

Prominence in French dual focus

Emilie Destruel (Iowa) and Caroline Féry (Frankfurt), October 2019

Abstract

This paper investigates how French signals prominence in prosody in the post-verbal domain of sentences with two objects or two adjuncts that vary in information status and prosodic length. The information status of particular interest here is dual focus, defined as the presence of two foci in a mono-clausal sentence, but other information states are investigated as well. The controlled production experiment we report on allows for a detailed examination of prosodic prominence. High boundary tones at the end of non-final prosodic phrases are pervasive, as has been documented in many studies before the present one. An important but less documented result is the variation in different prosodic curs, in particular in the number and position of high tones, as well as the particular scaling relationship between them, providing a powerful tool for the expression of (dual) focus. We also report on a perception experiment with our data, showing that overall, French speakers have difficulties to recognize the information structural context in which a sentence was produced, especially when compared to German. However, French listeners recognize dual focus better than other information states. Overall, this article provides elements of answers as to why French prosody is so difficult to pin down, and why contradictory results and analyses have been proposed for this language.

Keywords: French, prominence, dual focus, objects, adjuncts

1. Introduction

This article investigates phonetic and phonological prominence in French, with a special emphasis on mono-clausal sentences containing a so-called ‘dual focus’, i.e. sentences that answer interrogatives containing two ‘wh’-phrases. We understand phonetic prominence as a F0 raise, and phonological prominence as a high tone, at a local point in a sentence that signals prosodic highlighting of a word or a phrase (phrasal tone, boundary tone). Despite the growing body of experimental work on the interaction between prosody and information structure in French, the realization of dual focus is rarely examined. Yet, this strikes us as an important issue given that mainstream prosodic theories typically disallow

two main prosodic heads within one prosodic domain (Gussenhoven 2004; Selkirk, 1995; Truckenbrodt, 1995, as well as Kabagema-Bilan, López-Jiménez & Truckenbrodt, 2011 and Wang & Féry 2017 for an explicit formulation of the conflict). So, the issue comes down to understanding how phonetics and phonology cope with two foci within a single sentence, especially compared to single focus or to wide focus. The present study should also be understood as an extension of our previous work on postverbal given constituents, see Destruel & Féry 2019, where we showed that objects and adjuncts behave differently in this position: objects are phrased together with the preceding verb, but adjuncts are phrased independently, a difference in behaviour that is reflected in the prosodic correlates.

Given this backdrop, the specific goals of our study on French are, (i) to examine how prominence is realized in post-verbal sequences with two objects versus two adjuncts, (ii) to provide empirical data on dual focus via a newly conducted production experiment, (iii) to test whether French listeners are able to identify the context in which a sentence was produced and (iv) to compare the results with other types of focus (broad, initial and final focus), as well as what is known about dual focus realization in other languages, notably English, German and Mandarin.

The remainder of the introduction provides some background information relevant to our empirical study: focus and givenness, dual focus and prior research on French prosody. Section 2 presents the study, detailing our research questions and hypotheses, methods, and reporting on the results. Section 3 discusses individual variation in our data. Section 4 presents and reports on the perception study designed to test whether listeners could identify the intended focus structure of our target sentences. Section 5 relates the results to the literature on information-structure and prosody in French, as well as cross-linguistic results in the prior research on dual focus. Section 6 concludes.

1.1. Focus and givenness

In this paper, we follow Krifka (2008), Rooth (1992) and Schwarzschild (1999), whereby the notion of focus is taken to be a universal information-structural category that evokes a set of (explicit) alternatives that the speaker takes to be salient in the context. Focus is most commonly diagnosed in question-answer pairs, where it corresponds to the answering element associated with the ‘wh’-word in the congruent question. Thus, in (1), ‘Mary’ is

taken to be the focus of the sentence. In Germanic languages, prosodic prominence is realized on a lexically stressed syllable.

(1) Q: Who baked brownies?

A: [MARY]_F baked brownies.

The notion of givenness is defined as “already mentioned in the context”, since this is the only kind of givenness occurring in the production experiment reported below.

In German and English, the given elements display reduction of pitch range in the pre-focal region of the sentence, *deaccenting* or *compression* in the post-focal region of the sentence, see Ladd (1980, 2008) for English as well as Féry & Kügler (2008) and Kügler & Féry (2017) for German among others. In French, compression is also used for expressing givenness, but it may be less systematic than in the Germanic languages, and also confined to entire prosodic phrases, see next section.

1.2. Background on French prosody

Lexical stress does not exist in French. Thus, a core question concerns where and how prominence is exactly realized—if at all.¹ Prior work on French has shown that syntax-based strategies are frequent for expressing focus and givenness, especially those involving phrasing, with clefts being the device used by default, especially when focus falls on the grammatical subject (Clech-Darbon, Rebuschi & Rialland 1999; Destruel, 2013; Hamlaoui, 2009; Lambrecht, 1994). Nevertheless, scholars note that information structure does influence prosody. Post-focal compression is optionally found in corrective contexts (Welby 2006), but is rarely found in information focus (Vander Klock, Goad & Wagner 2018).

The phonological analysis of the prosodic pattern of French is subject to different interpretations. Most studies on prosody assume that an important prosodic constituent of French is what we call a ‘prosodic phrase’ (henceforth, Φ -phrases), or ‘phonological

¹ The question whether a H tone is prominent is also sometimes discussed for other languages, as in two other articles of this special issue, namely on Medumba (Franich, this volume) and on Taiwanese (Ou & Guo, this volume).

phrase' (Post, 2000).² As far as phonetic correlates of prominence are concerned, all researchers find a rising tonal excursion in the non-final position of a sentence (Hirst and Di Cristo 1996, Rossi 1980). Opinions differ, however, as to how to analyze the final rise; certain scholars argue that it should be analyzed as a pitch accent (Astésano, Bard & Turk, 2007; Beyssade, Marandin & Rialland, 2003; Delais-Roussarie, 1995; Delais-Roussarie et al., 2002; Portes, Beyssade, Michelas, Marandin & Champagne-Lavau (2014), while others argue that it is a demarcative tone, thus a boundary tone (Fónagy, 1979; Vaissière, 1980), or a phrasal tone (Féry, 2014) or both a pitch accent and a boundary tone (Jun & Fougeron, 2000).

Moreover, several authors also assume an optional phrase-initial prominence in the prosodic phrase, that is also realized with a pitch excursion (see D'Imperio & Michelas, 2010; Dohen & Loevenbruck, 2004; Jun and Fougeron, 2000; Post, 2000 among others). The exact location of this initial high tone is variable: it can be initial, or on the second or third syllable (see Welby, 2006). It also varies along different dimensions: it can be rhythmical or marking information structure (see Di Cristo 1998; Padeloup 1990 and Rossi 1980). A number of authors assume that it delimits the beginning of the focused constituent (see German and D'Imperio 2016).

Finally, our recent research has supplemented prior findings by investigating the phenomenon of F0 'compression' (i.e. a reduction of pitch range) in post-focal sequences. Destruel & Féry (2015, 2019) showed that while compression does occur in French, it mostly affects entire Φ -phrases: When given, only post-verbal adjuncts were phrased separately from the verb, thus being optionally realized with a compressed register, objects were not or insignificantly so. Taken together, these findings suggest that, although French resorts to prominence in information-structural contexts, it does so quite differently from Germanic languages, see O'Brien (2019) and Fanselow (2016) for an overview on the correlates of information structure in Germanic languages. Vander Klok et al. (2018) proposed that the difference between English and French realization of givenness lies in the semantics of focus. They propose that the reason for the difference in prosodic

² A prosodic constituent of this size has been given a variety of names in the literature, for instance 'rhythmic group' by Beyssade et al. (2003), and 'Accentual Phrase' by Jun & Fougeron (2000).

compression of a given constituent is to be found in the semantic and/or pragmatic content of the focus operator. French can only focus entire clauses, whereas English can focus any constituents. Post-focal constituents are not compressed because they are obligatorily part of the larger focus that encompasses entire sentences. We return to this proposal in Section 5.

1.3 Dual focus: Issue and evidence across languages

The main issue arising in connection with the realization of two foci in a single sentence is that, if every focus comes with its own prosodic head, two foci should correspond to two prosodic heads in a single intonation phrase. But can two equally prominent foci co-exist in one intonation phrase? Indeed, if this is the case, it would conflict with the ‘Culminativity Principle’ (Hyman, 2006), which requires a single and obligatory head per prosodic constituent—thus one per intonation phrase—or the formation of additional prosodic phrases that change the prosodic and tonal relationship of the sentence.

The past literature often presumed a negative answer, assuming that if two foci have to co-exist in one intonational phrase, one is more prominent than the other, resulting in a sequence of a subordinated secondary accent and a primary nuclear accent (Jackendoff, 1972; Truckenbrodt, 1995). Although empirical studies only exist for a few languages—i.e. English, German and Mandarin (see Eady, Cooper, Klouda, Mueller & Lotts, 1986; Kabagema-Bilan et al., 2011; Wang & Féry, 2015, 2017), these languages are shown to react differently to this conflict and to display different strategies as to how they realize dual focus. For instance, based on careful phonetic analyses, English and Mandarin allow two heads in a single prosodic domain of the size of an Intonation Phrase, while German adds the possibility of changing the phrasing, signaled by changes in the F₀ and by duration. Yet, these studies converge on the overall observation that the resulting prosodic structure of dual focus sentences amounts to more than just concatenating two single foci—the upcoming focus having a clear influence on how the first focus is realized, especially in the post-focal region. Anticipating our results slightly, we will see that this is not systematically the case in French. Although phrasing is prevalent in this language, information structure does not change it in any significant way. Section 2 presents the

experimental study we conducted to try and overcome the lack of a systematic study on French with respect to the prosodic correlates of prominence and phrasing.

2. Production experiment

2.1 Materials and participants

The written scripted material used to elicit production consisted of question-answer pairs. The experimental sentences contained two post-verbal sequences that varied according to their constituents, either two objects as in (2), or two adjuncts as in (3).

(2) [SV + object + object] (SVOO)

Jean-Marie a envoyé [un colis] [à ma sœur].

'Jean-Marie sent a package to my sister.'

(3) [SV + adjunct + adjunct] (SVAA)

Jean-Marie l'a envoyé [par la poste] [à Toulouse]

'Jean-Marie sent it via the post office to Toulouse.'

We note that neither the objects nor the adjuncts are assumed to be intrinsically right-dislocated.³ Both types of constituents are taken to be in their canonical post-verbal position as parts of the main clause. It is important to bear in mind that a right-dislocated object necessarily implies the presence of a clitic on the verb, whereas a right-dislocated adjunct has no clitic resumption whatsoever. Thus, it cannot be excluded that a given adjunct was right-dislocated in the speech of some informants, while that was impossible in the case of the object, since the sentence with a post-verbal object never contained an additional clitic pronoun. In other words, the question of whether the adjunct was realized as a right-dislocated element is not a primary concern. For now, we concentrate on the prosodic phrasing rather than on the syntactic structure of the sentences.

