another syntactic constraint must be added to account for the occurrence of non canonical word-orders like clefting, where the preverbal subject position is filled by a pronoun (e.g. 'ce' for the *c'est*-cleft). Indeed, the constraint EPP only prevents the preverbal subject position to be left empty, but does not make any prediction concerning the appropriateness of a non-canonical structure that has a dummy subject vs. a canonical sentence. In other words, EPP cannot make predictions between "C'est Marie qui a mangé les bonbons" and "Marie a mangé les bonbons". In order to achieve this, another constraint on syntactic faithfulness must be taken into account. This constraint, stated in (215), will require that no syntactic structure be added to the input. In other words, it will favor SVO word-order and penalize any syntactic permutations like clefts.

(215)　FAITHSYNTAX (FAITHSYN): Do not add syntactic material not present in the input.

　　　(*One violation for every layer of syntactic structure not present in the input*)

FAITHSYN will conflict with EPP since French frequently allows constituents to be realized in cleft constructions (or other non-canonical word orders), but never allows the canonical subject position to be left empty. The set of candidates considered here includes the candidates that differ only by their word-order , as illustrated in the tableau (6.6).

208

X ate a cookie. focus = subject	EPP	FAITHSYN
a. ☞ (Les enFANTS)$_{IP}$ (ont mangé un biscuit)$_{IP}$		
b. (Ce sont les enFANTS)$_{IP}$ qui ont mangé un biscuit)$_{IP}$		*
c. (ont mangé un biscuit)$_{IP}$ (les enFANTS)$_{IP}$	*!	

Table 6.6: EPP > FAITHSYN

The constraints discussed so far are brought together in the tableau (6.7), where the set of candidates differs only by word-order and IP phrasing only.

X ate a cookie. focus = subject	EPP	ALIGNHIR	FAITHSYN	STAY	*IP
a. ☞ (Les enFANTS)$_{IP}$ (ont mangé un biscuit)$_{IP}$				*	**
b. (C'est les enFANTS)$_{IP}$ (qui ont mangé un biscuit)$_{IP}$			*		**
c. (Les enFANTS ont mangé un biscuit)$_{IP}$		*!		*	*
d. (C'est les enFANTS qui ont mangé un biscuit)$_{IP}$		*!	*		*
e. (ont mangé un biscuit les enFANTS)$_{IP}$	*!		*	*	*

Table 6.7: EPP, ALIGNHIR > FAITHSYN > STAY, *IP

From now on, I will simply assume that a candidate with an empty preverbal subject is not felicitous and will be discarded as ever appearing as an output. But here again, there is still something missing to derive the correct empirical generalizations observed in French. Indeed, these constraints alone do not predict that the cleft will come out as the best candidate when the subject is a focus. Instead, they predict the canonical sentence as the output. One last type of constraint - a constraint on information structure - is needed to give a full account of focus marking and the alternation cleft/canonical in French.

6.5.3.4 Information structure constraint

This type of constraint govern the mapping of information structure and interpretation onto the prosodic constituents. Again, we need a constraint that will interact with the others in order to allow a cleft to emerge as the favorite candidate for subject focus marking, but not for object foci.

As a default, subjects have been argued by many scholars to be interpreted as topics in different languages (Chafe, 1976). Consequently, a constraint requiring this mapping is formulated in (216). If ranked high, this constraint will guarantee that subjects that cannot be interpreted as topics will have to be marked (Zimmermann, 2006). When ranked low, this constraint will not interfere with grammatical subjects being focused in canonical position.

(216) *SUBJFOC: The grammatical subject of a sentence must not contain a focus.

(*One violation for a narrowly focused grammatical subject*)

Since French prefers to add a layer of syntactic structure rather than in preverbal position where the element would be interpreted as a topic, *SUBJFOC conflicts with and dominates FAITHSYN. The set of candidates considered is identical in all respects except in the word-order, as illustrated in tableau (6.8).

X ate a cookie focus = subject	*SUBJFOC	FAITHSYN
a. ☞ (C'est les enFANTS)$_{IP}$ (qui ont mangé un biscuit)$_{IP}$		*
b. (Les enFANTS)$_{IP}$ (ont mangé un biscuit)$_{IP}$	*!	

Table 6.8: *SUBJFOC > FAITHSYN

6.5.4 Deriving the cleft/canonical alternation in neutral contexts

6.5.4.1 Subject focus

The data from the production experiment presented in the previous chapter show that in neutral contexts, participants mostly produce the form [Cleft S]$_F$VO (86% of the time). In the remaining cases (14%), participants produced the canonical form [S]$_F$VO. Tableau 6.9 illustrates that the constraint ranking ALIGNHIR > *SUBJFOC > FAITHSYN predicts exactly this statistical distribution. Let's note that I assume prosodic well-formedness at the PhP level: the primary accent falls on the last syllable of the focus "Marie", and the first syllable "Ma" can optionally be accented. However, it will not make a difference for the OT analysis of cleft/canonical interaction proposed here.

In Tableau 6.9 below, the candidates considered differ by three features, namely the number of IPs in the sentence, the alignment of prominence and the word-order. The optimal candidate (a) is realized in a cleft which creates two independent IPs, allowing the focus subject to receive prosodic prominence, thus obeying the alignment constraint ALIGNHIR but violating the lower ranked constraint *IP. The optimal candidate also violates the low-ranked constraint FAITHSYN because the focus subject is realized in a non-canonical word-order. In the pilot prosodic analysis I discussed in Chapter 5, the data suggests that grammatical subjects are not always realized in their own IP. This is also noted by Delais-Roussarie (1996) and Féry (2001), who argue that the size of the subject will influence its phrasing. If the subject focus is not realized within its own IP (candidate b), the authors argue it is realized within its own phonological phrase via a primary accent HL or LH on the last syllable, and is phrased with the following constituents at the IP level. Tableau 6.9 combines the set of candidates where the subject focus realized within its own

IP and the set of candidates where the subject is not. Candidate (b), where the subject focus is realized as its own IP in preverbal position is the second best output: it violates *SUBJFOC which requires a canonical subject not to be interpreted as a focus and violates the low ranked constraint *IP by creating an extra intonational phrase around the subject focus. *SUBJFOC being ranked slightly higher than FAITHSYN, candidate (b) is predicted to occur with a smaller frequency than candidate (a).

