Journal of French Language Studies **26** (2016), 299–326, © Cambridge University Press
doi:10.1017/S0959269515000265

Focus marking asymmetries in Colloquial and Standard French: A stochastic optimality-theoretic account

EMILIE DESTRUEL

University of Iowa

(Received June 2014; revised April 2015)

ABSTRACT

This article investigates the grammatical realization of the notion of focus in Colloquial French and Standard French. Based on two production experiments, the article reveals three findings: (i) focus marking is not as categorical as previously acknowledged, (ii) focus marking asymmetry for subjects vs. non-subjects is only supported in CoF and (iii) there is no strict relationship between focus realization and interpretation in either variety. I develop a stochastic optimality-theory analysis, which explains the canonical-cleft sentence alternation in terms of prosody and expands on past literature by accounting for the variation observed both within and across language variety.

1 INTRODUCTION

It is widely acknowledged that natural languages signal *focus* – the part of the sentence that evokes a set of (explicit) alternatives that the speaker takes to be salient in the context (Rooth, 1992; Krifka, 2008) – via different means. While English can shift prosodic prominence to match the position of focus, French is considered a less plastic language and is assumed to require the use of special syntactic constructions such as the *c'est*-cleft; a bi-clausal structure that expresses a single proposition into two separate clauses: a matrix clause and a relative-like clause (Lambrecht, 2001). This requirement is, moreover, argued to be mandatory under two specific contexts, illustrating two focus asymmetries: if the focus element is a grammatical subject (1), but not if it is a non-subject (*argument* asymmetry), and if the focus element conveys a strong interpretation such as a correction or a contrast (2), but not if it simply provides information to a wh-question (*focus type* asymmetry)

I would like to thank Fatima Hamlaoui, Leah Velleman and David Beaver for detailed discussion on previous versions on this paper, and the anonymous reviewers for their insightful comments. All remaining errors are mine.

(Vion and Colas, 1995; Katz, 1997; Hamlaoui, 2009; Skopeteas and Fanselow, 2010; Belletti, 2012). [1]

(1) *Context:* Qui est-ce-qui a ri?
  a. #[Jean]$_f$ a ri.
  b. C'est [Jean]$_F$ qui a ri.

(2) *Context:* Est-ce-que Jean a invité Paul?
  Non, c'est [Marc]$_F$ qu'il a invité.

To account for these marking patterns, many scholars have offered a categorical treatment of focus realization in French, treating the competition between canonical and cleft sentences as producing a single best output form given the grammatical function of the focused element and the focus-type being encoded (Lambrecht, 1987, 1994, 2001; Vion and Colas, 1995; Katz, 1997; Hamlaoui, 2009; Belletti, 2012), with Lambrecht even questioning whether SVO sentences do in fact constitute the canonical configuration in the language. Interestingly, the majority of these studies have either not used empirical methods to support their claims, or have used grammaticality judgements, and tend to restrict their examination to the colloquial (or spoken, demotic) variety of the language. But French is well-known for the significant differences that exist between its colloquial and standard variety (de Cat, 2007; Klein, 2012). In fact, the differences are such that certain scholars posit a diglossic analysis of the language, claiming that native speakers in fact possess two separate yet intersecting grammars (Massot, 2008; Zribi-Hertz, 2011, 2013).

In light of these observations, the present article sheds light on the following questions:

Q1: How categorical are the two focus asymmetries in colloquial French (CoF)? Is there indeed a strict one-to-one mapping between the focus element encoded and its realization in the grammar? And similarly, is there a strict one-to-one mapping between the type of focus expressed and its realization in the grammar?

Q2: Do native speakers of French systematically use different strategies to realize focus across language variety? Put slightly differently, are the two focus asymmetries also observed in standard French (StF)?

Q3: If variation and optionality occur in the data, both within and across language variety, how can this be explained?

The intended contribution of this article is both methodological and theoretical. From the methodology perspective, it provides empirical data from two semi-spontaneous production studies designed to test the effect of focus type and grammatical constituent on the way information-structure interacts with syntax in the expression of focus, for two language varieties, namely colloquial French (CoF) and standard French (StF). Results suggest, contra previous categorical accounts

[1] Throughout the article, the asterisk is used to indicate ungrammaticality, and the hash mark to indicate semantic or pragmatic infelicity. Focused elements are indicated via square brackets with a subscript F.

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

(Samek-Lodovici, 2005; Hamlaoui, 2009), that variation does occur, both within and across varieties: CoF allows non-subjects to be realized in a cleft, StF allows subjects to be realized *in situ*, and a corrective focus does not systematically trigger the use of a cleft. From the theoretical perspective, I explain these findings by proposing a prosodic analysis of focus marking couched in a stochastic optimality-theoretic framework (Boersma and Hayes, 2001). Under this analysis, CoF and StF possess different, yet intersecting grammars that are governed by similar constraints, which are ranked differently, thus providing speakers with different structural strategies to mark focus. The advantage of such an analysis is that it captures the variation observed in the data (i.e., variable outputs for the same input), and allows for a better comparison between the two language varieties.

The remainder of the article is structured as follows: Section 2 motivates the study. Section 2.1 defines the two language varieties under study and summarizes how the past literature has accounted for the differences. Sections 2.2 and 2.3 discuss the previous analyses of focus in French, analyzing how the two asymmetries have been treated. Section 3 and Section 4 present the experimental paradigm, discuss its limitation, and report on the results for CoF and StF respectively. Section 5 offers a general discussion about the findings, which are analyzed within a stochastic optimality theoretic framework in Section 6. Section 7 concludes.

2 BACKGROUND

2.1 Two language varieties

One undisputed characteristic of the French language lies in the divergence between its standard and its colloquial variety (see Benveniste, 1997; Katz and Blyth, 2007; and Hamlaoui, 2010, among others).[2] Standard French (StF) – or formal, traditional French – refers to the highly codified variety of French, which is learnt late in a more formal context and under the influence of literacy, for instance by being taught in schools and advocated by prescriptive grammars. While this variety is mostly present in the written modality, it is certainly not restricted to it, and also includes the spoken language of the educated used in formal settings such as, for example, in the media or in speeches. The standard variety is also preponderant in foreign language classrooms to meet teachers' normative expectancies. Indeed, Blyth (1999: 156) points out that 'relatively few foreign language materials make extensive use of authentic interaction [. . .] In fact, textbooks frequently fail to even mention or exemplify constructions that are prevalent in the spoken language'. Thus, language learners of French will often produce sentence structures that are pragmatically odd or sociolinguistically inappropriate because of the divide between the two language varieties. On the other hand, Colloquial French (CoF) – or demotic, informal French – is understood as the speech used by most native speakers in their everyday

[2] The divide between written and spoken language is also noted in languages such as English, and is discussed, with special emphasis on the use of cleft sentences, in Roland et al. (2007) and Calude (2009).

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

life interactions, which is acquired early, well and in a naturalistic environment. Here as well, CoF may not be restricted to the spoken medium, but can also be found in the written modality; in informal pieces such as comics that aim to simulate the colloquial language.

Researchers have identified several features that are characteristic of one or the other variety. Massot (2008) catalogues numerous areas of variation, including for example the negation particle '*ne*', which absence in speech is an extremely familiar aspect of colloquial French. Most relevant to the studies presented in this article is the complex system of interrogative forms, which different forms are governed by an intricate web of factors (see Myers, 2007), one of them being *formality* (Quillard, 2001; Coveney, 2002; Myers, 2007). Indeed, it is typically assumed that, for partial questions that include a wh-word and question only a portion of the proposition, an inverted form (i.e., *wh*-fronting with subject-verb-inversion) is restricted to StF and is still prevalent, for better or worse, in foreign language textbooks (Etienne and Sax, 2009). Other question forms, notably *est-ce que*, SVQ and QSV belong to colloquial speech (see Zribi-Hertz, 2013 for a more exhaustive list of morphological, phonological, syntactic and lexical properties that distinguish the two language varieties).

The differences between the two varieties are such —- often fundamental and systematized —- that the question about whether speakers actually possess two grammars has been raised, and caused much debate in the past decades. Variationist accounts, deriving from the work by Labov (1969; 1972), argue against this claim, defending that grammars of natural languages are inherently variable (Hornsby, 1998; Coveney, 2011; Rowlett, 2013). Other scholars posit that French speakers do not combine socio-stylistic features from StF and CoF and propose an explanation in terms of diglossia, arguing that speakers possess two overlapping, yet distinct, co-existing grammars. In the present article, I will side with the latter, taking clefts to be part of the CoF grammar and canonical sentences part of StF.