For each of the experimental sentences in (2) and (3), two additional factors were manipulated, Focus Type and Constituent Length (of the focused element). Focus Type

³ Right-dislocation is often realized in a separated prosodic phrase, separated from the main clause by a small break and presenting deaccenting. We return to this prosodic pattern in Section 3.3.

had four levels, illustrated in (4) for SVOO sentences. Thus, focus was tested (i) on the whole sentence “**all focus/AF**” (4a); (ii) on both post-verbal constituents “**dual focus/DF**” (4b); (iii) only on the first post-verbal constituent “**initial focus/IF**” (4c); (iv) only on the second post-verbal constituent “**final focus/FF**” (4d). The difference between AF and DF is that in an AF sentence, the two constituents under scrutiny are part of a larger focus, whereas each of the focused constituents in DF is a single narrow focus.

- | | | |
|-----|---|---------------|
| (4) | a. [Jean–Marie a envoyé un colis à ma sœur] _F
<i>‘Jean-Marie sent a package to my sister.’</i> | All-Focus |
| | b. Jean–Marie a envoyé [un colis] _F [à ma sœur] _F | Dual-Focus |
| | c. Jean–Marie a envoyé [un colis] _F à ma sœur. | Initial-Focus |
| | d. Jean–Marie a envoyé un colis [à ma sœur] _F | Final-Focus |

The second factor, Constituent Length, had two levels: the post-verbal elements were either both short (i.e., 3 or 4 syllables) as in (5a-b), or both long (i.e., 6 or 7 syllables) as in (5c-d).

- | | |
|-----|---|
| (5) | a. Short objects: Jean–Marie a envoyé [un colis] _F [à ma sœur] _F |
| | b. Short adjuncts: Jean–Marie l’a envoyé [par la poste] _F [à Toulouse] _F |
| | c. Long objects: Jean–Marie a envoyé [un colis important] _F [à sa voisine anglaise] _F
<i>‘Jean-Marie sent an important package to his British neighbour.’</i> |
| | d. Long adjuncts: Jean–Marie l’a envoyé [au fin fond de l’Argentine] _F [sans vraiment le faire exprès] _F
<i>‘Jean-Marie sent it in the middle of podunk Argentina without really doing it on purpose.’</i> |

To ensure that participants would interpret the experimental sentences with the intended information status on the target constituent(s), the different focus conditions were triggered via an explicit, congruent question. To illustrate for SVOO sentences, the AF condition (4a) was triggered via the broad question (6a), the DF condition with a question like (6b)

such that the two ‘wh’-words were always separated, the IF condition via (6c), and finally the FF condition via the question in (6d).

- (6) a. AF: **Qu’est-ce qu’il s’est passé?**
‘What happened?’
- b. DF: **Qu’est-ce que Jean-Marie a envoyé et à qui?**
‘What did Jean-Marie send and to whom?’
- c. IF: **Qu’est-ce que Jean-Marie a envoyé à ma sœur?**
‘What did Jean-Marie send to my sister?’
- d. FF: **À qui est-ce que Jean-Marie a envoyé un colis?**
‘To whom did Jean-Marie send a package?’

For each condition in this 2 x 4 x 2 design, we created four lexicalizations, which were pseudo-randomized with 20 fillers (~ 1:3 ratio) into two experimental lists (see Appendix I for all experimental material). Thus, each speaker uttered half of the 64 experimental sentences, and we obtained a total of 512 expected sentences (16 x 32), excluding fillers. However, some sentences had to be discarded. Indeed, one lexicalization of two long adjuncts was wrong in the IF condition (8 sentences): the object was not cliticized but appeared as an additional post-verbal constituent). Moreover, 17 sentences altogether contained disfluencies in the post-verbal domain in different conditions and lexicalizations. In sum, a total of 487 sentences entered the statistical analysis.

Sixteen female native speakers of Standard French (aged between 23 and 45) participated and were compensated monetarily for their time.

2.2 Recording procedure and analysis

Participants met a native speaker of French in a quiet laboratory space where they sat in front of a microphone and were given a handout that contained the experimental list to be read. The version of the handout given to the participants only contained the target sentences, and this to ensure that they paid attention to the question posed by the experimenter. Participants were encouraged to repeat the recording whenever they felt they made a mistake or produced an unnaturally or improperly read sentence. The recording

took about approximately an hour per participant. They were recorded at a sampling rate of 22.05 kHz with a 16-bit resolution with a head-mounted Shure microphone. The experimental sentences were extracted and saved as separate files.

The data was then automatically annotated for the words of interest by using the automatic phonetic alignment tool EasyAlign (Goldman, 2011). The phonetic correlates of phrasing investigated were F0 and duration. To obtain measurements on the target phrase, we used the script ProsodyPro (Xu, 2013). To extract continuous F0 contours, the vocal cycles were first calculated by Praat and then hand-checked for errors such as double-marking or pitch period skipping. While checking for spurious vocal pulse markings, segmentation labels were also added to mark the syllable boundaries. The duration of F0 periods was converted into F0 values automatically by ProsodyPro. The vocal pulse marking, segment labels, and F0 values were saved in separate text files for each utterance. In the next step, ProsodyPro calculated the highest and lowest F0 values as well as the duration of each syllable. For the graphical display of the intonational contours, the F0 values were smoothed using a trimming algorithm (Xu, 1999).

The data from our two dependent variables, F0max (discussed in section 2.6.1) and duration (see section 2.6.2), were log-transformed prior to statistical analysis in order to improve normality, and then analysed using generalized linear mixed-effects regressions implemented with the *lme4* library in the R environment (GPL-2j GPL-3, v.3.3.3; R Core Team, 2017). For the F0max variable, the measuring points considered for our analysis correspond to the F0 maxima measure taken on each constituent.

In all analyses, Participant and Lexicalization were included as random-effects. The three fixed factors Post-Verbal Sequence (OO, AA), Length (short, long) and Focus Type (IF, FF, DF, AF), were treatment-coded prior to analysis. For Focus Type, DF was always the baseline. To find out whether the fixed factors had an effect on the dependent variables in the prosodic marking of dual focus, we first built a full model that included the maximal random-effect structure (RES) justified by the data and the theoretical assumptions (i.e., random intercept and slopes by-item and by-participant for the fixed effects of interest, as well as their interaction), the main effects of the three fixed factors, the two-way interactions between each of the three fixed factors, and the three-way interaction between all of them. Thus, the full model had the following structure: Maximal RES + Post-Verbal

Sequence * Length * Focus Type. Then, in a stepwise fashion, we pruned off any non-significant interaction from the model, as long as the higher-order one was not significant either. We report on the final model by presenting estimates, standard errors and t-values, with any t-value exceeding 1.96 considered statistically significant with $p < .05$. P-values were obtained by likelihood ratio tests of the final models.

2.3 Research questions and hypotheses

Our study seeks to answer the following research questions:

- A. ***Prominence as a final high tone.*** How does prosodic prominence vary as a function of focus and givenness in the post-verbal domain? Given the consistent result in prior work on French that a non-final Φ is delimited by a rising tonal excursion and a longer duration at the end of Φ , we postulated that a non-final focused constituent should have a higher final high tone than a non-focused one. Moreover, the literature also finds an additional (but optional) high tone at the beginning of Φ , non-final and final ones alike. We hypothesized that this initial high tone should be present more often in a focused constituent than in a non-focused one. Duration should also be longer in a focused constituent than in a non-focused one.
- B. ***Effect of type of post-verbal sequence.*** Do objects and adjuncts differ in their phonetic correlates? Given results in our own prior work, we expected an adjunct to be more easily phrased independently from the preceding verb than an object. Consequently, there should be a higher F0max and a longer duration on the verb in case of adjuncts, suggesting the verb was phrased separately from the next constituent.
- C. ***Effect of prosodic length.*** How does prosodic prominence vary in long and short post-verbal constituents? Due to well-formedness constraints on the prosodic form of sentences, a long constituent is more prone to be phrased independently from the adjacent verb than a short one, and this for objects and adjuncts alike. Here again, if length does affect phrasing, as we found in our prior work, the verb is expected to end on a higher tone and to be lengthened when followed by a longer constituent than when followed by a short one.

- D. **The effect of information structure, especially dual focus.** Do sentences varying in their information structure vary in F0max and duration? Following results for Germanic languages, we may expect narrowly focused constituents to have higher F0max than given ones. And specifically for sentences containing a dual focus: do they amount to concatenating two single foci or do they present special properties not found in sentences with single focus?
- E. **Individual variation.** Because of the absence of lexical stress, a clear strategy for the expression of prominence is lacking in French: there is no designated syllable that obligatorily carries the highest F0 value. Therefore, we expect to see variation across speakers in the phonetic correlates of prominence in the location of high tones, but also in their scaling, the possibility to realize breaks, and the possibility to compress focused constituents: adjuncts should be more easily compressed than objects.

2.4 Results

2.4.1 F0 results

Before reporting on the statistical analysis, the results are illustrated via four figures, representing the pooled normalized F0 results of all speakers for the test sentences per Length and Focus type. Thus, the plotting points represent pooled averaged measurements on each constituent (subject, verb, post-verbal constituent 1 and post-verbal constituent 2), with ten measurements for each syllable per constituent (as provided by ProsodyPro). Figure 1 shows SVOO sentences in the short (left panel) and long condition (right panel).

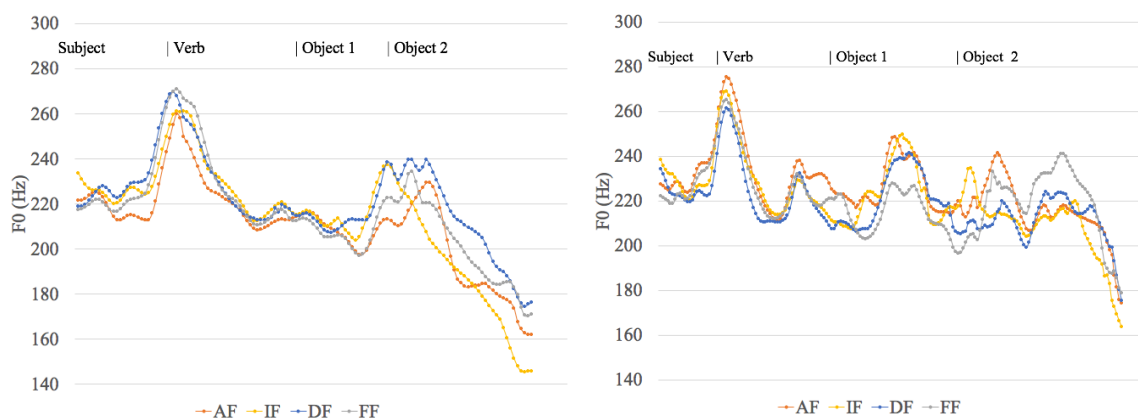


Figure 1. Pooled normalized intonational contours per focus condition for the SVOO sentences with two short objects (top), and two long objects (bottom).

