Anticipatory marking of (non-corrective) contrastive focus by the Initial Rise in French

Axel Barrault1, James S. German1, Pauline Welby1

1 Aix-Marseille Université, CNRS, LPL, Aix-en-Provence, France
axel.barrault@univ-amu.fr, james.german@univ-amu.fr, pauline.welby@univ-amu.fr

Abstract
This study addresses tonal marking of non-corrective contrastive focus in French. Speakers read aloud sentences composed of two parallel clauses, where the structure of the post-verbal constituent under focus was varied by the presence or absence in the second clause of the final adjective appearing in the first clause (signaling Noun-focus or NP-focus, respectively). This way, we were able to test whether the Initial Rise (i.e., LIH) is a marker of the span of an upcoming non-corrective contrast in the second clause. We posited that French speakers mark contrast tonally in the first and/or second clause.

Corroborating previous findings, we found that a faster speech rate is associated with fewer Initial Rises. More importantly, an Initial Rise occurs on the direct object of the first clause more often in narrow focus. Additionally, the height of the Initial Rise peak does not depend on focus structure. These results suggest that anticipatory use of the Initial Rise may signal an upcoming contrast and reveals additional complexity in the tonal encoding of the left edge of a contrastively focused constituent.

Index Terms: Intonation, French Initial Rise, anticipatory tonal marking, production, non-corrective contrastive focus.

1. Introduction

Different languages use different means to express contrastive focus. As compared to English, French is described as relying less on prosodic prominence to mark focus [1]. Prosodic encoding of contrastive focus has been claimed to be restricted to corrective contrast in French [2], whereas English requires a tonal marking for other types of focus, including question-answer congruence and non-corrective contrastive focus [1]. Indeed, most studies on French prosodic encoding of focus investigated corrective focus, often using the term contrastive to refer to correction as in [3] for instance. Previous findings, however, have established that French includes mechanisms for marking non-corrective focus prosodically [4].

At the same time, there is a wider range of factors that condition the distribution of precisely those prosodic prominences that are associated with focus marking in an Accentual Phrase (henceforth, AP). Among them, the Initial Rise (i.e., LIH) which occurs at the left edge of an AP [3,5], post-focal deaccenting [7], and the strengthening of prosodic boundaries [7] are all potential strategies for marking focus in French. However, each of them is sensitive to a wide variety of factors originating from different levels of representation. With respect to the Initial Rise, speaking rate [3,5,8], speaking style [9], syntactic constituency [10], phrase length [8] and segmental composition [8,9] are all known to have an influence on its distribution. For non-corrective contrastive focus, it has been claimed that French relies principally on specific syntactic structures, such as clefting of the constituent under focus [11]. Of course, these syntactic structures are always accompanied by specific tonal marking, but in such cases, it is difficult to differentiate the contribution of tonal structure from that of syntactic processes in the perceptual identification of contrastive focus.

Another syntactic device, which we henceforth refer to as parallelism, naturally involves contrastive focus based on the salience of similarities and differences in the clauses involved. In other words, part of the linguistic material is repeated while one constituent changes. Parallelism is interesting because it allows a more direct comparison between English and French in the prosodic encoding of focus, and more generally between Germanic and Romance languages [2,12]. Contrastive focus has been treated by some as an anaphoric process [13,14]. Under this type of account, therefore, the first instantiation of a contrasted pair of elements, to the extent that it is marked, can be viewed as a type of cataphora. Henceforth, we refer to this phenomenon as an anticipatory focus marking.

In this study, we used parallelism to study whether the Initial Rise contributes differently to the anticipatory marking of the first (i.e., anticipatory) element of a contrastive pair as a function of the span of the focus. We build on a production experiment [6] that first established quantitatively a probabilistic association between the Initial Rise and the left edge of a non-corrective contrastive focus.

For the present study, our working hypothesis is that French speakers can set up an upcoming prosodically encoded contrast, perhaps as part of a strategy help listeners anticipate a parallelism and facilitate lexical access. Our experimental hypothesis is that this prosodic encoding involves an anticipatory marking by an Initial Rise which should occur more often when the focused constituent is a Noun (i.e., a “narrower” focus constituent) as compared to when the focus constituent is a Noun Phrase (i.e., a “wider” focus constituent). Furthermore, we expect that, independently of focus structure, a faster speech rate will be associated with fewer Initial Rise [8]. Finally, since the Initial Rise is expected to be produced in both conditions – although more in the Noun focus condition than in the Noun Phrase condition – speakers may also rely on a higher Hi peak as a means of enhancing the marking of a narrower focus constituent. Therefore, we hypothesized that Hi peaks may be higher when the focused constituent is a Noun versus a Noun Phrase. To assess these hypotheses, we carried out a production experiment in which the span of the contrastive constituent in a first clause was manipulated through the type of parallelism induced by the structure and content of the clause that followed it.
2. Methods

2.1. Participants

We recorded 43 French native speakers (34 women, 9 men) who did not declare any language impairment. Their age ranged from 18 to 45 years old (mean = 23.7; SD = 7.7). All resided in France at the time of the study and received a paid compensation.