2.2 *Subject vs. non-subject focus marking*

The seminal work of Lambrecht (1987, 1994, 2001) claims that, although grammatically well-formed, canonical sentences (Subject-Verb-Object) are extremely rare in Colloquial French. The culprit of this infelicity is taken to be a restriction on the preverbal position: subject focus is categorically banned from appearing *in situ*, as seen in (1a). Non-subject focus, on the other hand, is generally realized in its default position. This apparent *argument asymmetry* is empirically documented in a few studies: Vion and Colas (1995) report on a forced-choice task where native speakers judge clefts as more natural with subjects than with objects, Reichle (2014) provides data from a quantitative study in the Europarl corpus (Koehn, 2005) demonstrating that subject clefts are significantly more frequent than object clefts, and more recently Féry (2013) presents a production task in

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

which participants never produce clefts with object focus.[3] Two questions arise: (i) what is the strategy used in CoF to cope with the apparent restriction on grammatical subjects? and (ii) how can the argument asymmetry be explained?

Let me address the first question. Unlike other Romance languages like Italian that can move a 'bad' subject to the rightmost position (3a), CoF strictly forbids an empty subject position (3b). In optimality-theoretic terms, this is seen as fulfilling the constraint SUBJECT in (4), similar to the EPP (Extended Projection Principle) hypothesis about the obligatoriness of subjects (Chomsky, 1982).

(3) Question: Qui a ri?
  a. *Italian*: Ha riso [Gianni]$_F$
  b. *French*: *a ri [Jean]$_F$

(4) SUBJECT: Highest A-Specifier must be filled.

Instead, CoF resorts to non-canonical word orders, and specifically the *c'est*-cleft (5).

(5) C'est [Jean]$_F$ qui a ri.

The fact that a more complex syntactic structure is favoured over a simpler canonical sentence shows that a constraint like *STRUCTURE (6) (adapted from Grimshaw, 1997), which prohibits the addition of syntactic structure not present in the input (Hamlaoui, 2009), is not crucial in CoF and can be violated.

(6) *STRUCTURE: Avoid structure not present in the input.

But crucially, Lambrecht's assumption concerning subject focus marking is disputed. In fact, some researchers take a canonical sentence like (1a) to be completely acceptable, and argue that subject focus can also be marked via prosodic means (Marandin, 2004; Delais-Roussarie, 2005). Féry (2001) is one of the few studies that have empirically tested this issue. Her article reports on a production task in which ten native speakers read the answering component of a question-answer pair. Although given the option to change the wording if judged unnatural, participants marked focus on grammatical subjects via prosodic phrasing (i.e., by realizing the subject in its own prosodic phrase), retaining the canonical order in most cases. Of course, one limitation of this study is that the participants are given scripts for the sentences to produce. One reason why they choose not to vary the sentence structure and produce a cleft instead could simply be that they are being polite and obedient. Yet despite this potential methodological problem, the study has the advantage to show that prosody is sensitive to information-structure in French – a conclusion now commonly acknowledged despite the debate on the exact prosodic correlates of focus. Similarly, Belletti (2009: 248) points out that French speakers do not totally exclude SV answers similar to the English type (although no empirical data is provided). In another study, although not reported in the discussion of their cross-linguistic experimental study, Skopeteas and Fanselow

[3] The same asymmetry is observed cross-linguistically in Spanish (Büring and Gutiérrez-Bravo, 2001), Italian (Féry, 2013), West Chadic languages (Zimmermann, 2006), Georgian, Hungarian (Skopeteas and Fanselow, 2010) and in Northern Sotho (Zerbian, 2007).

(2010: 187) also find that native speakers produce focused subjects in informational contexts in their canonical sentence-initial position at a significant rate (45.3% of the time). Finally, Samek-Lodovici (2005: 728) develops an optimality-theoretic account where *in situ* realization is derived as the optimal output for subject focus.

This shows that the expression of subject focus may be more complicated than previously acknowledged and that more thorough empirical work is still needed to understand the exact conditions under which clefting is required.

With regards to the second question on how to account for the marking asymmetry, the past literature has offered two main analyses. First, scholars like Lambrecht (1994, 2001) propose an account in terms of mapping between syntax and information-structure (see also Katz, 1997). The basic premise for this proposal is the observation that languages tend to reserve the preverbal position for topical material – material which is more readily accessible in the discourse (Chafe, 1976; Keenan, 1976). This idea is implemented in Aissen (1999) through the harmonic alignment of two scales, (i) a relational scale (Subject > Non-subject) and (ii) a focus scale (Topic > Focus), resulting in the generalization that foci in preverbal subject position are not harmonic, or in other words that subjects with a low thematic prominence (i.e., which are discourse new) should be avoided. In the same vain, Lambrecht (2001: 466) claims that 'clefts serve to mark as focal an argument that might otherwise be construed as nonfocal, or as nonfocal a predicate that might otherwise be construed as focal, or both'. This claim is formalized in Zerbian (2007), who accounts for a similar restriction in Northern Sotho, with the constraint SUBJ = TOPIC (7) encoding pressure on subject choice grounded in discourse prominence.

(7) SUBJ = TOPIC: The grammatical subject of the sentence must not be F-marked.

In contrast, the second analysis is framed in purely prosodic terms (Hamlaoui, 2009). Unlike Lambrecht's assumption that the cleft is a construction per se and therefore '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' (2001: 466), Hamlaoui follows Clech-Darbon et al. (1999) in positing a maximally simple structure: the cleft is made up of two propositions in which the focused constituent is directly merged as the complement of an identificational TP to which a CP is right adjoined. This structure is then mapped onto prosody via stress assignment rules that follow Selkirk's (1984) hierarchy (Intonational Phrase (ι) > Phonological Phrase (ϕ) > Prosodic word (ω)), and by which a strong label is assigned to the rightmost element at each level. The prosodic structure in Figure 1 is derived, with main stress occurring rightmost at the ι-level.

In this syntax-prosody mapping, the matrix and the relative-like clause each form an independent ι-phrase, and the entire sentence also forms one. Thus, the right-edge of the first ι-phrase (formed by the matrix clause) aligns with the post-copular element. This is a welcome result since it fulfills the language-specific requirement

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

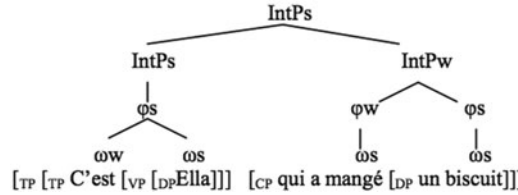

Figure 1. *Cleft syntactic structure mapped onto prosody, adapted from Hamlaoui, 2007)*

that demands main stress to occur at the right-edge of a ι-phrase (Jun and Fougeron, 2000), expressed via the constraint in (8) from Samek-Lodovici (2005).

(8) HEAD-I-R (HIR): Align the right boundary of every intonation phrase with its head.

Assuming that a focus constituent must receive prosodic prominence under the effect of the constraint in (9) (see also the Stress-Focus constraint in Truckenbodt, 1995 and in Szendröi, 2003, as well as the related Focus-Prominence constraint in Büring, 2009), Hamlaoui claims that no movement is necessary (*contra* Lambrecht, 2001), and that in fact, a subject focus is directly merged into the position where grammar assigns main stress, viz. the post-copular position. Put slightly differently, a subject focus is *not* realized *in situ* because the preverbal position does not match the position in which main stress is assigned by default in the language.[4] Under this view, an economy constraint that prohibits syntactic movement like STAY in (10) (see Grimshaw (1997) and Samek-Lodovici (2005) for this constraint) is not required since no movement is assumed.

(9) STRESS-FOCUS (SF): A focused phrase has the highest prosodic prominence in its focus domain.

(10) STAY: No traces.

Under Hamlaoui's account, the fact that non-subjects are focused *in situ* is simply explained: non-subjects are aligned by default with the right-edge on an ι-phrase, therefore there is no need for a special syntactic structure. In this article, I adopt her prosodic analysis but expand on it in two ways: first, by investigating the realization of subject and non-subject focus in StF, and second by showing that corrective focus (discussed in the following section) does not always induce the use of a *c'est*-cleft.

2.3 *Informational vs. corrective focus marking*

Within the literature on focus, researchers have generally distinguished between two major pragmatic types of focus (see Kiss, 1998 and Krifka, 2008 among many others). On the one hand, instances of *informational* focus simply express the non-presupposed information in a sentence; their sole function is to highlight the part

[4] See Féry (2013) for a similar account without reference to prosodic prominence, but purely in terms of alignment.