Similarly, Figure 2 illustrates SVAA sentences where post-verbal adjuncts are either both short (left), or both long (right).

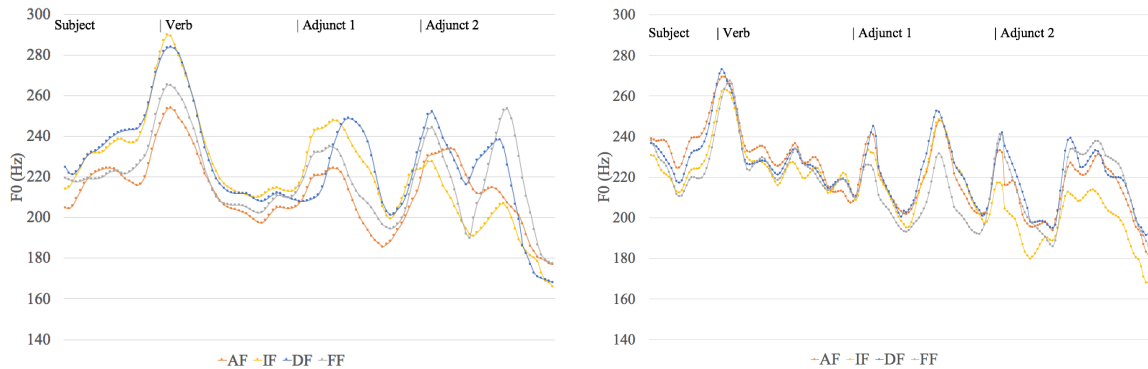


Figure 2. Pooled normalized intonational contours per focus condition for the SVAA sentences, with two short adjuncts (left) and two long adjuncts (right).

Visual inspection of Figure 1 and 2 does not lead to a straightforward interpretation, yet some important generalizations can be made. First, some properties of the tested sentences did not vary much:

- A striking property is that the subjects and the beginning of the verbs were realized in the same way in all conditions, except in the short adjuncts where the different focus conditions triggered different height on the subject and verb. In the other conditions, the variation in the pitch contours began at the boundary of the verb and was most obvious on the post-verbal constituents.
- A non-final constituent always bears a final high tone. This high tone was present in all focus conditions, and thus it seems to be dependent on phrasing rather than on information structure (Research question A).⁴ There may be a difference in the

⁴ In three of the SVAA sentences where this last tone is falling without that the next constituent is deaccented (see Section 3.3), the fall sounds odd, as if the speaker is making a mistake, or as an afterthought.

way the first post-verbal constituent was phrased relatively to the verb, but not in the way the two constituents were phrased relatively to each other: they were always in separate prosodic phrases as testified by the ubiquitous high boundary tone at the end of the first post-verbal constituent (see also the pitch tracks in Figures 3-5 for illustrations).

- When the constituent was long, there was also an additional high tone. We discuss these additional high tones in Section 3.1.
- All speakers realized the sentences as declaratives: all of them ended in a low tone (or in some instances at mid-level, see Figure 1a for an example). In cases where the last constituent was subject to compression (see Section 3.3), the first constituent ended with a falling contour.
- A further common property is that the IF condition often triggered post-focal compression. In the figures 1 and 2, the final contour of the yellow line is always lower than in the other conditions. As for the FF condition, it did not regularly trigger a higher contour on the focused constituent; it did in the long OO and in the short AA, but in the other cases, it resembles the AF and DF conditions.

Second, a couple of other obvious differences arose across the different conditions, the first and most important one being the number of high tone peaks in each constituent and their position (see Section 3.1 for more detail), and the second being the height of the high tones (see statistical analysis just below). For instance, the first post-verbal constituent always reached the highest value in the IF condition, and the second post-verbal constituent was sometimes the highest one in the FF condition.

Turning now to our statistical analysis, a first result concerns **the F0 height of the high tone on the final syllable of the verb before objects and before adjuncts**. We ran a mixed-effect linear regression to the normalized F0max (our first dependent variable) of the verb for whole data set (short and long SVOO and SVAA sentences altogether). The full model revealed no significant triple interaction nor any significant double interaction between the three fixed-effect predictors. After pruning these out, the final model revealed a main effect of Post-Verbal Sequence (PVS) ($\beta = -5.645$, $SE = 1.054$, $t = -3.153$, $p < 0.001$), suggesting that the verb's F0max was significantly lower when followed by an

object rather than an adjunct. There was also a main effect for Constituent Length ($\beta = 7.24, SE = 4.62, t = 2.52, p < .001$), suggesting that the F0max of the verb was higher when followed by a long versus a short Post-Verbal Sequence. Finally, there was no main effect of Focus type ($\beta = -6.31, SE = 5.34, t = -1.28, p < 0.2$).

A second result concerns the effect of Constituent Length, and more specifically **the F0 height of the PVS (Post-Verbal Sequence) itself, i.e. on the sequence that includes two long objects as compared to two short ones, and similarly for two adjuncts**. Here, we ran a mixed-effect linear regression to the normalized F0max of the first and the second post-verbal constituents. The full model revealed a significant double interaction between Constituent Length and PVS ($\beta = -8.101, SE = 5.76, t = -2.03, p < .05$), and a main effect of the following two individual factors; Constituent Length ($\beta = -17.30, SE = 6.01, t = -2.88, p < .001$) and PVS ($\beta = -8.24, SE = 7.99, t = -2.57, p < .001$). Given the significant interaction between PVS and Constituent Length, we repeated a similar analysis for the subset of SVOO and SVAA sentences independently (i.e. the statistical model conducted on each subset of the data contained RES (random-effect structure) + Constituent Length * Focus type). Here again, we found an effect of the sole factor Constituent Length for SVOO sentences ($\beta = 12.83, SE = 4.19, t = 5.34, p < .001$), indicating that the F0max on the two objects was consistently higher in the long condition compared to the short one. The same factor Constituent Length also had an effect for SVAA sentences ($\beta = 4.21, SE = 2.52, t = 3.04, p < .001$), although smaller than for objects. There was no significant interaction between the two factors Constituent Length * Focus type, neither for SVOO sentences ($\beta = 12.63, SE = 8.11, t = 1.16, p < .08$), or for SVAA sentences ($\beta = 10.54, SE = 8.20, t = 1.21, p < .08$). In sum, F0 on the post-verbal sequences is higher when the post-verbal constituents are long than when they are short.

The third and last result for F0 height concerns **the effect of Focus Type**. Recall that this analysis seeks to assess the role of dual focus (DF) as compared to the other focus conditions; AF, IF and FF. Here, we ran a mixed-effect linear regression to the pooled F0max data (for SVOO and SVAA sentences, both short and long altogether) at three points: (i) on the verb, (ii) on the first post-verbal constituent, and (iii) on the second post-verbal constituent. The final model revealed a significant interaction between PVS and Constituent Length ($\beta = 10.74, SE = 6.18, t = 2.09, p < 0.01$). There was also a main effect

of Constituent Length ($\beta = -11.74$, $SE = 4.22$, $t = -3.05$, $p < .001$) and of PVS ($\beta = 5.61$, $SE = 6.23$, $t = 2.31$), but no main effect of Focus Type ($\beta = -1.02$, $SE = 3.45$, $t = -0.43$, $p < .5$). Given the significant double interaction PVS * Length, a subsequent, similar analysis was conducted on SVOO and SVAA independently.

First, we discuss the results for SVOO sentences, illustrated visually in Figure 3 and 4, which represent the pool normalized F0max means for the VP domain with the two panels representing a different length condition for each sentence type (short PVS on the left, long PVS on the right).

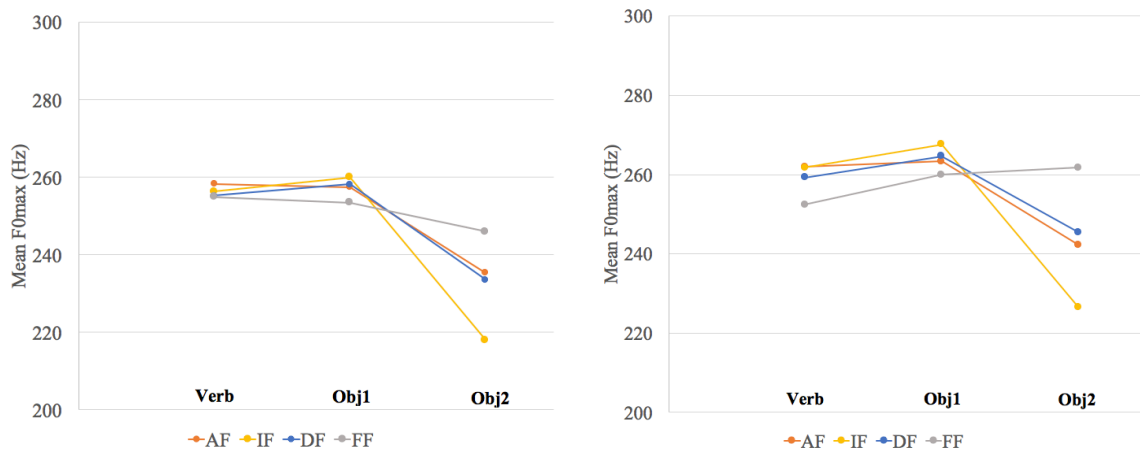


Figure 3. Pooled normalized means for F0max per focus condition for the SVOO short and long sentences (left and right, respectively).

Statistically, Table 1 shows the results of mixed-effect linear regressions ran on the normalized F0max data for the verb, the first and the second post-verbal object, with full models including RES + Focus Type * Constituent Length. Again, the DF condition served as a baseline and was compared to the other conditions: AF, IF and FF.

	F0max V			F0max Obj1			F0max Obj2		
	β	SE	t	β	SE	t	β	SE	t
Intercept	261.6	7.8	33.4	264	5.9	44.7	242.2	6.5	37.9
FocusTypeIF	4.8	5.3	1.4	-7.8	4.8	-1.6	-17.5	6.4	-2.68
FocusTypeFF	7.9	5.5	0.9	-0.4	4.8	-1.2	3.7	6.5	2.12
FocusTypeAF	5.1	5.5	1.2	0.1	4.6	0.36	5.9	6.4	0.57

LengthLong	5.5	5.3	2.4	18.6	4.7	3.82	39.9	9.1	6.2
FocusTypeIF:LengthLong	8.1	7.6	1.1	-8.2	6.9	-1.41	-3.8	9.1	-2.72
FocusTypeFF:LengthLong	1.2	7.6	0.7	-1.01	6.9	-1.15	4.04	9.2	2.42
FocusTypeAF:LengthLong	1.6	7.6	0.35	2.77	6.8	0.41	5.54	9.2	0.47

Table 1. Significance results for Focus Type (and its interaction with Length) for SVOO sentences in the normalized F0max data assessed for the verb, the first and the second post-verbal constituent (significant results are bolded).