	X bought a newspaper focus = subject	ALIGNHIR	*SUBJFOC	FAITHSYN	*IP
86% a.	(C'est maRIE)$_{IP}$ (qui a acheté un journal)$_{IP}$			*	**
14% b.	(maRIE)$_{IP}$ a acheté un journal.		*		**
c.	(C'est maRIE qui a acheté un journal)$_{IP}$	*!		*	*
d.	(maRIE a acheté un journal)$_{IP}$	*!	*		*

Table 6.9: Subject focus

In order to produce this variation where candidate (b) comes out as the optimal output 14% of the time, thus allowing a focus subject to be marked via prosody in a canonical sentence, the constraints *SUBJFOC and FAITHSYN must be understood as overlap. 86% of the time, the ranking *SUBJFOC FAITHSYN holds and the cleft is the favorite output, and 14% of the time, the opposite ranking holds and the canonical sentence is favored. In other words, when the candidates evaluated are in the upper part of the constraint FAITHSYN and the lower part of the constraint *SUBJFOC, the opposite ranking will hold and FAITHSYN will dominate *SUBJFOC and candidate (b) is favored (see figure 6.5). The variation across speakers can be explained by the strictness of the *SUBJFOC constraint. This strictness is a function of the distance between this constraint and the conflicting constraint. Speakers who have a loose version of *SUBJFOC, where the distance between the two constraints *SUBJFOC and FAITHSYN is small, will produce more canonicals. On the

other hand, speakers who have a more stringent version of *SUBJFOC will have a longer distance between the two constraints *SUBJFOC and FAITHSYN. As a result, they will produce canonicals less frequently than clefts.
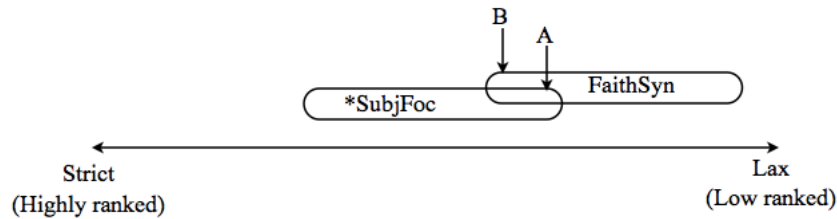

Figure 6.5: Overlapping constraints *SUBJFOC and FAITHSYN

6.5.4.2 Direct and indirect object focus

The production data shows that direct or indirect object foci are mainly realized in situ (77%) but can also be clefted (23%). Within the clefted cases, the difference in pragmatic context (neutral vs. correction) had a significant effect: more clefts were produced in the corrective context. This variation will need to be accounted for in the OT analysis.

But first let us explain the main cases when objects remain in situ. We shall note that, here, the constraint *SUBJFOC does not come into play, since the focus element is a non-subject. The analysis for direct object focus here is quite straightforward: in its original position, an object occurs rightmost in the position where main stress occurs by default. Therefore, there is no clash between syntax and phonology. Tableau 6.10 illustrates the ranking of constraints that generate the canonical sentence as the favorite output and the cleft as second best.

Candidate (a), which violates the low-ranked *IP constraint and the pragmatic constraint FOCCLEFT relevant for non-subjects comes out as the optimal candidate. A

	Mary bought X focus = direct object	ALIGNHIR	FAITHSYN	FOCCLEFT	*IP
75% a. ☞	(Marie a acheté un jourNAL)$_{IP}$			*	*
b.	(Marie a acheté)$_{IP}$ (un jourNAL)$_{IP}$			*	**
25% c.	(C'est un jourNAL)$_{IP}$ (que Marie a acheté)$_{IP}$		*		**
d.	(C'est un jourNAL que Marie a acheté)$_{IP}$	*!	*		*

Table 6.10: Direct objects focus

canonical sentence where the focus object is realized as its own IP is a possibility and is slightly worst than (a) because it violates *IP twice. A more in-depth phonological analysis must be conducted to investigate whether speakers always realize object foci in an independent IP as in candidate (b) or within the same IP as illustrated in (a). A clefted object such as candidate (d), where the cleft does not form its own IP, differing from candidate (a) only by the word-order used, will be completely ruled out because it violates the higher ranked constraint ALIGNHIR. Indeed, the clefted object, the head of HIR, is not aligned rightward. In order to get a better cleft candidate, the clefted object must be contained within its own IP, as illustrated in candidate (c), which in turn violates the low-ranked constraint *IP. Both clefted candidates (c) and (d) will violate the constraint FAITHSYN that demands faithfulness to the input form.

In order to get the variation observed between candidate (a) and (c), the constraints FAITHSYN and FOCCLEFT must be understood as overlapping as illustrated in Fig 6.6. 84% of the time, the constraint FAITHSYN will be ranked above FOCCLEFT to have the canonical sentence be the favorite output and 16% of the time, the ranking will be reversed to have the cleft emerge.

When a direct object is focussed and is followed by an indirect object or an adjunct, the same analysis holds. The data in tableau 6.11 is analogous to the one collected in the
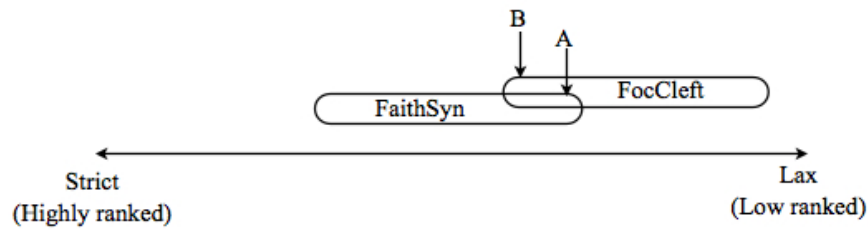
Figure 6.6: Overlap between FAITHSYN and FOCCLEFT

production experiment in Chapter 5, since participants were instructed to respond with full sentences. All stimuli included an indirect object, whether it was the focus targeted or not. Therefore, under the direct object focus condition, participants were expected to produce the sequence direct object-indirect object.

	Mary bought X at the newsstand focus = direct object	ALIGNHIR	FAITHSYN	FOCCLEFT	*IP
75% a. ☞	(Marie a acheté un jourNAL)$_{IP}$ (au kiosque)$_{IP}$			*	**
25% b.	(C'est un jourNAL)$_{IP}$ (que Marie a acheté au kiosque)$_{IP}$		*		**
c.	(Marie a acheté au kiosque un jourNAL)$_{IP}$		*	*	*
d.	(Marie a acheté au kiosque)$_{IP}$, (un jourNAL)$_{IP}$		*	*	**
e.	(Marie a acheté un jourNAL au kiosque)$_{IP}$	*!*		*	*
f.	(C'est un jourNAL que Marie a acheté au kiosque)$_{IP}$	*!*	*		*

Table 6.11: Direct object focus followed by indirect object or adjunct.

As previously discussed, empirical results show that participants realized direct objects in situ 75% of the time. The ranking of constraints below correctly predicts that a canonical sentence (candidate a) will be the optimal output. This candidate demonstrates that a focused direct object is marked by phrasing; by occurring at the right edge of an IP boundary. The indirect object or adjunct occurring post focus is deaccented (via a low plateau or a continuous fall) and/or dephrased. This is indeed observed in the preliminary prosodic analysis in Chapter 5, and also argued by many studies in the literature on French (Féry, 2001; Delais-Roussarie *et al.*, 2002; Claire *et al.*, 2004). If the focus direct

object was not phrased separately from post-focus material, a clefted candidate such as (b) would be a better output. Second best output is a tie between a cleft and a dislocation. Different studies including Zubizarreta (1998), Samek-Lodovici (2005b) and Hamlaoui (2008) already note that a sentence like candidate (c) is a marginal possibility. Yet, Hamlaoui (2008) concedes that she has no evidence from native speakers' judgments on the acceptability of (c). In my production experiment, candidate (c) where the direct object is extraposed was never produced. Candidates (d) and (e) where the focus does not occur rightmost are the least favorite.