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

of an answer that corresponds to the open variable instantiated by a congruent *wh*-question. On the other hand, instances of *identificational* or *corrective/contrastive* focus involve the exclusion of at least one relevant (and salient) focus alternative, and thus carry a stronger pragmatic interpretation of focus.

Certain scholars, and most notably Kiss (1998), argue that these two focus-types are fundamentally distinct: they must be systematically realized in different ways in the syntax. Identificational focus must be realized by means of a special marking, whether it is via syntactic, prosodic or morphological means. In English for example, this special marking is taken to be the *it*-cleft (Kiss, 1998). On the contrary, this is not required for informational focus, which is marked *in situ*. This *focus-type* asymmetry has the predictive power that no grammar should require a special grammatical marking for informational focus and not for identificational focus.

In the literature on French, the *c'est*-cleft is also assumed to convey a stronger focus interpretation (Lambrecht, 1994; Katz, 1997; Clech-Darbon et al. 1999; Belletti, 2005; De Cat, 2007). Thus, researchers have argued that focus-type has an effect on word order. This is mostly observable when non-subjects are focused since grammatical subjects are clefted on independent grounds (see Section 2.1): Non-subjects are realized *in situ* (11) except when conveying contrast/correction (see inference in 12b), in which case a cleft is required (12a).

(11) Informational non-subject focus:
Marie a embrassé [Jean]$_F$

(12) Identificational/Corrective non-subject focus:
a. C'est [Jean]$_F$ que Marie a embrassé.
b. '*John and no one else/ John, as opposed to Paul, was kissed by Mary.*'

But one caveat immediately appears. Indeed, if the focus-type asymmetry is understood in a categorical sense with a strict one-to-one relationship between the grammatical realization of focus and its interpretation, the *c'est*-cleft should not be used to signal instances of informational focus and should be reserved for stronger focus interpretations. However, as described in Section 2.1, clefts are also used in informational contexts, i.e. with grammatical subjects. Therefore, the following question arises: how can the difference between subject and object clefts be explained, or put slightly differently, how can the different interpretative possibilities the cleft exhibits be analyzed?

Belletti (2005, 2009, 2012) proposes a cartographic account in which she argues for the existence of two different Focus positions in the clausal map: 'the vP peripheral low Focus position in the matrix clause vP periphery of the copula, in the case of new information subject clefts' and 'the high left peripheral position in the reduced CP complement of the copula, in the case of contrastive/corrective object clefts – and also in subject clefts when they are interpreted/used contrastively/correctively [. . .]' (2012: 96). Therefore, under this view, the interpretational effects associated with the cleft derive directly from the syntactic position in which the focus is realized. In other words, the focus-type

asymmetry is explained by positing that specific positions in the functional structure are dedicated to the expression of different interpretations. Thus, the cleft is the optimal candidate to express instances of identificational focus.

However, there is a vast literature showing that there are clear prosodic correlates of contrastive focus in the language. Di Cristo (1998), Rossi (1999) and Jun and Fougeron (2000), to cite only a few, find that contrastive focus is characterized by a large and sharp rise in F0 and/or intensity on the focused constituent, an increased duration of the focused syllables and a global F0 and intensity compression in the post-focus sequence, with either a low plateau or a steady fall until the end of the utterance. In an unpublished study testing the role of prosody on information structure in different context types, Vander Klok et al. (manuscript) found that 'a shift occurs only when the antecedent is in a separate speech act that stands in rhetoric contrast to the present one, as is the case in a direct correction'. So the question as to whether a stronger use of focus necessarily triggers a cleft construction remains open, and the experiments presented in the following section aim to provide new empirical data to answer it.

3 EXPERIMENT 1: FOCUS MARKING IN COLLOQUIAL FRENCH

3.1 Participants

Twenty-one native speakers of Continental French (8 males and 13 females) from different regions in France (Paris, Lyon, Toulouse) who had lived in the United States for less than six months and were naïve as to the purpose of the task participated. Their age ranged between 28 and 42 years.

3.2 Procedure and material

The present experiment is designed to test how speakers realize different focus types on different grammatical constituent in CoF. For this task (adapted from Gabriel, 2010), participants sat in front of a computer screen where they saw a series of pictures presented as a PowerPoint slideshow. Each picture was accompanied by a short one-sentence description of the scene that appeared in written form under the visual display and provided participants with the relevant information to answer subsequent questions (Figure 2).

(13) Picture description:
Marie est allée au kiosque et elle a acheté un journal.

After carefully reading the description, participants pressed the return key to advance to the next slides, which contained a series of questions delivered one at a time in written form. The purpose of these questions was to ensure that the target sentences were produced with the intended information-structural meanings, thus eliciting a particular focus-type (informational vs. corrective) for a given grammatical function (subject vs. non-subject focus), giving a 2 × 5 Latin square design. *Informational* focus was triggered by a simple

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

Figure 2. *Sample pictorial stimulus*

wh-question, whereas *corrective* focus involved a preceding statement containing a faulty description and the question-tag '*no*?', making up experimental paradigms (14) and (15).

(14) Subject informational focus condition:
Qui est-ce-qui a acheté un journal au kiosque?
(15) Subject corrective focus condition:
On dirait que Jean a acheté un journal, non?

One clarification is in order: the condition 'corrective sentence-focus' was left out of the design as it seemed strange to have a context where the constituent to correct was the entire sentence. This does not have a negative impact on the experiment itself since the conditions that are most relevant to the research questions concern narrow focus on subjects vs. objects. Therefore, the final design involved nine conditions instead of ten. For each condition, five lexicalizations were created, leading to a total of 45 experimental items, which were pseudo-randomized with 20 distractors.

Finally, before the experiment started, participants were instructed to answer questions as if answering to a friend to whom they typically talk informally. Moreover, they were asked to avoid answering with single constituent fragments, but, as much as possible, to try and use full sentences. The instruction did not specify whether the answers needed to include given material in their lexical form so that speakers would feel free to use pronouns. Yet, instructions emphasized the importance to answer as naturally as possible. Participants delivered their responses orally and were recorded with the program Audacity. Trials where participants did not correctly identify the focused element, producing a mismatched answer, were excluded from the results.

Although this method has been largely used in prior elicitation tasks on focus (see for example Gabriel, 2010; Skopeteas and Fanselow, 2010 and Féry, 2013),

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

Table 1. *Colloquial French data set*

	Informational focus				Corrective focus			
	Canonical		Cleft		Canonical		Cleft	
	n	%	n	%	n	%	n	%
Subject	19	18.1	86	81.9	10	9.5	95	90.5
Dir.object	88	83.8	17	16.2	70	66.7	35	33.3
Ind.object	95	90.5	10	9.5	71	67.6	34	34.2
Predicate	105	100	0	0	105	100	0	0
Sentence	105	100	0	0	n/a	n/a	n/a	n/a

a cautionary note is necessary. Full sentences are not typically used in natural discourse, and ensuring that participants use such forms may have led to elicited data that contains a biased, non-optimal variant of participant's speech. This is indeed a potential drawback that will need to be addressed in future work. The present experiment constitutes a first step in the exploration of focus-marking strategies in different varieties of French that will certainly benefit from further empirical work exploring more naturally-occurring data (i.e., corpus studies and other elicitation methods like semi-directed interviews) to corroborate the current findings.

3.3 Results

The data collected were all analyzed using linear mixed-effects regression models implemented with the *lmer* program (Bates, Maechler and Boekler 2012) in R. In the models, the dependent variable *answer* (the form of the sentence produced by participants) is predicted by two independent variables, *grammatical function* and *focus-type*. Two random-effects predictors, participant ID and lexical item, were also included in the model to account for the variability in participants' performance and experimental item seen. To assess whether inclusion of a given predictor significantly improved the fit of the overall model, likelihood-ratio tests were performed that compared two minimally different models, one with the fixed effects predictor in question and one without, while keeping the random effects structure identical.

The results for CoF are presented in two ways: Table 1 presents raw counts and percentages, Figure 3a presents raw counts for the informational condition and Figure 3b for the corrective condition.[5]

First, I concentrate on the factor grammatical function, collapsing results for focus-type. As predicted, there is a main effect of this factor on the strategy used in the answer: the difference in distribution of answer form produced across the

[5] Note that the raw numbers for the condition Sentence focus on Figure 3b is smaller because this condition only appeared for the informational condition.