2.2. Materials

2.2.1. Stimuli

56 target sentences were created for the study. These involved two parallel clauses which were conjoined by the temporal expression et ensuite (’and then’), and in which a substantial part of the lexical material of the first clause was repeated in the second clause. The structure of these sentences thus varied in terms of the presence or absence in the second clause of a final adjective modifying the post-verbal noun phrase (i.e., direct object), appearing in the first clause, as illustrated in (2).

(2) Elle découpe le mirador bleu et ensuite, elle découpe [...] ‘She cuts the blue watchtower and then, she cuts [...]’

a. Noun focus condition

[...] le monument bleu. ‘the blue monument’

b. Noun Phrase focus condition

[...] le monument. ‘the monument’

The target sentences were matched, such that half (n=28) were designed to induce a contrastive focus that spans the two head nouns (in (2), mirador vs. monument) within the specified color set (bleu). The other half were also designed to induce a narrow contrast, though one which spans the entire Noun Phrase in each clause (i.e., mirador bleu vs monument).

In order to minimize fo perturbations, the 56 pairs of nouns used as targets were selected to include as few obstructant segments as possible. Using lexical information from Lexique 2 [15], preference was given to frequent and imageable words. Variability in sentence content across items was varied by means of five pronouns, ten transitive verbs, and eight color adjectives that vary in gender and number according to their head noun. An additional 24 “filler” sentences served both as distractors and as a baseline manipulation, the results of which are not reported here. Half of these involved parallelism and a narrow corrective focus. The other half were designed to induce only fully broad (i.e., clause-level) focus.

2.2.2. Procedure

Participants were recorded in an anechoic chamber at the Laboratoire Parole et Langage. During the task, they sat facing a laptop, and stimulus presentation was implemented the experiment builder OpenSesame [16]. Participants were told that they would see sentences appearing one by one and that each sentence would appear twice in a row: first with a yellow “thought bubble”, and then with a green microphone icon. When they saw the thought bubble, participants were asked to read the sentence silently. When they saw the green microphone, they were asked to read it out loud. Sentence presentation was self-paced. A two-second interval between slides was enforced in order to prevent skipping and to ensure they carried out the silent reading. The task lasted approximately 20 minutes.

2.2.3. Annotation

Prosodic annotation was carried out in two phases using Praat [17]. In the first phase, the first clause of each trial was annotated for (i) tonal targets including AP-initial rises (LHi) and AP-final contours (e.g., LHi), (ii) phrase boundaries (AP, intermediate phrase (ip), intonation phrase (IP)), and (iii) miscellaneous events such as pauses, hesitations, speech errors, etc. Crucially, the experimenter was prevented from hearing or viewing the second clause and was therefore blind to the experimental condition (i.e., Noun vs. Noun Phrase focus). In the second phase, the second clause was annotated following the same procedure as before. In a third step, all tokens were automatically aligned at the word- and phone-level using WebMAUS [18]. Finally, a script was used to automatically label the syntactic status of each word (i.e., pronoun, verb, determiner, noun or adjective) on a separate tier.

Annotations were carried out by the first author who is a native speaker of French. The presence or absence of LHi in the critical region (i.e., the final AP of the first clause) was determined as follows: First, auditory inspection was used to decide whether a perceivable prominence was present. This was followed by visual inspection to corroborate the auditory percept and to decide uncertain cases. For this step, a threshold of a 10 Hz rise was used following [10].

3. Results

3.1. Prosodic analyses

3.1.1. Descriptive statistics

A total of 2408 (56x43) target tokens were collected. Since the present paper concerns primarily anticipatory focus marking, only the results from the analysis of the first clause are presented here.

Overall, four different phrasing patterns for the first clause were observed. As expected, the most frequent phrasing consisted of two adjacent APs forming a single ip, such that the second AP encompassed the critical Noun Phrase region.

In order to compare the proportion of Initial Rise occurrences as a function of the focus span on the critical AP of the first clause, it is necessary to control for the phrasing level. Therefore, the following analyses focus on the subset of the data produced using the most frequently observed phrasing pattern. Table 2 shows that 94.8% of the data produced with the most frequent phrasing are quite evenly distributed across conditions.