As for direct objects, the same analysis holds for indirect objects as illustrated in Tableau 6.12. Everything else being equal and only differing by the word-order of the sentence, tableau 6.12 demonstrates that a canonical sentence is preferred over a cleft. Indeed, assuming that the focussed element is realized within its own IP, the optimal candidate is the canonical candidate (a): the indirect object occurs rightmost, so there is no conflict between syntax and prosody. The clefted form in (c) is marked as it violates FAITHSYN.

	Mary bought a newspaper at X focus = indirect object	ALIGNHIR	FAITHSYN	FOCCLEFT	*IP
90% a. ☞	(Marie a acheté un journal au KIOsque)$_{IP}$			*	*
b.	(Marie a acheté un journal)$_{IP}$ (au KIOsque)$_{IP}$			*	**
10% c.	(C'est au KIOsque)$_{IP}$ (que Marie a acheté un journal)$_{IP}$		*		**
d.	(C'est au KIOsque que Marie a acheté un journal)$_{IP}$	*!**	*		*

Table 6.12: Indirect objects focus

6.5.4.3 Sentence focus

Let's now turn to cases when the focus is *broad*, or sentence-wide. In these cases, the focus does not target one particular constituent, but rather the entire sentence. It is characterized by the absence of a topic-comment relation between the subject and the predicate, and by the absence of a focus-presupposition relation between an argument and an open proposition (Lambrecht, 2000). This type of focus is generally triggered by a QUD like "What happened?". In this case, the driving constraints *SUBJFOC, and FOC-CLEFT are not relevant, and it is in fact the other constraints that derive the optimal output. The results from the production experiment are very straightforward: speakers always realize sentence-focus via prosody, not via clefting. At the IP level, the position of main stress naturally matches with the rightmost edge of the sentence, therefore predicting candidate (a) as optimal. Furthermore, candidate (a) does not violate *SUBJFOC since the grammatical subject is not the focus of the sentence.

focus = sentence		ALIGNHIR	*SUBJFOC	FAITHSYN	FOCCLEFT	*IP
100%a. ☞	(Marie a acheté un journal au kiosque)$_{IP}$					*
b.	(C'est Marie qui a acheté un journal au kiosque)$_{IP}$			*!		*

Table 6.13: Sentence focus

Under this focus condition, a cleft candidate that would contain two IPs, such as (C'est Marie)$_{IP}$ (qui a acheté un journal au kiosque)$_{IP}$, would be infelicitous. Indeed, this division separates the pivot from the post-verbal material in a case where the interpretative unit containing the focus is larger than the pivot itself. In cases where the pivot is the element answering the QUD, this division is desired, since it marks the pivot as the focus constituent. However, in the case of a sentence focus, the pivot, here "Marie", is not the

217

(only) focus. Having the prosodic structure of a clefted candidate with two IPs, which is similar to the structure found for subject-focus, will also violate the undominated constraint SF because it predicts marking on the wrong constituent. Candidate (b), on the other hand, where the information in the relative clause contains part of the focus, only violates the low ranked constraint FAITHSYN. Claire *et al.* (2004) discuss the prosodic realization of clefted sentence-focus and argue that it is characterized by Major Phrases, each of them ending with a continuative H tone (instead of a boundary tone in narrow-focus clefts). To illustrate, the broad-focus cleft in (b) has the intonational characteristics in (217a), whereas a narrow-focus will have the intonational characteristics in (217b) with a boundary tone to the right of the pivot.

(217)   a. C'est Marie]H$_{continuative}$ qui a acheté]H$_{continuative}$ un journal au kiosque.]L%

      b. C'est Marie]L% qui a acheté un journal au kiosque.]L%

6.5.4.4   Overall distribution

Before examining stronger focus interpretations, here is the complete ranking of constraints that accounts for the majority of the data, and the free variation observed:

(218)   SF, EPP, ALIGNHIR, *SUBJFOC > FAITHSYN > FOCCLEFT, *IP

The lowest-ranked constraints (i.e. the ones that can be violated without causing rejection of the candidate) are found on the right hand-side where they are considered laxed, and on the left are equally highest-ranked constraints that will distinguish the optimal candidate, SF, EPP, ALIGNHIR and *SUBJFOC. The diagram in 6.7 shows the two pairs of constraints that overlap in order to create variation: *SUBJFOC overlaps with FAITHSYN

and FAITHSYN overlaps with FOCCLEFT. The next section in this chapter will introduce new constraints necessary to account for the variation triggered by semantically stronger interpretations.
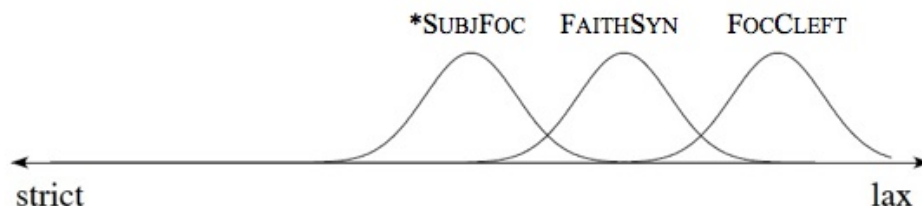

Figure 6.7: Overlapping constraints.

6.5.5 Deriving the cleft/canonical alternation in non-neutral contexts

At the beginning of the chapter, I discussed the fact that French displayed some strong preferences for marking focus, but also talked about the variation observed in the data. In the previous section, I accounted for the cleft vs. canonical alternation when marking focus in neutral/ informational contexts and the free variation occurring in these cases. I showed that subject focus is grammatically marked because of a prosodic constraint requiring main stress to occur rightmost, but that non-subjects are mostly left in situ. However, there exists another type of variation occurring with semantically stronger interpretations of focus. In chapter 5, I showed that the realization of focus is affected in cases where focus is *contrastive* and/or *corrective*. This is especially noticeable for non-subjects that tend to be clefted to a significantly higher degree than in neutral contexts. An analysis of focus realization must also account for this variation. In Chapter 3, I discussed another interpretation of the cleft, namely *exhaustivity*. Under this interpreta-

219

tion, I showed that the cleft was also preferred to a canonical. The forced-choice experiment in Chapter 3 showed that the cleft was more exhaustive than a canonical sentence. Again, this preference must be accounted for if we wish to provide a full explanation for the cleft/canonical alternation.

Let us first look back at *exhaustivity*. In Chapter 3, I claimed that this interpretation is not part of the truth-conditional meaning of the cleft, but is an implicature that arises from the context in which the sentence is used. I argued that the speaker will use a cleft if he wishes to convey to the hearer that the element in the cleft constitutes the total answer speaker commits to. Indeed, speaker did not choose to utter a stronger form containing an exclusive that would have asserted the exhaustive content of the sentence. Neither did he choose to utter a canonical sentence that would have lacked any exhaustive implication. The choice of the conveys that speaker wants to implicate exhaustivity without committing to the truth of it, e.g. without asserting that there are no stronger alternatives that exist. Hence, I argued that the competition between the pivot position and the preverbal subject position underlies the competition between the cleft and the canonical sentence. I proposed that the exhaustive effects observed in the cleft can be accounted for by a general markedness constraint FOCCLEFT$_{exhaustive}$ in (219) that has requirements for the pivot position but also for the preverbal subject position, and which is triggered when an exhaustive interpretation is demanded:

(219)   FOCCLEFT$_{exhaustive}$: A focus with an exhaustive reading must be clefted.