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

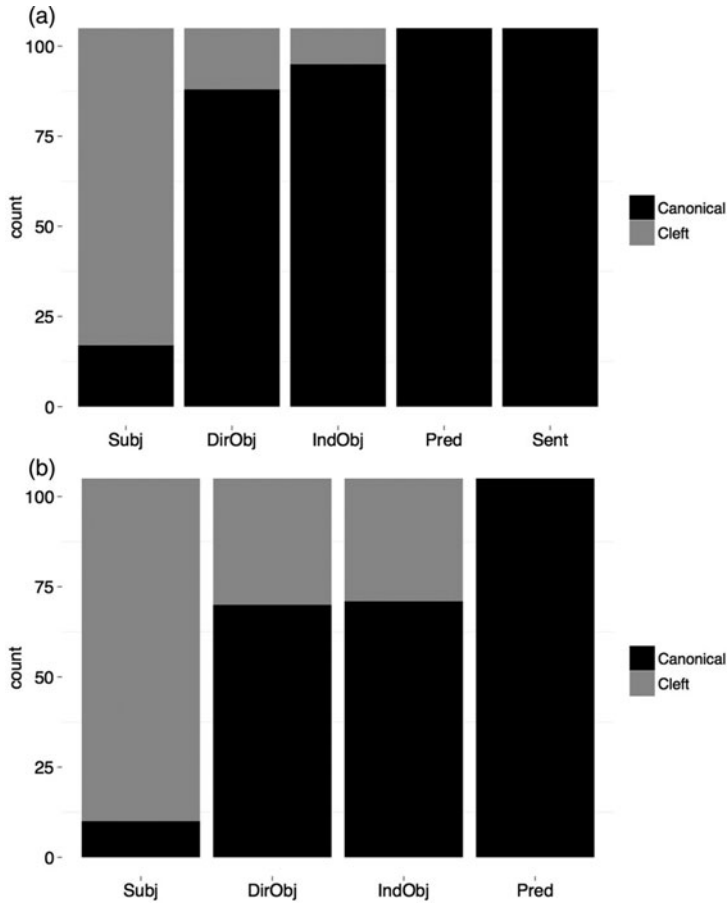

Figure 3a. *(a) Count per grammatical function in informational context in CoF. (b) Count per grammatical function in corrective context in CoF*

five grammatical functions is highly significant ($\chi^2(5) = 502.41$, $p < 0.0001$), and a model that includes the factor gives a much better fit to the data than one without it ($\chi^2(4) = 104.44$, $p < 0.0001$). However, almost all the variation can be traced back to the difference between grammatical subjects vs. non-subjects. This result clearly substantiates the presence of an argument asymmetry in CoF: participants used significantly more clefts when focusing a subject ($z = 14.83$, $p < 0.01$) and significantly more canonical sentences when focusing a non-subject ($z = 10.34$, $p < 0.01$ for direct objects and $z = 11.9$, $p < 0.01$ for indirect objects).

Second, let us turn to the results by focus-type. At first view, the number of clefts produced increases in corrective contexts, both for subjects and non-subjects. Yet,

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

when turning to the mixed-model regression analysis, the hypothesis according to which a stronger focus interpretation has an impact on syntax and requires a special marking such as the cleft (Zimmermann, 2006) is not confirmed. Indeed, a model including focus-type as a predictor performs only marginally better than a model without it ($\chi^2(4) = 2.89$, $p = 0.08$). Put slightly differently, participants were found to produce only slightly more clefts when expressing a correction than when marking informational focus ($\beta = -0.17$, SE = 0.10, $t = -1.69$, $p = 0.064$). In fact, the best fit to the data is given by a model that includes the interaction between the two predictors, grammatical function and focus-type ($\chi^2(4) = 110.36$, $p < 0.0001$), but here again, most of the variation is attributable to grammatical subjects.

4 EXPERIMENT 2: FOCUS MARKING IN STANDARD FRENCH

The second experiment investigates the effect of focus type and grammatical function on the realization of focus in Standard French. It uses the exact same experimental design as the one described in Section 3. However, in order to create a more formal setting, a few changes were made to the stimuli and the instructions, as described in Section 4.2.

The predictions for this study are as follows: If the restriction on subject focus occurring in preverbal position is specific to the colloquial grammar, canonical sentences should emerge in StF in response to a '*Qui/Who*' question, and clefting should be optional. Moreover, if there is a strict one-to-one mapping between a marked focus realization and a stronger focus interpretation, clefts should appear in corrective contexts, even in cases where the focus is a non-subject, and canonicals should be reserved for instances of informational focus.

4.1 Participants

Twenty-one French native speakers (11 males and 10 females), distinct from those who participated in the CoF experiment and who had lived in the United States for less than six months, took part in the study. Their age ranged between 22 and 53 years old. They were also naïve as to the purpose of the experiment and had uncorrected vision.

4.2 Procedure and material

For the purpose of the present experiment, I implemented three changes to trigger a more standard form of the French language. First, the instructions differed from the experiment on CoF in that it was emphasized that participants had to imagine a setting in which the questions were asked by a hierarchical superior to whom they generally respond in a formal way. Second, the form of the question was an inverted question (16a) as opposed to an *est-ce-que* question in the CoF experiment

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

and third, the sentence ended with a question-tag in corrective contexts '*n'est-ce pas?*' (16b).

(16) Sample stimulus in Standard French:
  a. Où Marie a-t-elle acheté un journal?
  b. Il me semble que Marie a acheté un journal au supermarché, n'est-ce-pas?

Eliciting the standard variety of the language while keeping the experimental design parallel to that in Experiment 1, for comparative purposes, is not a trivial task. The context that triggers different experimental conditions must be as systematically controlled as possible, differing just in the specific independent variables investigated (i.e., here, focus type and grammatical function of the focused element). In the present experiment, there is little 'room' for manipulations, which are restricted to adapting the context in which the question is asked, as well as the form of the question itself. Therefore, it can be questioned whether the changes adopted are significant enough to trigger a different language variety and if the speakers will respond using the same language variety than the one intended by the experimental design. Moreover, the features adopted to prompt StF could be debated as they may only represent the language that is taught as a stylistic instruction, rather than what speakers actually produce when found in authentic situations. Due to these potential weaknesses, the results collected with the current method will need to be further confirmed via more naturalistic observations.

4.3 Results

Results show an effect of grammatical function in the predicted direction: participants produced different answer forms according to the grammatical function of the focused element, with the largest difference being observed between grammatical subjects and non-subjects, the former being clefted systematically more than the latter ($\beta = 0.49$, SE $= 0.05$, $t = 10.80$, $p < 0.001$). Similar to CoF, a model with the predictor grammatical function performed better than a model without it ($\chi^2(4) = 85.32$, $p < 0.0001$). Within the subject condition, clefts were produced significantly more than canonical sentences ($z = -3.71$, $p < 0.001$). However, this result is only provisional since most clefts produced are actually found in the corrective condition and the difference between the two sentence forms is trivial in the informational context, as discussed hereafter. Regarding non-subjects, *in situ* realization is greatly preferred and unsurprisingly the difference in answer form is significant ($z = -16.98$, $p < 0.01$ and $z = -18.54$, $p < 0.01$, respectively).

I now turn to analyzing the data per focus-type. Table 2 reports on the raw counts and percentages of answers given per grammatical function and focus-type. Figure 4a and 4b present the data visually for the informational and the corrective context, respectively.

Concentrating on the informational condition first, results show that StF allows grammatical subjects to be focused *in situ* as well as in a cleft. The difference between the two forms is negligible ($z = 1.79$, $p = 0.07$). The prediction that canonical

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

Table 2. *Standard French data set*

	Informational focus				Corrective focus			
	Canonical		Cleft		Canonical		Cleft	
	n	%	n	%	n	%	n	%
Subject	59	56.2	46	43.8	27	25.7	78	74.3
Dir.object	102	97.1	3	2.9	87	82.8	18	17.2
Ind.object	103	98	2	2	92	87.6	13	12.4
Predicate	105	100	0	0	105	100	0	0
Sentence	105	100	0	0	n/a	n/a	n/a	n/a