Results of interest for SVOO can be summarized as follows: Overall, there is no significant difference between the F0max of the verb in any of the Focus Type conditions compared to DF. Furthermore, when comparing AF to DF, there are no distinct differences on either of the post-verbal constituents, and indeed when visually inspecting the bottom panels of Figure 3, AF and DF data look very similar. The effect of IF is however seen on the second post-verbal constituent when compared to the DF condition—the second object is significantly lower when it is given than when it is focused, and there is a significant interaction with the factor Length. Finally, comparing DF to FF, we do notice differences in the post-verbal domain as well. Indeed, the F0max of the second object in the FF condition is significantly higher than in DF, and there is a significant interaction with the factor Length.

Next, we discuss the results for SVAA sentences, as reported in Table 2 (see also Figure 4). Starting with the verb: The only condition that shows a significant difference on verb F0max with the DF condition is when the focus is final (FF). Indeed, the V F0max is much lower in the latter condition, speaking for cancelling the phrase boundary between the verb and a following adjunct when this adjunct is given. As for the post-verbal constituents, when visually inspecting the bottom panels of Figure 4, the data for AF and DF are again strikingly similar. But when comparing DF to IF, the effect of Focus Type is significant on the second adjunct, whereby its F0max is drastically lower in the latter condition. However, we do not find an interaction with the factor Length, suggesting that both short and long adjuncts behave similarly in that respect. Finally, comparing DF to FF, we see differences on both post-verbal constituents, with the F0max of the first adjunct being significantly lower and that of the second being significantly higher in the FF condition, but no significant interaction of Length and Focus Type.

	F0max V			F0max Adj1			F0max Adj2		
	β	SE	t	β	SE	t	β	SE	t
Intercept	255.5	6.9	36.9	265	6.8	38.9	236.6	8.5	27.5
FocusTypeIF	6.2	6.1	0.67	4.1	6.5	0.63	-14.4	8.6	-3.15
FocusTypeFF	7.9	6.2	-2.48	-0.4	6.5	-2.3	6.8	8.5	5.08
FocusTypeAF	6.2	6.1	1.02	-2.02	6.6	-0.3	5.1	8.4	0.37
LengthLong	9.7	6.2	2.11	23.8	6.6	3.57	14.4	9.6	1.46
FocusTypeIF:LengthLong	4.7	8.6	0.54	27.8	9.4	-1.41	-2.3	13	-1.38
FocusTypeFF:LengthLong	16.9	8.6	-2.74	17.9	9.1	1.97	10.5	13	1.09
FocusTypeAF:LengthLong	8.4	8.5	1.05	9.5	9.2	1.03	3.7	13	0.82

Table 2. Significance results for Focus Type (and its interaction with Length) for SVAA sentences in the normalized F0max data assessed for the verb, the first and the second post-verbal constituent (significant results are bolded).

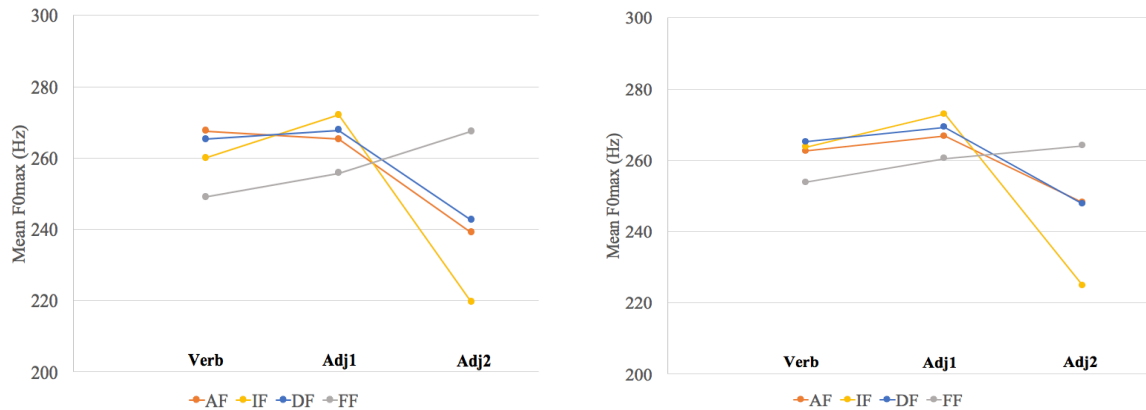


Figure 4. Pooled normalized means for F0max per focus condition for the SVAA short and long sentences (left and right, respectively).

2.4.2 Duration results

Similar to the analyses conducted for F0, we were interested in the effect of Post-Verbal Sequence and Focus Type on duration—our second dependent variable. We note that we did not analyse the effect of Constituent Length here since long constituents necessarily had a longer duration compared to short ones.

First, we examined **the effect of PVS (OO and AA) on the duration of the verb.** We conducted a linear mixed-effect regression to the normalized data set of V duration—the full model included the maximal RES (random intercept and slopes by-item and by-

participant for the two fixed effects of interest, as well as their interaction) and the interaction of PVS and Focus Type. After pruning off the non-significant interaction between the two factors, results from the final model show a main effect of the sole factor PVS ($\beta = 0.45$, $SE = 0.0756$, $t = 14.76$, $p < .001$); the verb was longer when followed by adjuncts (SVAA sentences) than when followed by objects (SVOO sentences).

Second, we were concerned with **the effect of Focus Type on the duration of each of the post-verbal constituents and the verb**. This comparison was statistically assessed via fitting three linear mixed-effect regression models—one on the data set for the duration of each constituent—for short and long sentences separately (so six models total). Here again, each model included the maximal RES, and the fixed-effect factors Focus Type and PVS and their interaction. The DF condition served as a baseline here too and was compared to the other three focus conditions.

Concerning the **duration of the verb** in short and long sentences, results from the final model revealed a main effect for PVS in both Length conditions (short: $\beta = 0.587$, $SE = 0.026$, $t = 2.6$, $p < .001$; long: $\beta = 0.423$, $SE = 0.035$, $t = 3.75$, $p < .001$), but no main effect of Focus Type (i.e. no significant difference between DF and any of the other three conditions). This suggests that the duration of the verb is really only affected by the type of constituent that follows, and this when the constituents are short or long.

Results of the final model for the **duration of the first post-verbal constituent in short sentences** revealed: A main effect of PVS ($\beta = 0.071$, $SE = 0.045$, $t = 2.45$, $p < .001$), a significant effect of Focus Type between DF and IF ($\beta = 0.032$, $SE = 0.016$, $t = 2.14$, $p < .001$), and a significant interaction between PVS and Focus Type for DF compared to IF ($\beta = 0.069$, $SE = 0.0156$, $t = 2.24$, $p < .001$), and DF compared to FF ($\beta = -0.065$, $SE = 0.016$, $t = -2.73$, $p < .001$). **For long sentences**, the final model revealed a significant effect for Focus Type between DF and IF ($\beta = 0.131$, $SE = 0.04$, $t = 3.46$, $p < .001$) and DF vs. FF ($\beta = -0.101$, $SE = 0.035$, $t = -2.48$, $p < .001$), and a significant interaction for Focus Type * PVS when DF was compared to IF ($\beta = 0.199$, $SE = 0.086$, $t = 2.71$, $p < .001$).

Finally, results from the final model on the **duration of the second post-verbal constituent in short sentences** revealed: A significant effect of Focus Type between DF and IF ($\beta = -0.008$, $SE = 0.0173$, $t = -2.27$, $p < .001$), and between DF and FF ($\beta = 0.0175$, $SE = 0.0168$, $t = 2.53$, $p < .001$). **For long sentences**, there was a significant effect of Focus

Type when DF was compared to IF ($\beta = -0.124$, $SE = 0.031$, $t = -3.59$, $p < .001$) and to FF ($\beta = 0.107$, $SE = 0.038$, $t = 2.72$, $p < .001$).

In sum, these results suggest that duration is not significantly affected by the focus type encoded in the sentence, but does vary according to the constituent that follows the verb along with its prosodic length.

3. Individual variation

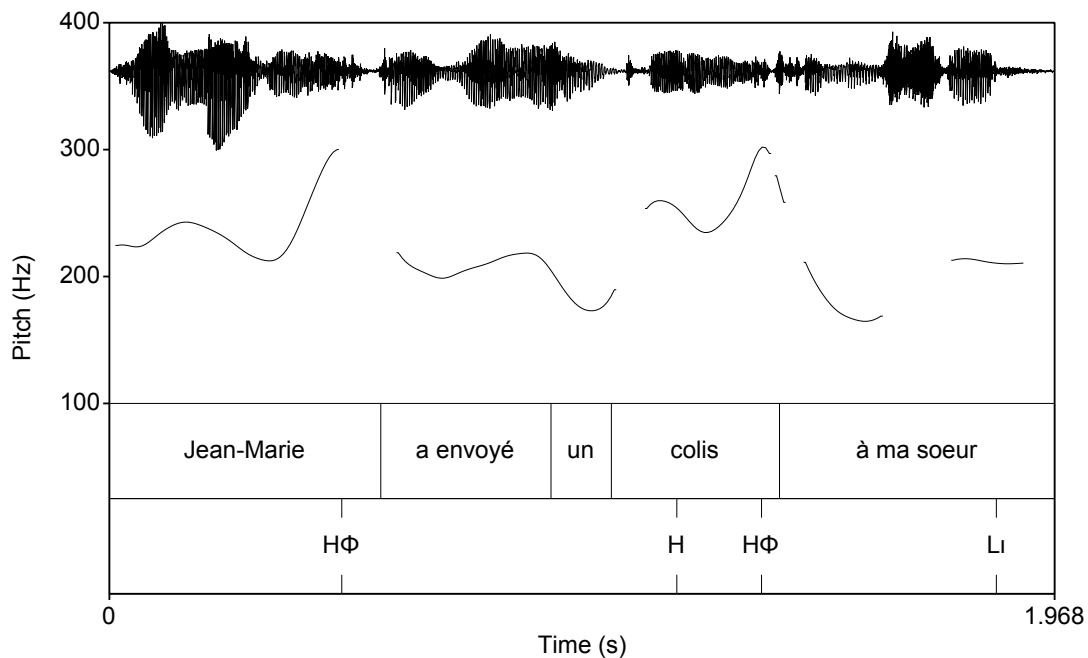
It is important to note that the pooled normalized results (Figures 1-4) and the statistical analyses can obliterate the individual differences that may appear among speakers. Yet, speakers use different strategies to realize focus prosodically, and cross-speaker variations appears pervasive in French. We address them in this section, as posited in the research question E, since in our view, variation is one of the reasons why it is so difficult to account for French prosodic structure in simple terms. An understanding of the source of this difficulty can only be achieved by a careful survey of variation. In the following discussion, we focus on the factors that were most affected by variation: Number and position of the high tones in the post-verbal constituents (section 3.1); tone scaling relationship between them (3.2); deaccenting of entire final phrases (3.3); and perceived breaks between the post-verbal constituents (3.4). In this section, 480 sentences were taken into consideration. From the expected 512 sentences, the same 25 mentioned in Section 2.2. were discarded, and an additional 7 sentences were also removed because they contained hesitations between constituents that did not affected the statistical results but did affect the results of the individual variation.