In a context where the focus must be interpreted as a total answer to the QUD, this constraint explains why the focus element is realized as a pivot. Let us take the case of an ob-

ject focus: the cleft will be used to convey exhaustivity even if this means that the canonical word order is not maintained, thus violating the faithfulness constraint on syntactic ordering FATIHSYN. In tableau 6.14, illustrating the ranking of FOCCLEFT$_{exhaustive}$ above FAITHSYN, candidate (b) is ruled out because, despite respecting the canonical word order SVO, it violates the constraint demanding that exhaustive elements be realized in a cleft.

	Pauline kissed X focus = direct object	FOCCL$_{exh}$	FAITHSYN
a. ☞	(C'est son MAri)$_{IP}$ (que Pauline a embrassé)$_{IP}$.		*
b.	(Pauline a embrassé) (son MAri)$_{IP}$	*!	

Table 6.14: Object focus in exhaustive context

The constraint in (219) also predicts why the cleft will not be used in cases where there is a very clear non-exhaustive reading, such as cases where a quantified NP is focused. Let us consider a focused subject for this example, since subjects are typically preferred realized in a cleft. Under the condition of a non-exhaustive reading, the focus subject is favored in a canonical sentence, appearing in preverbal position so that it can fulfill FOCCLEFT$_{exhaustive}$ as illustrated in Tableau 6.15. In this case, the FOCCLEFT$_{exhaustive}$ constraint conflicts with the *SUBJFOC which demands that a preverbal subject do not contain a focus element.

	X kissed Pauline ? focus = subject	FOCCL$_{exh}$	*SUBJFOC
a. ☞	(QUELques enfants)$_{IP}$ (ont embrassé Pauline)$_{IP}$.		*
b.	(C'est QUELques enfants)$_{IP}$ (que Pauline a embrassé)$_{IP}$	*!	

Table 6.15: Subject focus with quantified NP (non-exhaustive context)

The second notion that I looked at was *contrast/correction* in Chapter 4 and 5. More specifically, looking at a series of examples from corpus search, I argued that the cleft occurred when there is a mismatch between the speaker's and the addressee's expectations (or the expectations speaker has for the addressee). I argued that the canonical form did not convey this meaning and the cleft was used in order to signal this mismatch. I posited the constraint repeated in (220) to account for the use of the cleft in these contexts.[8] In Chapter 5, the results from the production experiment also confirmed that the cleft was significantly more likely to be used if the context conveyed a correction of a faulty assumption.

(220)   FOCCLEFT$_{contrastive}$: A focus element with a contrastive reading must be clefted.

Tableau 6.16 shows the interaction between the constraint FOCCLEFT$_{exh}$ and the constraint FOCCLEFT$_{contrast}$. It shows that the ranking FOCCLEFT$_{contrast}$ > FOCCLEFT$_{exh}$ predicts that a focus that includes quantifiers and other numerals do occur in the cleft. Indeed, as an answer to the QUD 'Did all the students come?', candidate (a) violates the constraint FOCCLEFT$_{exh}$ by not realizing a quantified focus in preverbal position but fulfills the requirement of having contrastive/corrective elements in pivot position. Thus, even though the focus is not interpreted exhaustively, it will be realized in pivot position in order to signal a faulty assumption. This shows that the constraint on pragmatics overrides the constraints on semantic acceptability.

---

[8]In order to account for the emergence of clefts in contrastive contexts, Hamlaoui (2008) posits an extension of the SF constraint: SF$_c$, which is based on a similar constraint proposed in Féry and Samek-Lodovici (2006). However, she does not look at the type of data I discuss in Chapter 4, where quantified pivots occur in pivot position.

	focus = subject	FOCCL$_{contrast}$	FOCCL$_{exh}$
a. ☞	Non, (ce sont QUELques étudiants)$_{IP}$ qui sont venus.		*
b.	Non, (QUELques étudiants)$_{IP}$ sont venus.	*!	

Table 6.16: Non-exhaustive Subject focus in contrastive context

Lastly, the Tableau 6.17 shows how FOCCLEFT$_{contrastive}$ predicts that a cleft will be used in a contrastive/corrective context for non-subjects as well.

	focus = direct object	FOCCL$_{contrast}$	FAITHSYN
a. ☞	Non, (c'est un jourNAL)$_{IP}$ (que Julie a acheté)$_{IP}$.		*
b.	Non, (Julie a acheté)$_{IP}$ (un jourNAL)$_{IP}$.	*!	

Table 6.17: Object focus in corrective context

Overall, the constraints FOCCLEFT$_{contrastive}$ and FOCCLEFT$_{exh}$ are a subset of the FOCCLEFT constraint. The overall ranking is illustrated below:

(221) SF, EPP, ALIGNHIR, *SUBJFOC > FAITHSYN > FOCCLEFT, FOCCLEFT$_{contrast}$, FOCCLEFT$_{exh}$, *IP

Chapter 7

Conclusion

This dissertation has examined the meaning and usage of the French *c'est*-cleft. The overarching research question that I explored is 'What does the French *c'est*-cleft add to a sentence that a canonical sentence alone cannot do?', to which I answered that the cleft marks a piece of information as most relevant and as a total answer in the discourse. The three major findings of the present work are that (i) the cleft does not systematically encode the uniqueness of an answer (Chp 3, *contra Clech-darbon* et al. *1999) but is used to convey the total answer that the speaker wished to commit to, (ii) the cleft interacts with various expressions like quantifiers and additives only when the context is clearly contrastive (Chp 4,* contra *Kiss 1998), and (iii) the cleft is not systematically used to mark focus, especially on non-subjects (Chp 5,* contra Lambrecht 1994).

I started my dissertation by exploring the **meaning** of the cleft. The main issue I tackled in that exploration is the particular meaning often associated with the cleft: Exhaustivity. The previous literature on French argues that the cleft systematically encodes the uniqueness of the focussed element. I also discussed the lively cross-linguistic debate that still exists about the nature of this exhaustivity by discussing the different arguments scholars propose on the issue. In an effort to provide empirical evidence for "how" exhaustive the French cleft is, I ran a forced-choice task comparing the level of exhaustivity of three sentence forms: exclusive-sentences, clefts and canonicals. This experiment was

224

adapted from a series of similar experiments on English, German and Hungarian. Results, similar to the ones found for the other languages, showed that the cleft does not behave like other sentence forms when it comes to being interpreted as exhaustive, which led me to argue against a semantic account of exhaustivity: the cleft does not systematically encode the uniqueness of the answer. Instead, I proposed that exhaustivity in the French cleft arises from a constraint on the way speakers use language, a conclusion which recalls the conversational implicature argument associated with Horn's early work (esp. Horn 1981).