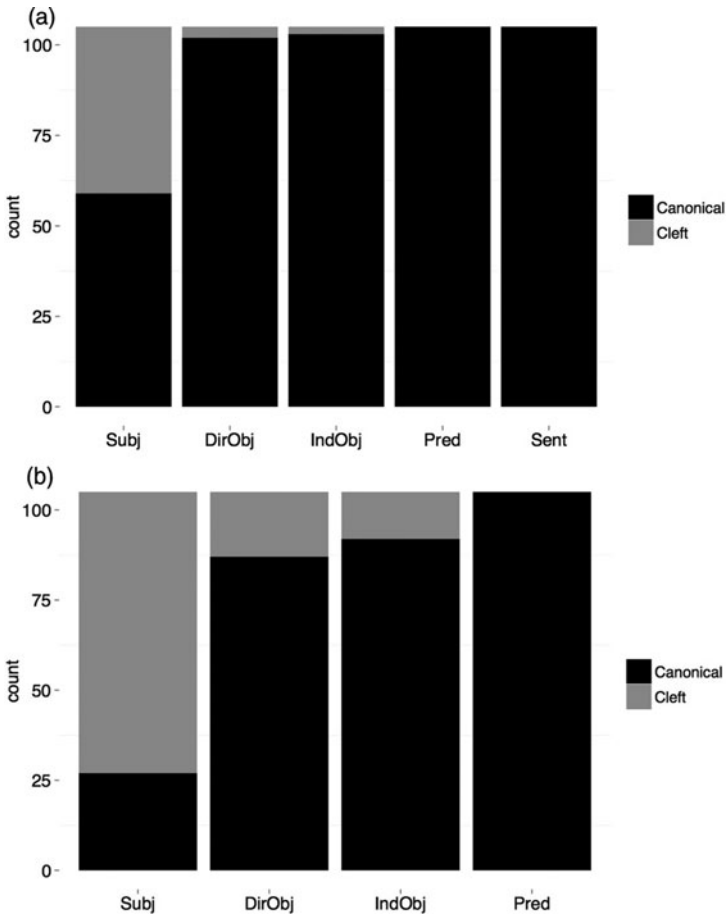

Figure 4. *(a) Count per grammatical function in informational context in StF. (b) Count per grammatical function in corrective context in StF*

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

sentences are more natural in the standard variety of the language is confirmed: grammatical subjects are focused in preverbal position significantly more in StF than in CoF ($z = 5.7$, $p < 0.01$). Finally, StF also displays a marked asymmetry between arguments. Although less distinct than in CoF, the argument asymmetry is also present in StF: the cleft is used significantly more to signal subject focus than non-subject focus. In fact, the rate of clefting with non-subjects is trivial (2.45%).

As opposed to CoF, focus-type is found to have an effect on the form produced in StF: the difference between the sentence form used in a corrective context vs. in a informational one is significant ($\beta = -0.16$, SE $= 0.07$, $t = -2.26$, $p < 0.05$) and a model that included the predictor focus-type gave a slightly better fit to the data than one without it ($\chi^2(1) = 5.04$, $p < 0.05$). But this is not the whole story. In fact, the effect observed is mainly due to the interaction of grammatical function with focus-type: most of the clefts produced in a corrective context in StF are produced with grammatical subject ($\beta = -0.16$, SE $= 0.055$, $t = -2.92$, $p < 0.01$). Corrective non-subjects, on the other hand, although clefted significantly more often than in an informational context, are nonetheless mostly realized *in situ*. A model with the interaction of the two predictors, grammatical function and focus-type, performs better than one without the interaction ($\chi^2(7) = 137.58$, $p < 0.001$).

5 SUMMARY AND DISCUSSION

5.1 *Discussion*

Based on the claims found in the previous literature and the well-documented differences between the colloquial and the standard variety of the language, I predicted that the grammatical function of the focus element, the focus-type encoded and the language variety (or standardity) were factors that would affect the sentence structure produced in the marking of focus. These predictions were mostly supported by the results, but two new findings appear: (i) a stronger focus interpretation did not systematically induce the use of a cleft (i.e., I found no strict one-to-one relationship between focus-type and its realization), and (ii) variation does exist, both within and across language varieties. Given the limitations considered for the methodology presented, it will be useful to keep the results in perspective, for they could be artifacts of the experimental design and may not generalize to the language overall.

Let us first discuss variation and what implications it has for past categorical account on focus marking. Within CoF, it was found that although small in proportion, instances of informational non-subject focus *can* occur in a cleft (12.9% for both direct and indirect objects). This result challenges the claim made in Belletti (2012), but also, as noted in Skopeteas and Fanselow (2010: 171), the fact that a strict implementation of the focus-type asymmetry excludes a grammar in which 'non-canonical structures occur with non-identificational instances of focus while identificational instances of focus are expressed through canonical structures'.

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

Whether these results can be taken to represent more than an artifact of a particular experimental design, that is, represent the reality of the speakers' language use, can be debated, yet the variation must be explained. Indeed, at first sight, a cleft construction seems like a costly strategy to focus non-subjects in informational contexts: It is more complex than a canonical sentence both syntactically (it adds a layer of syntactic material to the output) and phonologically (it creates an extra ι-phrase boundary that is not required to fulfill the phonological alignment constraint). From a pragmatic perspective, the use of a more complex structure when a simpler one is available can be seen as flouting Grice's (1975) maxim of Manner: 'be brief, avoid unnecessary wordiness'. Interestingly, a closer look at the data reveals that these results are distributed across just six participants (out of 21). This observation leads me to propose that the variation arises from properties of these speakers' grammars; certain speakers have a bias towards marking the category *focus* in a non-canonical sentence, overgeneralizing the use of these structures to the both subject and non-subject focus. Indeed, these speakers are also the ones who produce most of the clefts found in the corrective focus condition. I formalize this hypothesis via the constraint in (17) (borrowed from Zerbian, 2007), which maps the information-structural category focus to a specific syntactic position. Its interaction with other constraints is discussed in Section 6.

(17) Foc-Spec: Focused constituents must be in Specifier position.

Secondly, within the StF condition, optionality is observed in the marking of grammatical subjects. This is due to the fact that the grammar of StF and CoF differ in what Vallduvi (1992) calls 'plasticity' – the ability to mold prominence according to information status – and therefore provide speakers with different structural options to realize focus. While the default position for main stress in both language varieties is rightward, CoF strongly prefers to keep main stress in the default position, consequently altering the syntax when a mismatch occurs between the position of main stress and the focus element. The grammar of StF is similar to that of English: it allows for main stress to be moved to the position of the focused element (leftward in the case of subject focus) and is therefore stricter with regards to keeping the canonical word order intact.

Finally, previous accounts have derived clefting as the optimal strategy in corrective contexts, specifically for non-subjects (Hamlaoui, 2009), and cartographic accounts and purely semantic accounts of focus have claimed that the interpretational properties of the cleft are directly derived from its syntactic configuration (Kiss, 1998; Clech-Darbon et al., 1999; Belletti, 2009, 2012; Drubig, 2003). However, these accounts are questioned by the results presented here. Indeed, I have established that there is no strict relationship between focus-type and focus realization, showing that a corrective focus – although a potential condition – is not a sufficient one to trigger clefting. Although clefts may constitute a more transparent way to encode a stronger interpretation of focus, they also constitute a more complex way to do so. And if speakers have access to both *in situ* and *ex situ* strategies to encode the same focus-type, their choice should be mainly guided

by economy of form (i.e., the minimality condition formulated in Skopeteas and Fanselow, 2010): the structure with the lowest degree of structural complexity is expected to come out as the surface form. But variability is observed in this context too: for the vast majority, non-subjects are realized *in situ*, but clefting is also allowed, especially in CoF. Here again, I explain this variability by positing that the grammar of these speakers include the constraint in (17).

But another question remains unresolved: are focus constituents realized *in situ* in corrective contexts fundamentally different than focus constituents realized *in situ* in informational contexts? Although a detailed prosodic analysis is beyond the scope of this article and more research is needed on the phonetic correlates of different focus types in French, the vast literature on the prosodic correlates of contrastive/corrective focus suggests that there is indeed a difference (see Section 2.2). Thus, in this article, I hypothesize that despite the lack of evidence for a systematic difference in the syntax, the two types of focus should involve distinct prosodic properties, and therefore be distinguished in the grammar. In other words, instances of corrective focus that are expressed *in situ* should systematically differ from *in situ* informational focus in their prosody (observable on the duration, pitch and intensity on contrastive foci, see Katz and Selkirk (2011) for a study on English).

I choose to formalize this hypothesis by having the constraint on prominence (introduced in (9)) be factorized for focus-types (following Féry and Samek-Lodovici, 2006), as shown in (18). Crucially, with $\text{SF}_{\text{identificational}}$ ranked above $\text{SF}_{\text{informational}}$, a stronger interpretation of focus will receive higher prominence than an informational focus.

(18) STRESS-FOCUS$_X$: An x-focus phrase has the highest prosodic prominence in its focus domain.