3.1 Number and position of high tones in each post-verbal constituent

In the short constituents, there generally was one high tone per non-final constituent that lied on the last syllable of the constituent. This high tone is best analysed as a prominent phrase boundary (see Section 1 for references), and is annotated as H_{Φ} in the following figures, to indicate that it is the boundary tone of a prosodic phrase. In some cases, however, there was an additional high tone earlier in the phrase, as illustrated in the top panel of Figure 5, where the first syllable of *colis* ‘package’ has an additional high tone. Since this tone is not a boundary tone, we annotate it as H. This figure compares to the bottom panel

of Figure 5 where there is only one final high boundary tone on the last syllable of *colis*. Both sentences were produced in the all-focus condition and both of them show the final L_1 responsible for the final fall at the end of the sentence.⁵

We found six additional high tones in the first post-verbal object in the AF condition, out of 32 realizations, thus in 19% of such sentences. In the DF and IF conditions, we found more additional high tones on the first constituent, 15 and 20 respectively, see Table 5 for a survey. In the FF condition, only two such tones were produced on the given first constituent. Additional high tones were also found in the second post-verbal constituent, though not as frequently as in the first one. Only in the FF condition did we count 12 additional high tones. In sum, focused constituents had more additional high tones than given ones.



⁵ Both reviewers also noted a difference on the verb in these two pitch tracks. Indeed, in the top one, the verb has a final high tone, though much lower than the other ones, and is thus more clearly separated from the following object than in the bottom one. This difference is only relevant for the phrasing between the verb and the following object, but does not matter for the additional tone on the object discussed here.

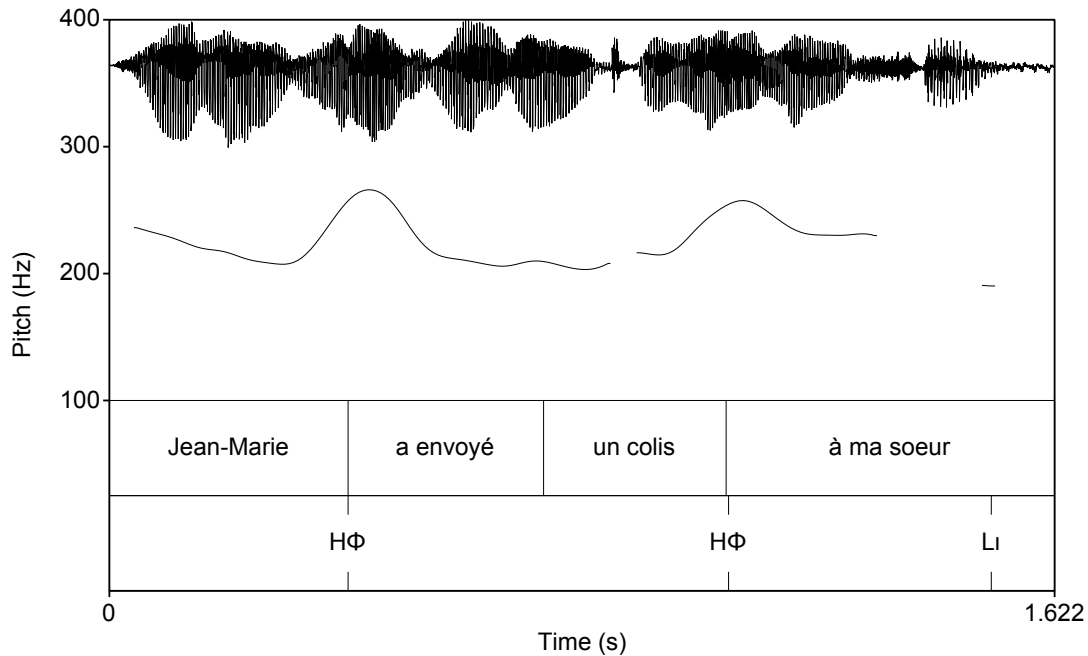


Figure 5. A high tone on the first syllable of *colis* ‘package’ and another one on the second syllable of the same word. This sentence ends at a mid-level, rendering the word *sœur* ‘sister’ prominent (top). A high tone on the last syllable of *colis* and a final falling contour on *sœur* (bottom).

The most interesting lexicalization is the one which has a disyllabic final word (*à leur rival* ‘to their rival’), the other three lexicalizations had a final monosyllabic word. Out of the eight utterances for the sentence containing the disyllabic word in the FF condition, an additional high tone was realized once on *à* ‘to’, twice on *leur* ‘their’, and three times on the first syllable of *rival* ‘rival’. Only two speakers did not add a high tone there. Similarly, the DF condition exhibited five additional high tones, all of them on *ri-* (of *rival*). In the other sentences, where the final word was monosyllabic, the last syllable was often longer in duration or the final word was pronounced with a rising-falling tone or with a mid-tone, all realizations being perceived as prominent. Some additional tones were present in these sentences too, usually on the preposition *à* or on the following possessive.

	AF 1 n=32	AF 2 n=32	DF 1 n=32	DF 2 n=32	FF 1 n=32	FF 2 n=32	IF 1 n=31	IF 2 n=31
Total	6 19%	4 12.5%	15 47%	6 19%	2 6%	12 37.5%	20 64.5%	2 6.5%

Table 5. Non-final high tones on the two constituents in sentences with two short objects. AF1 stand for first post-verbal constituent in the AF condition, etc.

The results for sentences with two adjuncts are partly similar to those with two objects (more additional tones in focused constituents), except for the fact that the additional high tones were altogether more numerous. This difference may correlate with the fact that there were more disyllabic final words in the sentences with adjuncts (and thus more space to realize an additional tone), or with the difference in phrasing between objects and adjuncts (Destruel & Féry 2019).

	AF 1 n=31	AF 2 n=31	DF 1 n=32	DF 2 n=32	FF 1 n=31	FF 2 n=31	IF 1 n=32	IF 2 n=32
Total	5 16%	6 19%	15 47%	29 91%	2 6.5%	28 90%	19 59%	3 9%

Table 6. Non-final high tones on the two constituents in sentences with two short adjuncts.

In sentences with long adjuncts, the post-verbal constituents under scrutiny mostly consisted of a prepositional phrase containing another (embedded) phrase, or of a noun and an adjective. They are thus syntactically and prosodically more complex than the sentences with two short constituents. The typical pitch contour has again a high boundary tone on each constituent, usually the last word, as a prominent H_φ boundary, but often also on the embedded constituent, as illustrated in Figure 6, a sentence with final focus. In this pitch track, there are several additional high tones. Downstep is interrupted on the last syllable of *américain*, suggesting that *américain* carries a more important boundary than *collègue*.

(8) Bernadette a présenté [son collègue [américain]] [à ma belle-sœur [canadienne]].

‘Bernadette introduced her American colleague to my Canadian sister-in-law.’

law.’

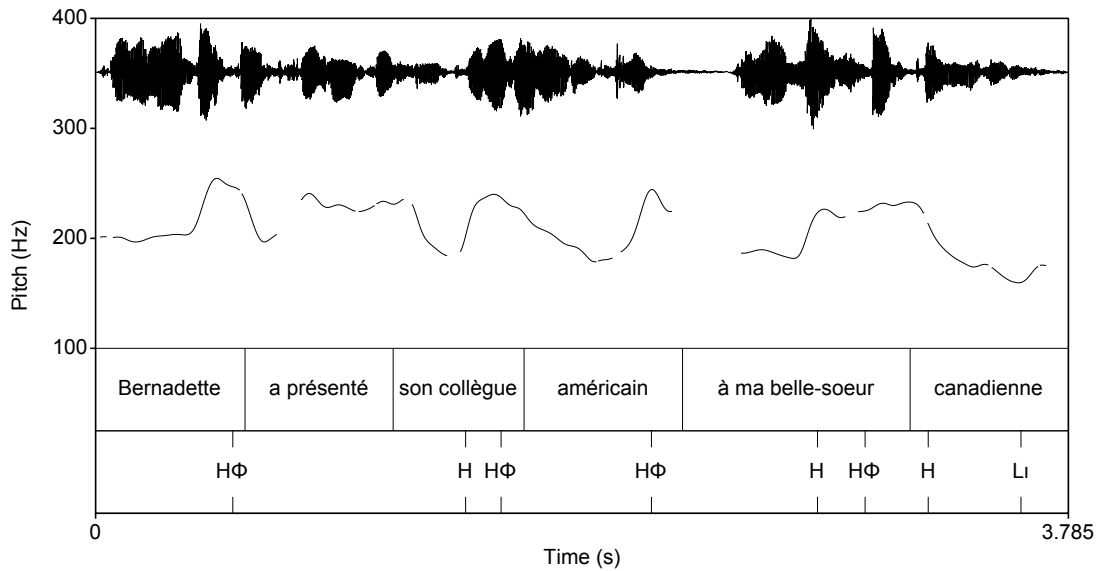


Figure 6. Additional high tones on a sentence with long objects.

We also counted the additional high tones in the long sentences. Results, as reported in Tables 7 and 8, reveal here again a tendency to realize more non-final high tones in adjuncts than in objects, but also more in focused constituents than in non-focused ones.

	AF 1 n=31	AF 2 n=31	DF 1 n=32	DF 2 n=32	FF 1 n=31	FF 2 n=31	IF 1 n=32	IF 2 n=32
Non-final	6 19%	4 13%	11 34%	5 16%	2 6.5%	7 23%	14 44%	1 3%

Table 7. Non-final high tones on the two constituents in sentences with two long objects.

	AF 1 n=28	AF 2 n=28	DF 1 n=27	DF 2 n=27	FF 1 n=27	FF 2 n=27	IF 1 n=20	IF 2 n=20
Non-final	8 28.5%	12 43%	8 30%	15 55.5%	3 11%	15 55.5%	14 70%	5 25%

Table 8. Non-final high tones on the two constituents in sentences with two long adjuncts.

To sum up, the most frequent occurrences of an additional high tone happened in a focused constituent, thus in both constituents in DF, in the last one in FF and in the initial one in IF, speaking for a cue of focus. Furthermore, more additional high tones were realized on the long constituents than in the short ones. Most interesting is that the high tone can appear in several locations, rather than in a single one, and speakers seem to behave quite freely with respect to which syllable they chose for hosting a non-final high tone.