This argument was supported further in Chapter 4 where I turned to examining more data from corpora and Google hits (naturally occurring examples). Cross-linguistically, some scholars like Kaitlin Kiss used the apparent ban of certain expressions in post-copular/ pivot position as an argument in favor of a semantic account of exhaustivity. However, I showed that these claims were not empirically substantiated. The major finding in Chapter 4 was that the ratio of canonicals/clefts for various given expressions like "all/tout" and "even/même" is non null. Essentially, what this means is that the semantics of the cleft does not clash with the semantics of these expressions. Thus, I concluded that exhaustivity was not semantic. Yet, interestingly, when looking at the context surrounding the occurrence of "cleft + expression", I found that all instances occurred in a rather specific context; a very highly contrastive one. This result was taken to show that context, specifically a contrastive and/or corrective one, plays a major role in the use of the *c'est*-cleft. This claim is explored further in the next chapter (Chp. 5), where *context* constitutes one of the independent variables of the production experiment.

As just mentioned, Chapter 5 changes the perspective from comprehension to

production by exploring the **usage** of the *c'est*-cleft. The question raised by this chapter is under what circumstances a cleft is produced. Recall that French is largely assumed to resort to clefting to mark focus on arguments (Lambrecht, 1994). Yet, some scholars such as Claire *et al.* (2004) dispute this view, claiming that prosody can be used, especially to mark focus on grammatical objects. Moreover, these same scholars also tend to dispute the realization of subject focus, arguing that it can also be marked via prosody. The production experiment presented in chapter 5 is a first step towards settling these disputes because it provides experimental evidence for the realization of different grammatical constituents in different contexts.

Results clearly show that the claim that French bans lexical focus subjects from sentence-initial position is too strong, since subjects are produced sentence initially with a non-zero frequency. The results are also interesting because they bring forward a clear marking asymmetry between subjects and objects: subjects are mainly clefted and objects are mainly realized in situ. This is even more interesting when put in a broader cross-linguistic context since it concurs with what is observed for many other languages (notably in Chadic languages by Zimmerman): focus subjects and objects are not typically marked the same way. Finally, the results seems to support the claim made by Lambrecht that *c'est*-clefts are used for narrow focus only since no cleft was produced to answer a broad-focus question 'What is happening?'. An interesting question for further research related to this production experiment is to what extent can hearers match a given sentence with the appropriate question. In other words, the goal would be to test whether hearers can as reliably match the sentences containing prosodic focus with the appropriate question as sentences that contain clefted focus. The hypothesis would be that clefts

226

facilitate and maybe even speed up the processing of focal information.

Finally, Chapter 6 proposed a unified account of the use of the cleft by bringing together the different constraints formulated at each step of the research. I couched my analysis in a Stochastic OT framework because it allowed me to explain the free variation observed in the data, something that had not been done in previous studies.

One way in which I intend to take this research further is by looking at the cleft/ canonical alternation from a second language acquisition perspective. Through this dissertation, I showed that French commonly uses the cleft, and many cross linguistic studies show that French uses them more frequently than, for example, English (Katz, 2000; Carter-Thomas, 2009; Dufter, 2009). Thus, an L2 learner who wants to speak French authentically and naturally must master the *c'est*-cleft, among other non-canonical sentences. Unfortunately, the reality is that, in most of today's classrooms, instructors do not teach clefts, nor other non-canonical forms. One reason may be that there is a tendency to want to teach the standard language, which in the case of French would correspond to what the language is in its written form. But the fact that written and spoken French differs greatly is well-known and recognized by many scholars (Katz and Blyth, 2007; Hamlaoui, 2009). Blyth (1999, 186) notes that "relatively few foreign language materials make extensive use of authentic interaction; scripted dialogues and scripted videos still rule the day … How is one supposed to teach the spoken language with materials that do not reliably reflect typical speech patterns? In fact, textbooks frequently fail to even mention or exemplify constructions that are prevalent in spoken language." More generally, students receive little information about the differences between the written and spoken language. Thus, one of my short term goals is to adapt the production experiment

to second language learners to test their proficiency in using clefts appropriately. I believe it will be interesting to conduct a cross-sectional study on three groups of learners (such as maybe Lower-Division Advanced, Upper-Division and Near-Natives) to investigate the emergence of the cleft as a focus marking strategy by comparing these groups controlling for who has none or some explicit instruction vs. direct exposure in the country.

Bibliography

Atlas, J. D. and Levinson, S. C. (1981). It-clefts, informativeness and logical form: Radical pragmatics. In *Radical pragmatics*, pages 1–62. Academic Press.

Beaver, D. (2011). It constructions.

Beaver, D. and Onea, E. (2011). Hungarian focus is not exhausted. In S. I. . D. L. Ed Cormany, editor, *Proceedings of Semantics and Linguistic Theory (SALT) 19*, pages 342–359.

Beaver, D. and Velleman, D. (2011). The communicative significance of primary and secondary accents. *Lingua*, **121**.

Beaver, D., Francez, I., and Levinson, D. (2005). Bad subject: (non-)canonicality and np distribution in existentials. In *Proceedings of Semantics and Linguistic Theory XV*, Ithaca, New York. CLC Publications.

Beaver, D., Roberts, C., Simons, M., and Tonhauser, J. (2011). What projects and why. *Proceedings of Semantics and Linguistic Theory XX*, pages 309–327.

Beckman, M. and Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, **3**, 15–70.

Belletti, A. (2005). Answering with a cleft. the role of the null subject parameter and the vp periphery. In *Proceedings of the XXX Incontro di Grammatica Generativa*. Cafoscarina.

Beyssade, C., Marandin, J.-M., and Rialland, A. (2001). Ground/focus revisited. a perspective from french. In *Selected papers of LSRL 2001*, Amsterdam. John Benjamins.

Beyssade, C., Hemforth, B., Marandin, J.-M., and Portes, C. (2010). Prosody and information structure in french. Tutorial presented at HPSG 2010: 17th International Conference on Head-Driven Phrase Structure Grammar. Paris.

Beyssade, C., Hemforth, B., Marandin, J.-M., and Portes, C. (2011). Prosodic markings of information focus in french. In H.-Y. Yoo and E. Delais-Roussarie, editors, *Actes d'Interface Discours et Prosodie 2009*, pages 109–122, Paris, France.

Blyth, C. S. (1999). Toward a pedagogical discourse grammar: Techniques for teaching word order constructions. In J. Lee and A. Valdman, editors, *Form and Meaning: Multiple perspectives. AAUSC Issues in Language Program Direction*, pages 183–229. Heinle, Boston.

Boersma, P. and Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, **32**, 45–86.

Büring, D. (2001a). Let's phrase it!—focus, word order, and prosodic phrasing in german double object constructions. In G. Müller and W. Sternefeld, editors, *Competition in Syntax*, number 49 in Studies in Genitive Grammar, pages 101–137. de Gruyter, Berlin & New York.

Büring, D. (2001b). Let's phrase it! focus, word order, and prosodic phrasing in german double object constructions. In G. Müller and W. Sternefeld, editors, *Competition in*

Syntax, number 49 in Studies in Genitive Grammar, pages 101–137. de Gruyter, Berlin & New York.