6 A STOCHASTIC-OT ANALYSIS TO FOCUS MARKING IN COF AND STF

As mentioned in the introduction, most of the prior studies on focus marking in French offer an account of the canonical-cleft alternation that derives a single best output form in CoF. The theoretical contribution of this article lies in that I develop a stochastic-OT analysis to formalize the variation observed especially the fact that, in contrastive contexts, the cleft is not necessarily produced.

6.1 Theoretical framework

The stochastic model of Optimality-theory developed by Boersma and Hayes (2001) differs from the classic version (Prince and Smolensky, 1993) in two major ways. First, instead of adopting an ordinal ranking, it adopts a ranking along a continuous scale. Consequently, the distance between the constraints is meaningful and can vary.

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

Figure 5. *(Colour online) Categorical constraint ranking (no overlap) with ranges of variation*

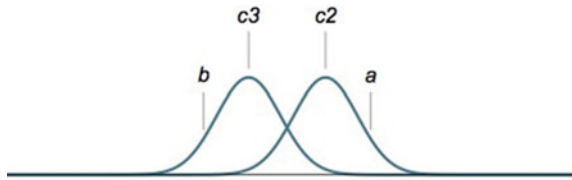

Figure 6. *(Colour online) Free constraint ranking (overlap) with ranges of variation*

Second, every time a candidate is evaluated, the position of each constraint is perturbed by some random noise, which is defined to have a Gaussian distribution with a standard deviation of 2. 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, which are interpreted as normal probability distributions. In other words, the constraints are associated with a range of values rather than a single point. If the ranges of for example C1 and C2 do not overlap, the more highly ranked constraint will strictly dominate the lower ranked constraint (Figure 5).

If the constraints have overlapping selection points, such as C2 and C3 for example, they can be ranked freely. At the time of evaluation, the selection point can be selected within the ranges of both constraints. In that case, the two constraints are not too distant and variable outputs can be generated. In a situation where the candidates evaluated are in the upper part of C2 and lower part of C3, the resulting ranking is C2 > C3: candidate *a* is favoured over candidate *b* (see Figure 6).

As the distance between two crucially ranked constraints increases, their distributions overlap less and less, and a categorical output emerges. Second, in cases where the candidates evaluated are in the upper part of C3 and the lower part of C2, the opposite ranking holds: C3 > C2. Here, candidate *b* comes out as the preferred output (see Figure 7).

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

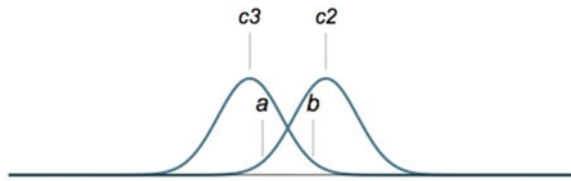

Figure 7. *(Colour online) Reversal of constraint dominance*

6.2 Input, candidates and constraints

While there is little debate on how the input should be structured in OT phonology – it consists of the underlying phonological representation of a word – defining the input in OT syntax proves more complicated because, according to Heck et al. (2002), syntax is information preserving while phonology is not. Scholars have made different proposals; whether the input is not, partially (Grimshaw, 1997) or highly structured (Legendre et al., 1998). Here, I follow Grimshaw (1997) in assuming the input is syntactically structured; it consists of a lexical head and its arguments, (b) an assignment of lexical heads to its arguments, (c) information on tense, and (d) semantically meaningful auxiliaries. I further follow Legendre et al. (1998) in 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. These assumptions have direct repercussions for the candidates: the set generated includes candidates that are semantically equivalent (Grimshaw, 1997; Kager, 1999; Feldhausen and Vanrell, 2014) and have the same focus domain (narrow vs. broad) as well as the same focus-type (informational vs. corrective). Thus, candidates only diverge from the input by their syntactic structure and their information-structural status.

The analysis I put forward relies on the constraints in (19a)-(19f):

(19) a. STRESS-FOCUS$_X$ (SF$_X$): An $_X$ focus phrase has the highest prosodic prominence in its focus domain.
b. SUBJECT: Highest A-specifier must be filled.
c. *STRUCTURE (*STR): Do not add syntactic material not present in the input.
d. HEAD-I-R (HIR): Align the right boundary of every intonation phrase with its head.
e. *)ι: Do not insert ι-boundaries
f. FOC-SPEC: Focused constituents must be in Specifier position.

6.3 Accounting for the Argument asymmetry

6.3.1 Subject focus

Concentrating on informational contexts, the ranking in (20) accounts for the results from experiment 1 (for the most part in CoF, subjects are clefted and non-subjects remain *in situ*). I argue that the variation observed within CoF and in StF

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

Tableau 1. *Realization of informational subject focus in CoF*

/Marie, acheter, un journal/ foc = subject type = informational		SF_{info}	SUBJ	___*overlap*___ HIR	*STR	*)ι
a. 82%	x)ι x)ϕ x)ϕ C'est [Marie]$_F$ qui a acheté un journal				*	**
b. 18%	x)ι x)ϕ x)ϕ [Marie]$_F$ a acheté un journal			**		*
c.	x)ι x)ϕ x) ϕ a acheté un journal [Marie]$_F$		*!			*
d.	x)ι x)ϕ x)ϕ Marie a acheté [un journal]$_F$	*!				*
e.	x)ι)ι x)ϕ x)ϕ C'est [un journal]$_F$ que Marie a acheté	*!				**

can be fully explained by positing that the distance between the two constraints HIR and *STRUCTURE is small, and that these constraints overlap and can be reversed, as illustrated in ranking (21). Tableau 1 illustrates the realization of subject focus in Colloquial French.

(20) sf_X > subject > HIR > *structure > *)ι
(21) sf_X > subject > * structure > HIR > *)ι

The two candidates *d* and *e* differ from the input in their logical form. They never emerge as output because they violate the constraint $SF_{informational}$ that penalizes candidates that fail to mark the appropriate focus element more prominently than a non-focus. This constraint dominates all others since non-identity between the focus in the input and the focus in the output and inconsistencies in the discourse are strictly prohibited. The violation imputed by candidate *c* is fatal: CoF rules out sentences without an overt preverbal subject, therefore attributing great importance to the constraint SUBJECT. The emergence of the cleft candidate *a* over the canonical candidate *b* arises from the ranking HIR > *STRUCTURE, with the candidates being evaluated in the upper part of HIR and the lower part of *STRUCTURE (as illustrated in Figure 8).

In 82% of cases, the constraint HIR is stricter than *STRUCTURE, requiring the focus element to be aligned rightward of a ι-phrase. In that case, the clefted candidate *a* emerges since its bi-clausal structure creates an extra ι-boundary. In 18% of cases, candidate *b,* which is more faithful to the input and therefore conforms to *STRUCTURE, occurs as the surface form. This is due to the overlap of the two constraints HIR and *STRUCTURE and the evaluation of candidates being done in the upper part of the latter constraint (see Figure 9). This overlap is explained by

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

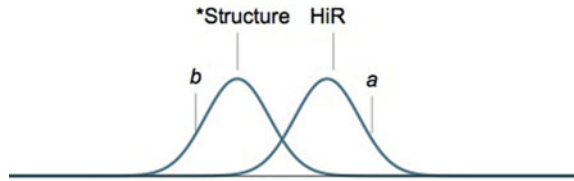

Figure 8. *(Colour online) Free variation in CoF*

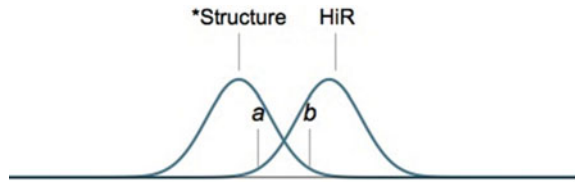

Figure 9. *(Colour online) Reversal of constraint dominance in CoF*

Tableau 2. *Realization of informational subject focus in StF*

/Marie, acheter, un journal/ foc = subject type = informational		SF_{info}	SUBJ	___overlap___ HIR	*STR	*)ι
a. 56.2%	x)ι x)ϕ x)ϕ [Marie]$_F$ a acheté un journal				*	**
b. 43.8%	x)ι)ι x)ϕ x)ϕ C'est [Marie]$_F$ qui a acheté un journal			**		*

the strictness of the constraint HIR, which is a function of the distance between this constraint and the conflicting constraint *STRUCTURE: Speakers with a more laxed version of HIR in their grammar, i.e. where the distance between HIR and *STRUCTURE is smaller, will produce more canonical sentences than speakers who have a more stringent version of HIR.