3.2 Scaling relation between the two post-verbal constituents

Let us turn to pitch scaling, a correlate that is seldom considered in the literature on French prosody. We think it is a crucial cue in French that should be studied carefully. As in other better studied languages, the height of high tones is an indicator of prominence, and even though a high tone does not have the same function as in Germanic languages, where it signals the head of a prosodic domain, it does render a word or a constituent prominent in French as well. And the higher it is, the more prominent it becomes. Inversely givenness can decrease the height of a high tone. For these reasons, tonal scaling is a prosodic reflex of information structure, see for instance Kügler & Féry (2017) for post-focal downstep in German.

First, notice that a large downstep typically takes place between the subject and the verb, but none between the verb and the next constituent. Visual inspection of Figures 1 and 2 reveals a default downstep pattern in the two post-verbal constituents, except for DF and FF. In AF and IF, the second constituent is lower than the first one—even more so in the IF condition.

Turning to individual variation, in the sentences with short objects, most sentences presented downstep, see Table 9 for quantification. In the following, downstep means a lower F0max in the second post-verbal constituent than in the first one. The difference between the two must be of at least 5 Hz for counting as downstep, considering the highest F0 in both phrases. Surprisingly, downstep is also often found in the DF condition, as often as in the AF condition. Only the FF condition cancels downstep of the second constituent more often than in the other conditions. In this case, the highest tone is at the same level or

higher as the highest tone of the first constituent in 18 cases (out of 31 cases). However, there is a difference in the amount of downstep in the different conditions.

In the case of two adjuncts, downstep also takes place in the IF condition (except in two cases). In the AF and DF conditions, downstep is less regular than in the sentences with two objects, as can be seen in the right-hand column of Table 9.

	Downstep of the final constituent (obj)		Downstep of the final constituent (adj)
AF (n=32)	27 (84%)	AF (n=31)	13 (42%)
DF (n=32)	24 (75%)	DF (n=32)	16 (50%)
IF (n=32)	32 (100%)	IF (n=31)	29 (93.5%)
FF (n=31)	13 (42%)	FF (n=31)	15 (48%)

Table 9. Downstep of the second phrase relative to the first one in two short constituents.

In the long sentences, downstep is less frequent altogether, especially in the sentences with adjuncts, although it is not rare. It is again most frequent in the IF condition, and least frequent in the FF condition. In AF and DF, it is much less frequent than in the short sentences.

	Downstep of the final constituent (obj)		Downstep of the final constituent (adj)
AF (n=31)	19 (61%)	AF (n=28)	16 (57%)
DF (n=32)	12 (37.5%)	DF (n=27)	10 (37%)
IF (n=31)	22 (71%)	IF (n=20)	15 (75%)
FF (n=31)	11 (35.5%)	FF (n=27)	6 (22%)

Table 10. Downstep of the second phrase relative to the first in two long constituents.

These data suggest that tonal scaling across constituents is an important cue to focus and givenness—a cue probably as important as the presence of an additional tone early in the

constituent. The length of the constituents also plays an important role: in the long conditions, downstep is less frequent than in the short conditions.

3.3 Deaccenting of entire final phrases

We call ‘deaccenting’ or ‘compression’ a realization in which an entire final constituent is realized with a low and flat contour. Compression can be considered an extreme case of downstep because the entire constituent is realized on a much lower level than the preceding one. Figure 7 illustrates the sentence in (9), realized in the IF condition. As is visible in the pitch track, the last word of the first post-verbal constituent *poste* ‘post office’ has a falling contour; it is not downstepped relative to the preceding verb and is thus perceived as prominent.

- (9) Jean-Marie l’a envoyé par la poste à Toulouse.
‘Jean-Marie sent it via the post-office to Toulouse.’

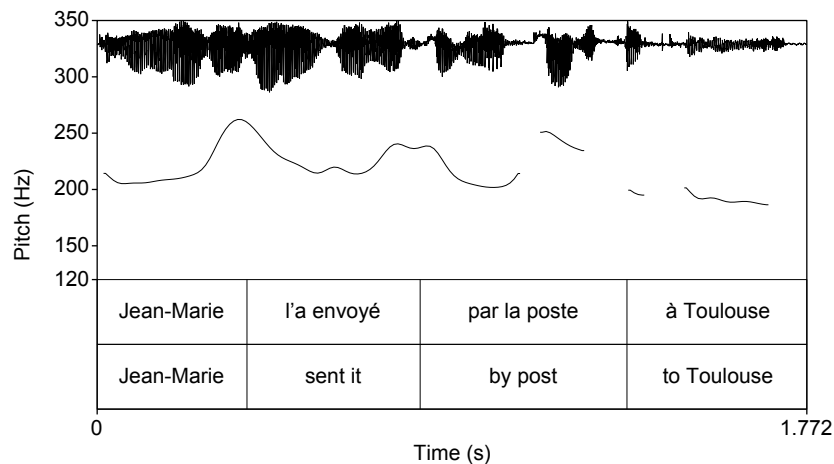


Figure 7. Deaccenting of a final constituent in the IF condition.

Compression of the final constituent was often present in the IF condition, but not always (59 times out of 114 possible contexts, thus 52%), see Table 11 for details. It is much more frequent in the short sentences than in the long ones, but there is no difference between objects and adjuncts.

Constituent & Length condition	Compression of the final constituent
Short objects (n=32)	24 (75%)
Short adjuncts (n=31)	22 (71%)
Long objects (n=31)	7 (22%)
Long adjuncts (n=20)	6 (30%)

Table 11. Compression of the final constituent in the IF condition for all lexicalizations.

Interestingly, compressing the final constituent in other conditions than IF was not infrequent. It happened 16 times altogether: 6 times in AF, 7 times in DF, and even 3 times in FF. One speaker, as seen in Figure 7, compressed the final constituent especially often (9 times) who also often ended her focused constituent at mid-level. This speaker is one of two speakers who always compressed the last constituent in the IF condition. Six other speakers compressed one or two sentences each in IF. In sum, only two of the 16 speakers always compressed a final given constituent—most of the speakers did it optionally, but more often when the constituents were short than when they were long.

We conclude that, although this is the preferred context, compressing a second post-verbal constituent does not necessarily mark givenness. It may be the reflex of right-dislocation, which is itself not necessarily triggered by givenness.

3.4 Perceived breaks between the post-verbal constituents

The last factor subject to variation is an audible break between the post-verbal constituents. It is realized with a short silence, or a glottal stop or an extra high tone on the preceding final syllable, that is then also a bit lengthened. A break is realized a lot more often between long constituents than between short ones, and also more often after a focused constituent than in the AF constituent or before the final focus of FF, as can be seen from the numbers reported in Table 12.

Perceived break between the post-verbal constituents	AF	DF	IF	FF
Short objects	2 (n=32)	17 (n=32)	11 (n=32)	8 (n=31)
Short adjuncts	6 (n=31)	20 (n=32)	20 (n=31)	16 (n=32)

Long objects	21 (n=31)	29 (n=32)	24 (n=31)	22 (n=31)
Long adjuncts	25 (n=28)	25 (n=27)	17 (n=20)	15 (n=27)
Total	54 (n=122)	91 (n=123)	72 (n=114)	61 (n=121)

Table 12. Audible breaks between the two post-verbal constituents in all conditions.

Altogether, this strategy for marking phrasing was extremely frequent in our data, but with large differences among the conditions. The main differences appear between the short sentences (100 breaks total) and the long sentences (186 breaks total). However, among the short sentences, there is another divide between objects and adjuncts. There were more breaks between adjuncts than between objects. And among objects, the most probable context for a break is after a focused constituent (DF and IF).

To sum up this subsection, the presence of a short break between the post-verbal constituents appears to be a further indication of focus.

4. Perception study

4.1 Methods

The perception experiment was conducted to assess whether French listeners make use of the differences in prosodic and phonetic realization to distinguish between different focus contexts. The experimental sentences elicited and recorded in the production experiment served as target items in this experiment—they were presented acoustically and in writing to the participants, together with the four questions corresponding to the four focus conditions tested (IF, FF, DF, AF). Only one of the questions was congruent with the sentence heard—participants had to select which one they thought was the appropriate one. An illustration of the test screen appears in Fig.8, where the questions correspond to dual focus, all focus, initial focus and final focus, respectively (see also examples (4) and (6) for other sample stimuli).



6. Jean-Marie a envoyé un colis important à sa voisine anglaise

- Qu'est-ce que Jean-Marie a envoyé et à qui?
- Qu'est-ce qui s'est passé?
- Qu'est-ce que Jean-Marie a envoyé à sa voisine anglaise?
- À qui est-ce que Jean-Marie a envoyé un colis important?

Figure 8. A sample trial for dual focus in the perception experiment.

A total of 26 sentences from the short-short objects (7 sentences), short-short adjuncts (6 sentences), long-long objects (6 sentences) and long-long adjuncts (7 sentences) were used as experimental sentences in this study. All four conditions were included: There were 6 AF, 6 IF, 6 FF and 8 DF sentences. For each sentence, we chose 3 different realizations each, thus resulting in a set of 132 utterances. These were organized in two lists of 66 sentences each, along with fillers. The study was conducted online, on the free SoSciSurvey (<www.soscisurvey.de>) platform and took approximately 20-30 minutes. A total of 79 French native speakers participated, but only 50 completed the entire study, the responses of whom we report in the next section.

4.2 Results

The results appear in Figure 9, which represents the percentages of responses for each Focus Type condition.

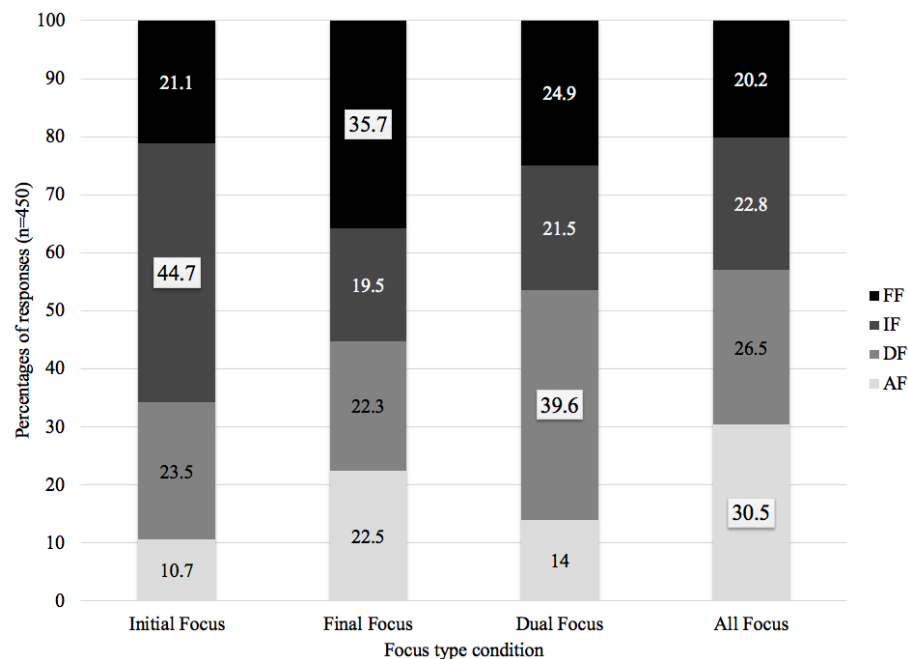


Figure 9. Results of perception study in percentages per Focus Type condition

Visual inspection shows that in the Initial, Final and Dual Focus conditions, speakers display a preference for the correct response, with the best results obtained for the IF condition (44.7% rate of correct responses). In the AF condition, on the other hand, percentages of responses are more evenly distributed across the four levels of responses. Statistically, we restricted our attention on the Focus Type condition of interest—Dual Focus—for which we examined whether the differences in the responses selected were significant. To do so, we ran a series of logistic mixed-effect models, each concentrating on two responses. Results revealed a significant effect of Response when comparing the rate of response between DF and IF ($\beta = 0.408$, $SE = 0.219$, $z = 1.86$, $p < .05$), between DF and FF ($\beta = 0.449$, $SE = 0.224$, $z = 2.22$, $p < .05$), and even more so when comparing DF and AF ($\beta = 1.477$, $SE = 0.347$, $z = 4.25$, $p < .001$). This suggests that listeners tend to select the congruent question in DF Focus Type condition, that is, the question that contains two wh-words (see 6b). This is a welcome result because it also indicates that the speakers in our production study completed the task well, with the intended information structure.