Büring, D. (2009). Towards a typology of focus realization. In M. Zimmermann and C. Féry, editors, *Information Structure*, pages 177–205. Oxford University Press, Oxford.

Büring, D. (2010a). Conditional exhaustivity presuppositions in clefts (and definites). Manuscript, UCLA.

Büring, D. (2010b). Towards a typology of focus realization. In M. Zimmermann and C. Féry, editors, *Information Structure*, pages 177–205. Oxford University Press, Oxford.

Büring, D. and Gutiérrez-Bravo, R. (2001). Focus-related word order variation without the nsr: A prosody-based crosslinguistic analysis. In S. M. Bhloscaidh, editor, *Syntax at Santa Cruz 3*, pages 41–58.

Carter-Thomas, S. (2009). The french c'est-cleft: Functional and formal motivations. In J. O. David Banks, Simon Eason, editor, *La linguistique systémique fonctionelle et la langue française*. L'Harmattan.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. Li, editor, *Subject and Topic*, pages 25–55. Academic Press, New York.

Chomsky, N. and Halle, M. (1968). *The sound pattern of English*. Harper and Row, New York, NY.

Claire, Delais-Roussarie, E., Doetjes, J., Marandin, J., and Rialland, A. (2004). Prosody and information in french. In F. Corblin and H. de Swart, editors, *Handbook of French Semantics*, pages 477–500. CSLI.

Clech-darbon, A., Rebuschi, G., and Rialland, A. (1999). Are there cleft sentences in french? In G. Rebuschi and L. Tuller, editors, *The Grammar of Focus*, pages 83–118. John Benjamins, Amsterdam.

Comrie, B. (1981). *Language Universals and Linguistic Typology*. University of Chicago Press, Chicago.

Cottell, S. (2002). *The comparative syntax of cleft constructions*. Ph.D. thesis, PhD dissertation, University of Wales, Bangor.

Cowles, W. (2003). *Processing Information Structure: Evidence from Comprehension and Production*. Ph.D. thesis, Unpublished PhD dissertation, UC San Diego.

David, L. G. and Yoo, H.-Y. (2000). Intonative structure of focalization in french and greek. In *Proceedings of the 14th Symposium on Romance Linguistics "Going Romance 2000"9*, Utrecht, The Netherlands.

de Cat, C. (2002). *French dislocation*. Ph.D. thesis, University of York, UK.

de Cat, C. (2007). *French dislocation: Interpretation, Syntax and Acquisition*. Oxford University Press.

Delais-Roussarie, É. (1995). *Phonologie et Grammaire : Études et modélisation des interfaces prosodiques*. Ph.D. thesis, Université de Toulouse-le Mirail. Mémoire de synthèse présenté pour l'obtention d'une Habilitation à Diriger des Recherches.

Delais-Roussarie, É. (1996). Phonological phrasing and accentuation in french. In M. N. . N. Smith, editor, *Dam Phonology : HIL Phonology Paper II*. Holland Academic Graphics.

Delais-Roussarie, É. and Post, B. (2008). Unités prosodiques et grammaire de l'intonation: une nouvelle approche. In *Actes des XXVIIème Journées d'études sur la Parole (JEPTALN 2008).*

Delais-Roussarie, E. and Rialland, A. (2007). Metrical organization, tonal association and focus in french. In S. Baauw and F. Drijkoningen, editors, *Romance languages and linguistic theory 2005*, pages 73–98. John Benjamins Publishing.

Delais-Roussarie, E., Rialland, A., Doetjes, J., and Marandin, J. (2002). The prosody of post-focus sequences in french. In B. B. et I. Marlien, editor, *Proceedings Speech Prosody 2002*, Aix en Provence.

Delin, J. (1992). Properties of em It-cleft presupposition. *Journal of Semantics*, **9**(4), 289–306.

Delin, J. and Oberlander, J. (1995). Syntactic constraints on discourse structure: The case of it-clefts. *Linguistics*, **33**(3).

di Cristo, A. (1999). Le cadre accentuel du français contemporain: Essai de modélisation: deuxième partie. *Langues*, **2**(4), 258–267.

Doetjes, J., Rebuschi, G., and Rialland, A. (2004). Cleft sentences. In F. Corblin and H. de Swart, editors, *Handbook of French Semantics*, pages 529–552. CSLI.

Dufter, A. (2009). Clefting and discourse organization: Comparing germanic and romance. In A. Dufter and D. Jacob, editors, *Focus and Background in Romance Languages*, pages 83–121. John Benjamins.

Edmonds, J. E. (1976). *A Transformational Approach to English Syntax.* Academic Press, New York.

Erteschik-Shir, N. (1997). *The Dynamics of Focus Structure.* Cambridge University Press, Cambridge.

Féry, C. (2001). Focus and phrasing in french. In *Audiatur Vox Sapientiae. A Festschrift for Arnim von Stechow*, pages 153–181, Berlin. Akademie-Verlag.

Féry, C. and Samek-Lodovici, V. (2006). Focus projection and prosodic prominence in nested foci. *Language*, **82**, 131–150.

Féry, C., Hörnig, R., and Pahaut, S. (2010). Correlates of phrasing in french and german from an experiment with semi-spontaneous speech. In C. Gabriel and C. LleŮ, editors, *Intonational Phrasing at the Interfaces: Cross-Linguistic and Bilingual Studies in Romance and Germanic.* Benjamins.

Frege, G. (1892). Uber Sinn und Bedeutung. *Zeitschrift fur Philosophie und philosophische Kritik*, pages 25–50.

Gabriel, C. (2010). On Focus, Prosody and Word Order in Argentinian Spanish: A minimalist OT account. *ReVEL*, **4**.

Grimshaw, J. (1997). Projection, heads, and optimality. *Linguistic inquiry*, **28**, 373–422.

Gussenhoven, C. (2007). Types of focus in english. In C. Lee, M. Gordon, and D. B§ring, editors, *Topic and Focus: Cross-linguistic perspectives on meaning and intonation; Studies in Linguistics and Philosophy, vol. 82.* Springer.

Halliday, M. A. K. (1967). Notes on transitivity and theme. *Journal of Linguistics*, pages 199–244.

Halvorsen, P.-K. (1989). *The syntax and semantics of cleft constructions.* Ph.D. thesis, University of Texas at Austin.

Hamlaoui, F. (2007). French cleft sentences and the syntax-phonology interface. In *Actes du congrès annuel de l'Association canadienne de linguistique 2007*. Canadian Linguistic Association.

Hamlaoui, F. (2008). Focus, contrast, and the syntax-phonology interface: the case of french cleft-sentences. In *Current Issues in Unity and Diversity of Languages. Collection of the papers selected from the 18th International Congress of Linguists (CIL18)*, Seoul. Linguistic Society of Korea.

Hamlaoui, F. (2009). *La focalisation à l'interface de la syntaxe et de la phonologie: le cas du français dans une perspective typologique.* Ph.D. thesis, Paris 3 Sorbonne Nouvelle.