Finally, it shall be noted that, although native speakers possess other non-canonical structures to signal information-structure, and in particular the presentational *avoir*-cleft (Il y a Marie qui a acheté un journal au kiosque), these structures do not enter the competition because they differ from the input in their focus domain. The *avoir*-cleft for example is used in broad-focus contexts where both the subject and the predicate are in focus (Lambrecht, 1994, 2001).

Experiment 2 established that StF offers speakers the possibility to express narrow subject focus *in situ*, which I argued comes from the possibility for StF to shift main stress to the position of focus, disfavouring more complex structures that diverge from the input. So, in this language variety, the two overlapping constraints *STRUCTURE and HIR are ranked default *STRUCTURE > HIR (see Tableau 2). This leads

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

Tableau 3. *Realization of informational non-subject focus in CoF*

	/Marie, acheter, un journal / foc = dir.obj type = informational	SF_{info}	SUBJ	*STR	HIR	F–SP	*)ι
	x)ι					*	*
a.	x)ϕ x)ϕ						
83.8%	Marie a acheté [un journal]$_F$						
	x)ι)ι			*			**
b.	x)ϕ x)ϕ						
16.2%	C'est [un journal]$_F$ que Marie a acheté						

to the emergence of the canonical form in 56.2% of cases (candidate *a*). For the clefted candidate *b* to surface as the output (43.8%), the ranking HIR > *STRUCTURE must hold, with the evaluation of the candidates taking place in the higher part of HIR.

6.3.2 Non-subject focus

The ranking in (20) also accounts for the realization of non-subject focus *in situ*, the most frequent marking strategy observed in CoF and StF. Indeed, under the prosodic account proposed here, *in situ* realization is expected since the canonical position of non-subjects is by default aligned with the position of main stress (rightward of an intonational phrase). Thus, the prosodic constraint HIR is never violated. The crucial constraint that determines the output is the economy constraint *STRUCTURE: the canonical candidate *a* is favoured because it is the least structurally complex. But, as discussed in Section 4.4, CoF displays a small amount of variation with certain speakers resorting to clefting. This variation, because specific to a certain subset of speakers, is explained by positing that the constraint FOC–SPEC (F–SP) is present in the grammar of certain speakers and enters the competition. It will be ranked higher than *STRUCTURE in order for the cleft to emerge as the output (see Tableau 3). Since clefting is so rare in StF (2.8%), I argue that this constraint is not active.

6.4 Accounting for the Focus-type asymmetry

This last section aims to formalize the realization of focus in contexts where it is associated with a stronger interpretation, specifically here *correction*.

While previous OT accounts on French (Hamlaoui, 2009) derive clefting as the best strategy to signal corrective focus, the results from experiments 1 and 2 show that the two are not strictly correlated. In fact, while instances of corrective subjects are clefted on independent grounds, instances of non-subjects corrective focus are mostly realized in their canonical position, both in CoF and StF. However, in line with studies on the prosodic correlates of corrective focus (Di Cristo, 1998; Jun and Fougeron, 2000; Dohen et al., 2006), I hypothesize that *in situ* instances of corrective focus should systematically differ from their informational counterparts

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

Tableau 4. *Realization of corrective non-subject focus in CoF*

/Marie, acheter, un journal / foc = dir.obj type = identificational (corrective)		SF_{ident}	SF_{info}	SUBJ	*STR	HIR	F–SP	*)ι
a. 66.7%	X)ι x)ϕ x)ϕ Marie a acheté [un JOUrnal]$_F$						*	*
b. 33.3%	X)ι)ι x)ϕ x)ϕ C'est [un JOUrnal]$_F$ que Marie a acheté				*			**
c. Marie a acheté [un journal]$_F$	x)ι x)ϕ x)ϕ	*!						*
d.	x)ι)ι x)ϕ x)ϕ C'est [un journal]$_F$ que Marie a acheté	*!			*			**

in their prosodic realization. Thus, a crucial part of my proposal is that corrective focus is unambiguously marked, but instead of using syntactic means, CoF and StF allow this marking to mostly be done via prosody (for example via an *accent d'insistance*). This claim is implemented in optimality terms by having the constraint STRESS-FOCUS factorized for focus-type, and more specifically having $SF_{identificational}$ ranked above $SF_{informational}$. This ranking encodes that a stronger focus type must be more prominent than a neutral (informational) focus. Tableau 4 illustrates the competition of constraints involved in corrective contexts:

Candidate *c* and *d* show that, because they have the same prosodic structure as instances of informational focus, violate $SF_{identificational}$ and are ruled out. Candidate *a* is the most common output because it only violates the lower ranked constraint foc-spec, but fulfills $SF_{identificational}$ by having a special prosodic marking on the focused element. Candidate *b* emerges as the output in 33.3% of cases in CoF in the grammar of speakers that includes the constraint FOC-SPEC, which overlaps with *STRUCTURE and favours focus elements in a higher syntactic position.

7 CONCLUSION

This article argued that cartographic and categorical OT accounts of focus realization do not account for the full range of data found within and across language varieties in French. Based on the results from two controlled tasks, I showed that there exists an importance difference between CoF and StF: the way in which subject focus is realized. If the results can be taken to represent more than an artifact of a particular experimental methodology, they support the claim that

https://doi.org/10.1017/S0959269515000265 Published online by Cambridge University Press

speakers have access to two grammars (Zribi-Hertz, 2013). I argued in favour of a unified Stochastic OT analysis arguing that these two grammars are governed by the same constraints, which can be reversed to illustrate the difference on strictness towards the position of main stress: CoF demands that the rightward position of main stress and focus be aligned, leading to the use of a cleft structure which provides two intonational phrases, whereas StF is more lenient and allows main stress to shift, thus allowing canonical sentences to emerge. As for the variation found with corrective focus, I have proposed to account for it by positing a factorized constraint with respect to focus type.

Address for correspondence:
Department of French and Italian
University of Iowa
111 Phillips Hall
Iowa City
Iowa 52245
United States
e-mail: e-destruel-johnson@uiowa.edu

REFERENCES

Aissen, J. (1999). Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory*, 17: 673–711.

Bates, D., Maechler, M. and Bolker, B. (2012). *lme4*: Linear mixed-effects models using S4 classes. Retrieved from http://CRAN.R-project.org/package=lme4 (R package version 0.999999-0)

Belletti, A. (2005). Answering with a cleft. The role of the null subject parameter andthe vp periphery. In: L. Brugè, G. Giusti, N. Munaro, W. Schweikert and G. Turano (eds.), *Proceedings of the Thirtieth 'Incontro di Grammatica Generativa'*. Cafoscarina, pp. 63–82.

Belletti, A. (2009). *Structures and Strategies*. New York: Routledge.

Belletti, A. (2012). Revisiting the CP of Clefts. In: E. Zimmermann and G. Grewendorf (eds.), *Discourse and Grammar. From Sentence Types to Lexical Categories* (*Studies in Generative Grammar 112*). Berlin: De Gruyter. pp. 91–114.

Blanche-Benveniste, C. (1997). *Approches de la langue parlée en français: L'essentiel français*. Paris: Ophrys.

Blyth, C.S. (1999). Toward a pedagogical discourse grammar: Techniques for teaching word order constructions. In: J. Lee and A. Valdman (eds.), *Form and Meaning: Multiple Perspectives. AAUSC Issues in Language Program Direction*. Boston: Heinle, pp. 183–229.

Boersma, P. and Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, 32: 45–86.

Büring, D. (2009). Towards a typology of focus realization. In: M. Zimmermann and C. Féry (eds.), *Information Structure: Theoretical, Typological and Experimental perspectives*. Oxford: Oxford University Press, pp. 177–205.