We speculate that the post-focal compression that was realized by a number of speakers was decisive in the recognition of an initial focus, whereas the other conditions did not vary that much. This is especially obvious in comparison to results for the same experience conducted in German (see Wang & Féry 2017), where the preference for the correct response is much larger in each condition. In German as well, initial focus elicited the best results, but the other conditions were also much better recognized than for French. We take these data to suggest that French speakers do use prosodic cues for communicating the information structural context, but that the prosodic cues are not strong enough to straightforwardly disambiguating the information-structural context in which the sentence occurred. This explains the mixed results of the perception experiment, especially when compared to a language where the prosodic cues are much stronger, i.e. German.

5. Discussion

In this section, we provide elements of answers to the research questions formulated in Section 2.4.

Research question A: *Prominence as a final and/or initial high tone*. Except for the cases

of final compression, a final high tone was always present at the end of the first constituent, speaking in favour of an analysis of this high tone as an indicator of phrasing rather than as prominence *stricto sensu*. But, as far as statistical averaged measures are concerned, we could not always find a clear correlation between the focused status of the first post-verbal constituent and the height of this final high tone. We could also not find a systematic initial high tone in a focused constituent. What we could find, however, was a difference in height of the second constituent in dependence of single focus: it was lower in IF and higher in FF than in AF and DF. Other clear correlates of focus were additional high tones in focused phrases—not always initial, and not obligatorily present, additional breaks between the constituents and suppression of downstep. In the IF condition, post-verbal compression was observed much more often than in the other conditions, although it was occasionally found in all conditions. Correlating these results with the literature on French prosody discussed in Section 1.2, the final high tone that is nearly always realized in a non-final prosodic phrase is not a pitch accent. It is what we call a demarcative “phrasal” tone indicating the end of the phrase. Its variation in height does not correlate with focus of the phrase itself, but rather with the strength of the phrase boundary separating it from the following prosodic phrase. As for the initial high tone, we found an optional additional high tone on one of the non-final syllables that varied in its exact location and its presence. A long prosodic phrase is more prone to have an additional high tone than a short one, and a focused phrase is also more prone to have one. However, this additional high tone is subject to individual variation and cannot be considered as an obligatory correlate of focus, as has been sometimes proposed in the literature on French. The rich variation in prosodic cues that the French speakers use is, in our view, one of the elements of the partly contradictory views on French prosody that has been analysed in different ways by different authors.

Research question B: *The effect of type of post-verbal sequence*. Another question of interest concerned the role of phrasing of the first post-verbal object, as compared to the first post-verbal adjunct. In line with our previous studies (Féry 2014, Destruel & Féry 2019), we expected that the post-focal object would be phrased with the preceding verb ([V + Object]_φ) but the adjunct would be phrased separately from the verb ([V]_φ + [Adjunct]_φ), following the syntactic structure that shows this difference in phrasing. The

main cues for this difference lied in the height of the prosodic boundary on the preceding verb, as well as its duration. This result could be confirmed in the present experiment. Here too, an effect of type of post-verbal sequence was found in the realization of the high tone on the final syllable of the verb which was significantly higher in the case of a following adjunct than in the case of a following object, as well as in the duration of the verb, which was significantly longer in the case of a following adjunct than in the case of a following object. We interpret these results as showing that the prosodic separation between verb and adjunct is larger than between verb and object. As for the second constituent, it was consistently downstepped relative to the first one, but we could not find a difference between object and adjunct here. Both types of constituent were individually phrased in the second position. In view of these results, we can safely assume that the results of Féry (2014) and Destruel & Féry (2019) are confirmed.

Research question C: *The effect of prosodic length*. Correlates of phrasing were clearly affected by length of the prosodic constituents. Indeed, there were higher boundary tones in long constituents, more additional high tones, less downstep, less occurrences of deaccenting, and more breaks separating the two constituents, as documented in the tables of Section 3. Length has been shown to affect the correlates of phrasing in a number of other studies, see for instance Jun & Fougeron (2000), Welby (2006), Vander Klok, Wagner & Goad (2018) and Destruel & Féry (2019) for French.

Research question D: *The effect of information structure in general and dual focus in particular*. Dual focus was of particular interest in this study, as it was the first time that it was investigated in French. We could not find any clear correlate of dual focus as compared to all-focus in the statistical data for F0 and duration as compared to AF that can be considered a baseline: the F0max value of DF did not differ from AF in either of the two constituents. When considering the results of individual variation, however, focus elicited more additional high tones, more breaks and less downstep in DF than in AF. The difference is more a question of degree than an absolute one. Crucially, there is no single preferred strategy in French for the expression of focus, unlike in English or German, but rather a collection of individual correlates, all optional. Interestingly, the perception test delivered a rather high and significant percentage of correct answers for dual focus, in

contrast with all-focus and final focus for which the performance were rather poor. This indicates that the number of individual cues listed in Section 3 help listeners to interpret a dual focus correctly.

However, in comparison with the other languages for which experimental evidence exist for the realization of dual focus, our data suggests a clear difference from German and English, see Wang & Féry (2017) for German. The listeners in the latter language prove to be much more reliable at discriminating across the four focus types encoded in the (non)-congruent questions presented in the experimental task. Dual focus is often realized with a special phrasing eliciting a falling tone both on the initial and on the final focus, but in French, phrasing was stable in all information contexts. There was no change in the direction of the medial boundary tone. In all focus conditions, the first post-verbal constituent was phrased independently from the second one, and the boundary between them was always rising, except in the cases of deaccenting of the final constituent (often in the IF condition, but not exclusively).

Moreover, Eady et al. (1986), who conducted the first production experiment investigation dual focus in English, showed that, in this language, the F0max and word duration of both focused words increased to the same degree as in the corresponding single focus conditions. Furthermore, both DF and IF conditions exhibited falling contours on focused words, and this differed from the FF and the AF conditions, which had initial rising contours. The main difference between IF and DF was that post-focal F0 was significantly lowered after an initial focus. According to Eady et al. (1986), the lack of post-focal F0 lowering in dual-focus sentences represents an anticipatory influence of the additional focus at the end of the sentence. In Mandarin (Wang & Féry 2015), there was also less final compression after the first focus in dual focus than in initial focus. No such F0 lowering was found after the first focus in a dual-focus sentence in French.

Research question E: ***Individual variation***. Because of the absence of lexical stress, a clear prosodic strategy for the expression of prominence is lacking in French. There is no designated place for a pitch accent that can be increased for the sake of marking focus, like we find in Germanic languages. The final high tone found in a non-final prosodic phrase is always there—however, it is not an indicator of focus or prominence but instead of

phrasing. What we observed and documented is that, in a number of phonetic correlates, individual variation was the rule rather than the exception: Number and position of additional high tones, downstep, audible breaks and compression were all a matter of degree. The most striking element of variation is the fact that, besides high boundaries at the end of prosodic phrase, French speakers may realize additional high tones as they want, and where they want. The presence of these additional high tones is a powerful indicator of prominence, although it is far from being obligatory. We suspect a still undiscovered systematicity of the preference for specific locations of this additional high tone, some of its occurrences can be compatible with the ‘initial focus tone’ that some authors are assuming.

Even though we were unable to detect a stable prosodic indicator of focus, the fact that listeners were able above chance to indicate a correct congruency in the case of dual focus speaks in favor of acoustic cues in the signal beyond F0-max and obligatory final H-tones. However, these cues appear to be variable and optional, which relates to the less systematic congruency ratings compared to Germanic languages.

6. Conclusion

To the best of our knowledge, this article contributes the first experimental study on dual-focus for French—an information-structural condition that is largely lacking from the past literature on French prosody, and more generally, scarcely investigated cross-linguistically. Even though we cannot exclude that our results may turn out to be specific to the type of sentences tested, some aspects of French prosody were distinctly revealed. Indeed, we found that a sequence of two foci does not display the same effects as corresponding single focus on each of them. We note that French differs from the other languages investigated so far in that the language does not change the phrasing of a focused constituent and does not appear to allow two equally prominent pitch accents to co-exist in one intonation phrase, as in German. In sum, our study brings to the forefront the importance of individual variation in French, suggesting that prominence can be achieved in different ways. It is important to keep in mind that French speakers may favour non-prosodic correlates for the communication of information structure, like cleft sentences, word order and ellipsis, rendering the use of prosody marginal in some cases.

This study highlights the fact that the individual variation as to how realize focus in French may partly explain the different interpretations that are found in the literature. French has a truly different type of prosodic structure, it has no lexical stress (like German) and of course no lexical tone (like Mandarin) and thus no designated syllable in a word or in a prosodic phrase for a pitch accent that can be realized with more prominence. There is thus no obligatory marking of focus, no obligatory correlate of focus. Rather, we found a very stable syntax-based phrasing and a multitude of different cues for indicating focal prominence- We could show in a perception experiment that listeners are aware of these different cues, although they do not perform as well as speakers of a language with obligatory cues for focus.

Acknowledgments

Thanks are due to Dominik Thiele, Sebastian Bredemann, Luise Kloß and Johannes Messerschmidt for technical help and Bei Wang for precious advice and comments, and review, to Annie Rialland for her review and help for the perception experiment, as well as to an anonymous reviewer and, last but not least, to Frank Kügler the editor of this special issue who gave us very helpful comments. We also benefited from conversations with Fatima Hamlaoui and Michael Wagner.