Table 1: First clause phrasing patterns

<table>
<thead>
<tr>
<th>Phrasing</th>
<th>%</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>((Pro Verb)AP (Det Noun Adjective)AP)AP</td>
<td>94.8</td>
<td>2283</td>
</tr>
<tr>
<td>((Pro Verb)AP (Det Noun)AP (Adjective)AP)</td>
<td>4.4</td>
<td>106</td>
</tr>
<tr>
<td>(((Pro Verb)AP (Det Noun Adjective)AP)ip</td>
<td>0.7</td>
<td>17</td>
</tr>
<tr>
<td>(Pro Verb)AP (Det Noun Adjective)AP</td>
<td>0.1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>2408</td>
</tr>
</tbody>
</table>

Table 2: Main phrasing across conditions

<table>
<thead>
<tr>
<th>Focus span condition</th>
<th>%</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>49.9</td>
<td>1139</td>
</tr>
<tr>
<td>Noun Phrase</td>
<td>50.1</td>
<td>1144</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>2283</td>
</tr>
</tbody>
</table>
As shown in Figure 1, utterances bearing either LLH*H- or LHiLH*H- on the ip-final AP represent 98.2% (n=2240) of the utterances produced with the most frequent phrasing (n=2283), whereas other intonational patterns combined represent only 1.9% (n=43).

Figure 1: Proportion of tonal patterns across focus conditions

Indeed, we have no a priori predictions regarding how these other tonal patterns relate to the span of narrow focus marking, and in any case, their very low proportional representation would not be clearly interpretable in a statistical model. We therefore excluded these patterns from subsequent analyses. As shown in Table 3, these remaining 2240 utterances were quite evenly distributed across conditions. To characterize how the Initial Rise relates to focus span, we examined the proportion of LHiLH*H- as compared to that of LLH*H- on the ip-final AP of the first clause as a function of focus span. As shown in Table 4, LHiLH*H- is very frequent regardless of the focus span.

Table 3: Tonal patterns distribution across condition

<table>
<thead>
<tr>
<th>Tonal patterns</th>
<th>% Noun condition</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>LHiLH*H-</td>
<td>52.3</td>
<td>1619</td>
</tr>
<tr>
<td>LLH*H-</td>
<td>44.9</td>
<td>621</td>
</tr>
<tr>
<td>Total</td>
<td>50.2</td>
<td>2240</td>
</tr>
</tbody>
</table>

Table 4: Distribution of LHiLH*H- in the first clause.

<table>
<thead>
<tr>
<th>Focus span condition</th>
<th>LHiLH*H-</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>75.2</td>
<td>1125</td>
</tr>
<tr>
<td>Noun Phrase</td>
<td>69.3</td>
<td>1115</td>
</tr>
<tr>
<td>Total</td>
<td>72.3</td>
<td>2240</td>
</tr>
</tbody>
</table>

3.1.2. Statistical model

In order to statistically assess whether the choice of LHiLH*H- versus LLH*H- depends on the span of focus, a generalized linear mixed model was fit to the relevant subset of the collected data presented above using the `glmer` function (lme4 package version 1.1-23) in R [19], with the intonational patterns (i.e., LHiLH*H- vs. LLH*H-) as the binary dependent variable. Focus span (Noun vs Noun Phrase), along with a variable splitting the task in two equal parts (i.e., first vs. second half of the task) were included as fixed factors, as well as their interaction. This second fixed factor was included in order to account for a change of production behavior during the task (e.g., due to test-retest effect). The maximal converging random effect structure included random intercepts for participants and item pairs. Model comparisons using likelihood ratio test with stepwise removal were used to assess which factors contributed significantly to the model fit. The optimal fitted model showed a main effect of focus span condition on the probability of observing LHiLH*H- (β= -0.4289, z= -3.745, p<0.0001) and a main effect of the variable split-half (β= 0.2683, z= 2.260, p=0.023843) showing more LHiLH*H- in the second half of the task; with no significant interaction (χ²=0.0847, p=0.771). The model reflected a higher variability between participants (σ² = 2.7045) as compared to between items pairs (σ² = 0.5609), which shows a relatively good homogeneity between item pairs.

3.2. Exploratory analyses

3.2.1. Speech rate analysis

According to previous findings, a faster speech rate is associated with fewer Initial Rises [3,5,8]. To test whether the present data reflect this effect we measured speech rate within the critical region used in the above analyses (i.e., the ip-final AP of first clause). Speech rate was computed as the number of syllables (always five), divided by the duration of the critical window. About 8.5% (n=191) of utterances were removed because of disfluencies that would artificially bias the speech rate calculation.