Büring, D. and Gutiérrez-Bravo, R. (2001). Focus-related word order variation without the NSR: A prosody based cross-linguistic analysis. In: J. McCloskey (ed.), *SASC*

3: Syntax and Semantics at Santa Cruz. Linguistics Research Center, University of California, Santa Cruz.
Calude, A. (2009). *Cleft Constructions in Spoken English*. VDM Verlag.
Chafe, W.L. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In: C. Li (ed.), *Subject and Topic*. New York: Academic Press, pp. 25–55.
Chomsky, N. (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.
Clech-Darbon, A., Rebuschi, G. and Rialland, A. (1999). Are there cleft sentences in French? In: G. Rebuschi and L. Tuller (eds.), *The Grammar of Focus*. Amsterdam: John Benjamins, pp. 83–118.
Coveney, A. (2002). *Variability in Spoken French: A Sociolinguistic Study of Interrogation and Negation*. Bristol: Elm Bank.
Coveney, A. (2011). A language divided against itself? Diglossia, code-switching and variation in French. In: F. Martineau et T. Nadasdi (eds), *Le français en contact*. Québec: Presses de l'Université Laval, pp. 51–85.
de Cat, C. (2007). *French Dislocation: Interpretation, Syntax and Acquisition*. Oxford: Oxford University Press.
Delais-Roussarie, E. (2005). *Phonologie et Grammaire: étude et modélisation des interfaces prosodiques*. Université de Toulouse 2. Habilitation thesis.
Di Cristo, A. (1998). Intonation in French. In: Hirst, D.J. and Di Cristo, A. (eds), *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press, pp. 195–218.
Dohen, M., Loevenbruck, H. and Hill, H. (2006). Visual correlates of prosodic contrastive focus in French. In: *Proceedings of Speech Prosody*. pp. 221–4.
Drubig, H.B. 2003. Toward a typology of focus and focus constructions. *Linguistics*, 44: 1–50.
Etienne, C. and Sax, K. (2009). Stylistic variation in French: Bridging the gap between research and textbooks. *Modern Language Journal*, 93: 584–606.
Feldhausen, I. and del Mar Vanrell, M. (2014). Prosody, focus and word order in Catalan and Spanish: An Optimality Theoretic approach. In: *Proceedings of the 10th International Seminar on Speech Production* (ISSP), 5–8 May 2014, Cologne.
Féry, C. (2001). Focus and phrasing in French. In: C. Féry and W. Sternefeld (eds), *Audiatur Vox Sapientiae. A Festschrift for Arnim von Stechow*. Berlin: Akademie Verlag, pp. 153–181.
Féry, C. (2013). Focus as alignment. *Natural Language and Linguistic Theory*, 31: 683–734.
Féry, C. and Samek-Lodovici, V. (2006). Focus projection and prosodic prominence in nested foci. *Language*, 8: 131–50.
Gabriel, C. (2010). On focus, prosody and word order in Argentinian Spanish: A minimalist OT account. *ReVEL*, 4: 183–222.
Grice, H.P. (1975). Logic and conversation. In: P. Cole and J. Morgan (eds.), *Syntax and Semantics, Vol. 3: Speech Acts*. New York: Academic Press.
Grimshaw, J. (1997). Projection, heads, and optimality. *Linguistic Inquiry*, 28: 373–422.
Hamlaoui, F. (2007). French cleft sentences and the syntax-phonology interface. In: M. Radisic (ed.), *Actes du Congrès Annuel de l'Association Canadienne de Linguistique 2007*. Canadian Linguistic Association.
Hamlaoui, F. (2009). *La focalisation à l'interface de la syntaxe et de la phonologie: le cas du français dans une perspective typologique*. PhD. thesis, University Paris 3 Sorbonne Nouvelle.

Hamlaoui, F. (2010). La structure de l'information et la prosodie dans les questions partielles : une analyse contrastive du français démotique et du français classique tardif. *Manuscript*.

Heck, F., Müller, G., Fischer, S., Vikner, S., Schmid, T. and Vogel, R. (2002). On the nature of the input in optimality theory. *The Linguistic Review*, 19: 345–376.

Hornsby, D. (1998). The dynamic model and inherent variability: the case of northern France. *Journal of Applied Linguistics*, 6: 19–36.

Jun, S.-A. and Fougeron, C. (2000). A phonological model of French intonation. In: A. Botinis (ed.), *Intonation: Analysis, Modeling and Technology*. Dordrecht: Kluwer Academic Publishers, pp. 209–42.

Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.

Katz, S. (1997). *The syntactic and pragmatic properties of the* c'est-*cleft construction*. Ph.D. thesis, University of Texas at Austin.

Katz, S. and Blyth, C. (2007). *Teaching French Grammar in Context*: *Theory and Practice*. New Haven: Yale University Press.

Katz, J. and Selkirk, E. (2011). Contrastive focus vs. discourse-new: evidence from prosodic prominence in English. *Language*, 87: 771–816.

Keenan, E. (1976). Toward a universal definition of subject. In: C. Li (ed.), *Subject and Topic*. New York: Academic Press, pp. 303–33.

Kiss, K.E. (1998). Identificational focus versus information focus. *Language*, 74: 245–73.

Klein, W. (2012). The information structure of French. In: M. Krifka and R. Musan (eds.), *The Expression of Information Structure*. Berlin: De Gruyter, pp. 95–126.

Koehn, P. (2005). *Europarl*: A Parallel Corpus for Statistical Machine Translation. *Conference Proceedings: the Tenth Machine Translation Summit*: 79–86. AAMT.

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55: 243–76.

Labov, W. (1969). Contraction, deletion and inherent variability of the English copula. *Language*, 45: 715–62.

Labov, W. (1972). *Sociolinguistic Patterns*. Trad. fr. *Sociolinguistique*. Paris: Minuit.

Lambrecht, K. (1987). On the status of SVO sentences in French discourse. In: R. Tomlin (ed.), *Coherence and Grounding in Discourse*. Amsterdam: John Benjamins, pp. 217–62.

Lambrecht, K. (1994). *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.

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

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 (eds.), *Is the Best Good Enough? Optimality and Competition in Syntax*. Cambridge, MA: MIT Press, pp. 249–89.

Marandin, J.-M. (2004). Pour une approche dialogique du contexte et de la structure informationnelle. On-line access: http://llf.linguist.jussieu.fr/llf/Gens/Marandin/Focus-HermesJuin04.pdf

Massot, B. (2008). *Français et diglossie. Décrire la situation linguistique française contemporaine comme une diglossie: arguments morphosyntaxiques.* PhD thesis, University Paris-8. http://tel.archives-ouvertes.fr/tel-00726999.

Myers, L.L. (2007). *Wh-interrogatives in spoken French: A corpus-based analysis of their form and function*. PhD dissertation, University of Texas at Austin.

Prince, A. and Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers Center for Cognitive Science Technical Report.
Quillard, V. (2001). La diversité des formes interrogatives: Comment l'interpréter? *Langage et Société*, 95: 57–72.
Reichle, R.V. (2014). Cleft type and focus structure processing in French. *Language, Cognition and Neuroscience*, 29: 107–24.
Roland, D., Dick, F. and Elman, J.L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, 57: 348–379.
Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1: 75–116.
Rossi, M. (1999). *L'intonation, le système du francais: description et modélisation*. Paris: Ophrys.
Rowlett, P. (2013). Do French speakers really have two grammars? *Journal of French Language Studies*, 23: 37–57.
Samek-Lodovici, V. (2005). Prosody-Syntax Interaction in the expression of focus. *Natural Language and Linguistic Theory*, 23: 687–755.
Selkirk, E. (1984). *Phonology and Syntax*. Cambridge, MA: MIT Press.
Skopeteas, S. and Fanselow, G. (2010). Focus types and argument asymmetries: A cross-linguistic study in language production. In: C. Breul (ed.), *Contrastive Information Structure*. Amsterdam and Philadelphia: John Benjamins, pp. 169–98.
Szendrői, K. (2003). A stress-based approach to the syntax of Hungarian focus. *The Linguistic Review*, 20: 37–78.
Truckenbrodt, H. (1995). *Phonological Phrases: their relation to syntax, focus and prominence*. PhD dissertation, MIT.
Vallduví, E. (1992). *The Informational Component*. New York: Garland.
Vander Klok, J., Wagner, M. and Goad, H. (manuscript). *The Prosodic Marking of Focus in English, Québec French, and European French*. Mc Gill University.
Vion, M. and Colas, A. (1995). Contrastive marking in French dialogue: Why and how? *Journal of Psycholinguistic Research*, 24: 313–31.
Zerbian, S. (2007). Subject/object-asymmetry in Northern Sotho. In: P. Schwabe and S. Winkler (eds.), *Information Structure and the Architecture of Grammar: A Typological Perspective*. Amsterdam: John Benjamins, pp. 323–45.
Zimmermann, M. (2006). Focus in Western Chadic: A unified OT-account. In: C. Davis et al. (eds.), *Proceedings of NELS 36*. Amherst: University of Massachusetts Press.
Zimmermann, M. (2011). The grammatical expression of focus in West Chadic: Variation and uniformity in and across languages. *Linguistics*, 49: 1163–213.
Zribi-Hertz, A. (2011). Pour un modèle diglossique de description du français: quelques implications théoriques, didactiques et méthodologiques. *Journal of French Language Studies*, 21: 1–26.
Zribi-Hertz, A. (2013). De la notion de grammaire standard dans une optique diglossique du français. *Journal of French Language Studies*, 23: 59–85.