References

- Astésano, C., Bard, E. G., & Turk, A. (2007). Structural influences on initial accent placement in French. *Language and Speech*, 50(3), 423–446.
- Beyssade, C., Marandin, J-M., & Rialland, A. (2003). Ground/Focus revisited. A perspective from French. In R. Nuñez-Cedeno, L. López & R. Cameron (Eds.), *A Romance perspective on language knowledge and use: Selected papers of LSRL 2001* (pp. 83–98). Amsterdam, The Netherlands: John Benjamins Publishing Company.
- Clech-Darbon, A., Rebuschi, G., & Rialland, A. (1999). “Are there cleft sentences in French?” In *The grammar of Focus*, ed. by Georges Rebuschi and Laurice Tuller, 83-118. Amsterdam, The Netherlands: John Benjamins Publishing Company.
- Delais-Roussarie, E. (1995). *Pour une approche parallèle de la structure prosodique: Etude de l'organisation prosodique et rythmique de la phrase française*. PhD Thesis, Université de Toulouse, Le Mirail. [In French.]
- Delais-Roussarie, E., Rialland, A., Doetjes, J., & Marandin, J-M. (2002). The Prosody of post-focus sequences in French. In B. Bel & I. Marlien (Eds.), *Proceedings of Speech Prosody 2002* (pp. 239–242). Aix en Provence, France.
- Destruel, E. (2013). An empirical investigation of the meaning and use of the French *c'est*-cleft. Ph.D. Thesis, University of Texas at Austin.
- Destruel, Emilie & Caroline Féry (2015) Compression in post-verbal sequences in French. *Proceedings of the 18th ICPHS*. Glasgow.
- Destruel, E. & Féry, C. (2019) Compression in French: Effect of length and information status on the prosody of post-verbal sequences. In Elsig, M., Feldhausen, I., Kuchenbrandt, I., & Neuhaus, M. (eds.) *Romance Languages and Linguistic Theory 15. Selected papers from 'Going Romance' Frankfurt 306*. London: John Benjamins. 157-176.
- Di Cristo, Albert (1998). Intonation in French. In: *Intonation systems: A survey of twenty languages*, ed. by Daniel Hirst and Albert Di Cristo, 195–218. Cambridge: Cambridge University Press.
- D’Imperio, M. & Michelas, A. (2010). Embedded Register levels and prosodic phrasing in French. In M. Hasegawa-Johnson (Ed.), *Proceedings of the International Conference on Speech Prosody 2010* (pp. 1–4). Chicago, USA.
- Dohen, M. & Loevenbruck, H. (2004). Pre-focal Rephrasing, Focal Enhancement and Post-focal Deaccentuation in French. In *Proceedings of the 8th International Conference on Spoken Language Processing* (pp. 785–788), Jeju Island (Korea). http://www.isca-speech.org/archive/interspeech_2004.
- Eady, S. J., Cooper, W. E., Klouda, G. V., Mueller, P. R., & Lotts, D. W. (1986). Acoustical characterization of sentential focus: Narrow vs. broad and single vs. dual focus environments. *Language and Speech*, 29, 233–250.

- Fanselow, G. (2016). Syntactic and Prosodic Reflexes of Information Structure in Germanic. In Féry, Caroline; Ishihara, Shinichiro (eds) *Oxford Handbook of Information Structure*. Oxford: Oxford University Press, 621–641.
- Féry, C. (2014). Final Compression in French as a Phrasal Phenomenon. In S. Katz Bourns & L. L. Myer (Eds.), *Perspectives on Linguistic Structure and Context: Studies in Honor of Knud Lambrecht* (pp. 133–156). Amsterdam, The Netherlands: John Benjamins Publishing Company.
- Féry, C. & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, 36, 680–703.
- Fónagy, Ivan (1979). L'accent français: accent probabilitaire. *Studia Phonetica*: 15: 123–133.
- Franich, K (this volume) Uncovering Tonal and Temporal Correlates of Phrasal Prominence in Medumba.
- German, James S., D'Imperio, M. (2016). The status of the initial rise as a marker of focus in French. *Language and Speech* 59(2): 165-195.
- Goldman, J-P. (2011). EasyAlign: An Automatic Phonetic Alignment Tool Under Praat. Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH. 3233-3236.
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge, MA: Cambridge University Press.
- Hamlaoui, F. (2009). Focus, contrast and the syntax- phonology interface: The case of French cleft sentences. In The Linguistic Society of Korea (Ed.), *Current Issues in Unity and Diversity of Languages. Collection of Papers Selected from the 18th International Congress of Linguistics (CIL18)*. Seoul, Republic of Korea: Dongam Publishing Co.
- Hirst, Daniel and Albert Di Cristo (1996). Ya-t-il des unités tonales en français?. In *Actes des XXIèmes journées d'étude sur la parole*. 223–226.
- Hyman, L. M. (2006). Word-prosodic typology. *Phonology*, 23, 225–257.
- Jackendoff, R. S. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Jun, S.-A & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), *Intonation: Analysis, modeling and technology* (pp. 209–242). Dordrecht, The Netherlands: Kluwer.
- Kabagema-Bilan, E., Lopez-Jimenez, B., & Truckenbrodt, H. (2011). Multiple focus in Mandarin Chinese. *Lingua*, 121(13), 1890–1913.
- Krifka, M. (2008). Basic Notions of Information Structure. *Acta Linguistica Hungarica* 55, 243–276.
- Kügler, F. & Féry, C. (2017). Postfocal downstep in German. *Language and Speech*, 60, 260–288.

- Ladd, D. R. (1980). *The Structure of Intonational Meaning: Evidence from English*. Bloomington, IN: Indiana University Press.
- Ladd, D. R. (2008). *Intonational phonology*. 2nd edition. Cambridge, UK: Cambridge University Press.
- Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus and the mental representations of discourse referents*. Cambridge, UK: Cambridge University Press.
- Nespor, M. & Vogel, I. (1986). *Prosodic phonology*. Dordrecht, The Netherlands: Foris.
- O'Brien, M. G. (2019). Intonation in Germanic. In: Putnam, M. & R. Page (Eds). *Cambridge Handbook of Germanic Linguistics*. Cambridge: Cambridge University Press.
- Ou, S. & Guo, Z. (this volume). The Language-specific Use of F0 Rise in Segmentation of an Artificial Language - Evidence from Listeners of Taiwanese Southern Min.
- Pasdeloup, Valerie (1990). *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*. Doctoral Dissertation, Institut de Phonétique d'Aix-en-Provence, Université de Provence.
- Portes C., Beyssade C., Michelas A., Marandin J.-M. & Champagne-Lavau, M. 2014. The dialogical dimension of intonational meaning: Evidence from French *Journal of Pragmatics* 74. 15-29.
- Post, B. (2000). *Tonal and Phrasal Structures in French Intonation*. The Hague, The Netherlands: Holland Academic Graphics.
- R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria. <https://www.R-project.org/>.
- Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1, 75–116.
- Rossi, M. 1980. Le français, langue sans accent? In: Fónagy, I., Léon, P. (Eds.). *L'accent en français contemporain*. Didier, Montréal, *Studia Phonetica* 15, 13--51.
- Schwarzschild, R. (1999). GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics*, 7, 141–177.
- Selkirk, E. O. (1995). Sentence Prosody: Intonation, Stress and Phrasing. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 550–569). Hoboken, USA: Wiley-Blackwell.
- Truckenbrodt, H. (1995). *Phonological Phrases: Their Relation to Syntax, Focus and Prominence*. PhD. Thesis. Cambridge, Mass: MIT.
- Vaissière, Jacqueline. 1980. La structuration acoustique de la phrase française. In: Pacini-Mariotti (Eds.), *Annali Della Scuola normale Superiore Di Pisa*. Tipografia, Pisa, pp. 530--560.
- Vander Klok, J., Goad, H., & Wagner, M. (2018). Prosodic focus in English vs. French: A scope account. *Glossa: A Journal of General Linguistics*, 3(1), 71. DOI: <http://doi.org/10.5334/gjgl.172>

- Wang B. & Féry, C. (2015) Dual-focus Intonation in Standard Chinese: 18th Oriental COCOSDA/CASLRE Conference during October 28-30 2015 in Shanghai Jiao Tong University, Shanghai, China, and for publication in the Conference proceedings.
- Wang, B. & Féry, C. (2017) Prosody of Dual-focus in German: Interaction between Focus and Phrasing. *Language and Speech* 61(2) 303-333.
- Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics*, 34(3), 343–371.
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics*, 27, 55–105.
- Xu, Y. (2013). ProsodyPro — A tool for large-scale systematic prosody analysis. Paper presented at the *Proceedings of Tools and Resources for the Analysis of Speech Prosody* (TRASP 2013) (pp. 7–10). Aix-en-Provence, France.

Appendix I. Reading material

Due to space restrictions, we only present here the material in condensed form, in French and the English translation, for each of the four lexicalizations.

Lexicalizations

Lexicalization 1: Jean-Marie (I') a envoyé [...] / 'Jean-Marie sent (it) ...'

short object: [...] un colis, à ma sœur / *'a package, to my sister'*

short adjunct: [...] par la poste, à Toulouse / *'via the post-office, to Toulouse'*

long object: [...] un colis important, à sa voisine anglaise / *'an important package, to his British neighbor'*

long adjunct: [...] sans vraiment le faire exprès, au fin fond de l'Argentine / *'without really doing it on purpose, to the middle of podunk Argentina'*

Lexicalization 2: Bernadette (I') a présenté [...] / 'Bernadette presented (it) ...'

short object: [...] son collègue, à leur rival / *'his colleague, to their rival'*

short adjunct: [...] dans le couloir, pendant la pause / *'in the hallway, during the break'*

long object: [...] son collègue américain, à ma belle-sœur canadienne / *'his American colleague, to my Canadian sister-in-law'*

long adjunct: [...] dans la salle de réunion, au cours du dîner d'adieu / *'in the meeting room, during the farewell party'*

Lexicalization 3: Benjamin (I') a commandé [...] / 'Benjamin ordered (it) ...'

short object: [...] des noix de macadamia, à son chef / *'some chesnuts, to his boss'*

short adjunct: [...] en automne, par hasard / *'in autumn, by chance'*

long object: [...] des noix de macadamia, à son chef de cuisine / *'some macademia nuts, to his cooking chef'*

long adjunct: [...] en plein cœur de l'hiver, par inadvertance / *'in the middle of the winter, inadvertently'*

Lexicalization 4: Ségolène (I') a caché [...] / 'Segolene hid (it) ...'

short object: [...] un trésor, à sa mère / *'a treasure, to her mother'*

short adjunct: [...] au grenier, en secret / *'in the attic, in secret'*

long object: [...] le trésor de Rackam-le-Rouge, à la famille de son gendre / *'Rackam-le-Rouge's treasure, to her son-in-law's family'*

long adjunct: [...] dans un placard abîmé, au milieu du mois d'avril / *'in a damaged closet, in the middle of the month of April'*