Heck, F., Müller, G., Vogel, R., Vikner, S., and Smichd, T. (2002). On the nature of the input in optimality theory. *The Linguistic Review*, **19**, 345–376.

Hedberg, N. (2000). On the referential status of clefts. *Language*, **76**, 891–920.

Heim, I. (1991). Articles and definiteness, published in german as Òartikel und definitheitÓ. In A. von Stechow and D. Wunderlich, editors, *Semantics: An International Handbook of Contemporary Research*, page 487Ð535. de Gruyter.

Hirst, D. and di Cristo, A. (1992). Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français. *TIPA*.

Holtgraves, T. (1997). Yes, but ...: Positive Politeness in Conversation Arguments. *Journal of Language and Social Psychology*, **16**, 222–239.

Hoot, B. (2012). Narrow focus on pre-nominal modifiers in spanish: An optimality-theoretic analysis. In K. Geeslin and M. Diaz-Campos, editors, *Selected Proceedings of the 14th Hispanic Linguistics Symposium*. Cascadilla Proceedings Project.

Horn, L. R. (1981). Exhaustiveness and the semantics of clefts. In V. Burke and J. Puste-jovsky, editors, *Proceedings of the Eleventh Annual Meeting of the North Eastern Linguistics Society*, page 125Đ142. University of Massachusetts, Amherst.

Hupet, M. and Tilmant, B. (1986). What are clefts good for? Some consequences for comprehension. *Journal of Memory and Language*, **25**, 419–430.

Hupet, M. and Tilmant, B. (1990). Effets du degré de conviction de l'interlocuteur sur l'usage de phrase clivées. *L'année psychologique*, **90**, 213–229.

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. MIT Press.

Jespersen, O. (1927). *A modern English grammar on historical principles. Part III: Syntax*. Heidelberg.

Jun, S.-A. and Fougeron, C. (2000). A phonological model of french intonation. In A. Botinis, editor, *Intonation: Analysis, Modeling and Technology*, pages 209–242. Kluwer Academic Publishers.

Jun, S.-A. and Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, **14**, 147–172.

Katz, S. (1997). *The syntactic and pragmatic properties of the* c'est-*cleft construction.* Ph.D. thesis, University of Texas at Austin.

Katz, S. (2000). Categories of *C'est*-cleft constructions. *Canadian Journal of Linguistics*, pages 253–273.

Katz, S. and Blyth, C. S. (2007). *Teaching French Grammar in Context: Theory and Practice*. Yale University Press, New Haven, CT.

Kayne, R. S. (1994). *The Antisymmetry of Syntaxs*. MIT Press, Cambridge, MA.

Keenan, E. (1976). Toward a universal definition of subject. In C. Li, editor, *Subject and Topic*, pages 303–333. Academic Press, NY.

Keenan, E. (1985). Relative clauses. In T. Shopen, editor, *Language Typology and Syntactic Description, vol. 2; Complex Constructions*. Cambridge University Press, Cambridge and New York.

Kenesei, I. (1986). On the logic of hungarian word order. In W. Abraham and S. de Meij, editors, *Topic, focus, and configurationality*, pages 143–159. John Benjamin.

Kiss, K. (2009). Structural focus and exhaustivity. In M. Zimmerman and C. Féry, editors, *Information Structure: Theoretical, Typological, and Experimental Perspectives*. Oxford University Press.

Kiss, K. E. (1998). Identificational focus versus information focus. *Language*, **74**, 245–273.

Koehn, P. (2005). Europarl: A Parallel Corpus for Statistical Machine Translation. In *Conference Proceedings: the tenth Machine Translation Summit*, pages 79–86. AAMT.

Kratzer, A. and Selkirk, E. O. (2007). Phase theory and prosodic spellout: the case of verbs. *Linguistic Review*, **24**, 93–135.

Krifka, M. (2007). Basic notions of information structure. In C. Féry and M. Krifka, editors, *Interdisciplinary Studies of Information Structure 6*. Postdam.

Ladd, D. R. (1996). *Intonational Phonology*. Cambridge University Press.

Lambrecht, K. (1986). Pragmatically motivated syntax: Presentational cleft constructions in spoken french. *CLS 22. Parasession on Pragmatics and Grammatical Theory*, pages 115–126.

Lambrecht, K. (1987). On the status of svo sentences in french discourse. In R. Tomlin, editor, *Coherence and grounding in discourse*, pages 217–262. Benjamins, Amsterdam.

Lambrecht, K. (1994). *Information structure and Sentence form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge University Press.

Lambrecht, K. (2000). When subjects behave like objects: An analysis of the merging of s and o in sentence focus constructions across languages. *Studies in Language*, **24:3**, 611–682.

Lambrecht, K. (2001). A framework for the analysis of cleft constructions. *Linguistics*, pages 463–516.

238

Lambrecht, K. and Michaelis, L. (1996). Toward a construction-based theory of language function: the case of nominal extraposition. *Language,* pages 215–247.

le Gac, D. and Yoo, H.-Y. (2002). Intonative structure of focalization in french and greek. In *Proceedings of the 14th Symposium on Romance Linguistics Going Romance 2000,* Utrecht, Netherlands. Benjamins.

Legendre, G., Wilson, C., Smolensky, P., Homer, K., and Raymond, W. (1995). Optimality and wh-extraction. In J. Beckman, L. W. Dickey, and S. Urbanczyk, editors, *Papers in Optimality Theory,* pages 607–636. University of Massachusetts, Amherst.

Legendre, G., Smolensky, P., and Wilson, C. (1998). When is less more? faithfulness and minimal links in wh-chains. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky, editors, *Is the best good enough? Optimality and competition in syntax,* pages 249–289. MIT Press, Cambridge, Massachusetts.

Martin, P. (1987). Prosodic and rhythmic structures in french. *Linguistics,* **25**, 925–949.

McCarthy, J. J. and Prince, A. S. (1993). Generalized alignment. In G. Booij and J. van Marle, editors, *Yearbook of Morphology 1993,* pages 79–153. Kluwer.

McCawley, J. (1989). Parentheticals and discontinuous structure. *Linguistic Inquiry,* **13**, 91–106.

Mikkelsen, L. H. (2002). Reanalyzing the definiteness effect: evidence from danish. *Working Papers in Scandinavian Syntax,* **69**, 1–75.

Moreau, M.-L. (1976). *C'est.* Editions Universitaires de Mons, Mons.

Nespor, M. and Vogel, I. (1986). *Prosodic Phonology*. Foris, Dordrecht.

Onea, E. and Beaver, D. I. (2008). Exhaustiveness of hungarian focus. experimental evidence from hungarian and german. In E. Onéa and A. Riester, editors, *Proceedings of Focus at the syntax semantics interface Stuttgart 2008*.

Percus, O. (1997). Prying open the cleft. In K. Kusumoto, editor, *Proceedings of the 27th Annual Meeting of the North-East Linguistics Society (NELS)*, pages 337–351.

Post, B. (1993). A phonological analysis of french intonation.

Potts, C. (2002). The syntax and semantics of *as*-parentheticals. *Natural Language and Linguistic Theory*, **20**, 623–689.

Potts, C. (2003). *The Logic of Conventional Implicatures*. Ph.D. thesis, PhD dissertation, University of California at Santa Cruz.

Potts, C. (2005). *The Logic of Conventional Implicatures*. Oxford University Press.

Prince, A. and Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers Center for Cognitive Science Technical Report.

Prince, E. (1981a). topicalization, focus movement and yiddish movement: a pragmatic differentiation. *BLS*, **7**, 249–264.

Prince, E. (1981b). Toward a taxonomy of given/new information. In P. Cole, editor, *Radical Pragmatics*, pages 223–255. Academic Press, New York.

Prince, E. (1983). A comparison of topicalization and left-dislocation in discourse.

Prince, E. F. (1978). A comparison of *it*-clefts and WH-clefts in discourse. *Language*, **54**, 883–906.

Reichle, R. V. (accepted). Cleft type and focus processing in french. *Language and Cognitive Processes.*

Rialland, A., Doetjes, J., and Rebuschi, G. (2002). What is focused in *c'est xp qui/que* cleft sentences in french? In B. B. et I. Marlien, editor, *Proceedings Speech Prosody 2002*, Aix en Provence.

Riegel, M., Pellat, J.-C., and Rioul, R. (2006). *Grammaire méthodique du francais.* Presses Universitaires de France, Paris.

Roberts, C. (1996). Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. *OSU Working Papers in Linguistics*, **49**. Papers in Semantics.

Roberts, C. (1998). Focus, the flow of information and universal grammar. *Syntax and Semantics*, **29**, 109–160.

Rochemont, M. (1986). *Focus in generative grammar*. John Benjamins, Amsterdam.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, **1**, 75–116.

Rooth, M. (1996). Focus. In S. Lappin, editor, *The Handbook of Contemporary Semantic Theory*, pages 271–297. Basil Blackwell, London.

Rossi, M. (1985). L'?intonation et l?'organisation de l?'énoncé. In *Phonetica 42*, pages 135–153.

Rossi, M. (1999). *L'intonation, le système du francais: description et modélisation.* Ophrys, Paris.

Samek-Lodovici, V. (2005a). Prosody-Syntax Interaction in the expression of focus. *Natural Language and Linguistic Theory*, **23**, 687–755.

Samek-Lodovici, V. (2005b). Prosody-syntax interaction in the expression of focus. *Natural Language and Linguistic Theory*, **23**, 687–755.

Schwarschild, R. (1999). Givenness, avoidf and other constraints on the placement of accent. *Natural Language Semantics*, **7**, 141–177.

Selkirk, E. O. (1978). On prosodic structure and its relation to syntactic structure. In T. Fretheim, editor, *Nordic Prosody II*, Trondheim: TAPIR.

Selkirk, E. O. (1984). *Phonology and Syntax.* MIT Press, Cambridge.

Selkirk, E. O. (1986). On derived domains in sentence phonology. In *Phonology Yearbook, vol. 3*, pages 371–405.

Selkirk, E. O. (1995). The prosodic structure of function words. In *Papers in Optimality Theory*, pages 439–470. GLSA, University of Massachusetts, Amherst.

Selkirk, E. O. (1996). Sentence prosody: Intonation, stress, and phrasing. In J. A. Goldsmith, editor, *The Handbook of Phonological Theory*, pages 550–569. Basil Blackwell, London.

Selkirk, E. O. (2000). The interaction of constraints on prosodic phrasing. In M. Horne, editor, *Prosody: Theory and Experiment*. Kluwer Academic Publishers.

Selkirk, E. O. (2005). Comments on intonation phrasing in english. In S. Frota, M. Vigario, and M. J. Freitas, editors, *Prosodies*. Mouton de Gruyter, Berlin.

Selkirk, E. O. and Shen, T. (1996). Prosodic domains in shanghai chines. In *The phonology-syntax connection*, pages 313–337. University of Chicago Press, Chicago.

Spector, B. (2006). *Aspects de la pragmatique des opérateurs logiques.* Ph.D. thesis, Université Paris 7 - Denis Diderot, Paris, France.

Stalnaker, R. (1974). Pragmatic presuppositions. In M. Munitz and U. P., editors, *Semantics and Philosophy*. New York University Press.

Strawson, P. F. (1950). On referring. *Mind*, **59**, 320–344.

Szabolcsi, A. (1981). The semantics of topic-focus articulation. In T. J. Jan Groenendijk and M. Stokhof, editors, *Formal methods in the study of language*, pages 513–541. Matematisch Centrum.

Truckenbrodt, H. (1995). *Phonological Phrases: their relation to syntax, focus and prominence.* Ph.D. thesis, PhD dissertation, MIT.

Truckenbrodt, H. (1999). On the relation between Syntactic phrases and phonological phrases. *Linguistic Inquiry*, **30**, 219–255.

Truckenbrodt, H. (2005). A short report on intonation phrase boundaries in German. *Linguistische Berichte*, **203**, 273–296.

Vallduví, E. (1992). *The Informational Component*. Garland Publishing, New York, NY.

Vion, M. and Colas, A. (1995). Contrastive marking in french dialogue: Why and how. *Journal of Psycholinguistic Research*, **24**, 313–331.

von Fintel, K. (2004). Would you believe it? the king of france is back! (presuppositions and truth-value intuitions). In A. Bezuidenhout and M. Reimer, editors, *Descriptions and Beyond: An Interdisciplinary Collection of Essays on DeÞnite and IndeÞnite Descriptions and other Related Phenomena*, pages 315–341. Oxford University Press, Oxford.

von Stechow, A. (1991). Current issues in the theory of focus. In A. von Stechow and D. Wunderlich, editors, *Semantik*, pages 804–825. Walter de Gruyter.

Wedgwood, D. (2005). *Shifting the Focus. From static structures to the dynamics of interpretation.* Elsevier, Amsterdam.

Zerbian, S. (2006). Expression of information structure in the bantu language northern sotho. *ZASPiL*, **45**.

Zimmerman, M. (2008). Contrastive focus and emphasis. In *Acta Linguistica Hungarica*, pages 347–360.

Zimmerman, M. and Onea, E. (2011). Focus marking and focus realization. *Lingua*, **121**, 1651–1670.

Zimmerman, M., Fiedler, I., Hartmann, K., Reineke, B., and Schwarz, A. (2010). Subject focus in west african languages. In M. Zimmerman and C. Féry, editors, *Information Structure. Theoretical, Typological, and Experimental Perspectives*, pages 234–257. Oxford University Press.

Zimmermann, M. (2006). Focus in western chadic: A unified ot account. In *Proceedings of NELS 36*.

Zimmermann, M. and Drenhaus, H. (2009). How exhaustive are you? an ERP study on it-clefts, only-foci, and scalar implicature.

Zubizarreta, M. L. (1994). On some prosodically governed syntactic operations. In J.-Y. P. L. R. Y. Cinque, J. Koster and R. Zanuttini, editors, *Path Toward Universal Grammar. Studies in Honor of Richard S. Kayne*, pages 473–485. Georgetown University Press, Washington DC.

Zubizarreta, M. L. (1998). *Prosody, focus and word order*. MIT Press, Cambridge, MA.