Overall, the speech rate ranged from 2.21 to 7.65 syllables per second (mean = 5.09; SD = 0.79). To assess statistically whether a higher speech rate decreases the probability of Initial Rise occurrence, a generalized linear mixed model was fit to this subset of the data using the `glmer` function (lme4 package version 1.1-23) in R [19], with the presence (i.e., LHiLH*H-) or absence (i.e., LLH*H-) of the Initial Rise as the binary dependent variable. Speech rate, focus span (Noun vs Noun Phrase), and their interaction were included as fixed factors. The maximal converging random effect structure included random intercepts for participants and item pairs. Model comparisons using likelihood ratio test with stepwise removal were used to assess which factors contributed significantly to the model fit. The optimal fitted model showed a main effect of the speech rate (β= -1.4518, z= -11.247, p<0.0001) and a main effect of the condition (β= 0.4905, z= 3.902, p<0.0001) on the relationship between the speech rate and the presence of the Initial Rise on the target AP; with no significant interaction (χ²=0.0072, p=0.9323). Again, the fitted model reflected a higher variability between participants (σ² = 2.4330) as compared to between item pairs (σ² = 0.5387). In line with the literature [3,5,8], we found that a faster speech rate is associated with fewer Initial Rises on the target AP.

3.2.2. Phonetic analysis of Hi peaks

Considering the results outlined in §3.1.2, additional analyses were conducted to investigate whether focus span influences the phonetic implementation of the Initial Rise. Specifically, we tested whether Hi peaks of the target AP were higher in the Noun focus than in the Noun Phrase focus condition. Specifically, we assessed whether mean f0 differed across the two conditions.

Of the bearing LHiLH*H- on the target AP, the f0 values of five Hi peaks were removed because they were produced with a creaky voice could not be reliably measured. Contrary to our hypothesis, mean f0 for both conditions was almost identical (Table 6).
Table 6: Average f0 values of Hi peaks per condition.

<table>
<thead>
<tr>
<th>Focus span condition</th>
<th>n</th>
<th>mean f0 value (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>843</td>
<td>216</td>
</tr>
<tr>
<td>Noun Phrase</td>
<td>771</td>
<td>217</td>
</tr>
<tr>
<td>Total</td>
<td>1614</td>
<td>217</td>
</tr>
</tbody>
</table>

Additionally, to explore how the scaling of Hi peaks relates to the focus span, we looked at their scaling relative to the upcoming H* (i.e., H*H-). From the previous subset (n=1614), 44 additional utterances were excluded because of H* produced with a creaky voice. As shown in Table 7, the proportion of Hi peaks scaled higher than their H*H- were evenly distributed across condition.

Table 7: Scaling of Hi as compared to H*

<table>
<thead>
<tr>
<th>Scaling</th>
<th>% Hi&lt;H*H-</th>
<th>Total n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>22</td>
<td>821</td>
</tr>
<tr>
<td>Noun Phrase</td>
<td>23</td>
<td>749</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>1570</td>
</tr>
</tbody>
</table>

4. Discussion

Overall, the results of the present study support our main experimental hypothesis. Namely, the Initial Rise (i.e., LH) occurs more often when the focused constituent is a Noun (e.g., (2a)) as compared to when it is a Noun Phrase (e.g., (2b)). In other words, in a parallelism structure, the Initial Rise appears to provide anticipatory marking of the first instantiation of a contrastively focus constituent as a function of the focus span. Furthermore, speakers tended to produce fewer Initial Rises as the speech rate increased but with no interaction with the focus conditions. This result corroborates previous findings for French, which suggest that this effect is rather robust across different experimental tasks. However, the exploration of the phonetic implementation did not reveal any difference with respect to the focus condition in this study contrary to our exploratory hypothesis. Descriptively, neither the scaling of Hi peaks relative to the following H*, nor the height of Hi peaks varied as a function of focus conditions. This suggests that speakers did not rely on continuous means to mark focus span in this task.

The main result is also compatible with an account where the use of prosodic prominence is restricted to corrective statements in French [1]. In this account, the required focus antecedent would come from a prior utterance that stands in a corrective discourse relation to the one being uttered [20]. In this reasoning, the present result could be interpreted as reflecting the tendency for speakers to reconstruct a prior corrective context as a way to make sense of the sentence and thereby produce the test materials in a felicitous manner. Given that a third of filler trials explicitly involved corrective foci (e.g., she did not start the tree, she washed the tree) at different syntactic positions (i.e., subject, verb or post-verbal adjective), these may have served as appropriate primes for such corrective meaning. This view, however, raises the question of why participants would use corrective focus more in the Noun focus cases than in the Noun Phrase focus cases. Further studies will be needed to address this question.

In line with previous work, the observed effect of the focus span on the occurrence of the Initial Rise should not be interpreted as suggesting that the IR is a marker of focus in the traditional sense of a one-to-one mapping between intonational pattern and discourse function, especially given that speech rate has a stronger impact on the Initial Rise distribution. Nevertheless, the present finding is statistically robust and corroborates the characterization of the association between the Initial Rise (IR) and contrastive focus as “weak” [6].

Furthermore, the main effect size of the focus span on the probability of Initial Rise may underestimate the daily use of anticipatory marking by the Initial Rise for several reasons. First, producing anticipatory marking requires planning the entire utterance in this case. Given the rather high cognitive load of the material, we could reasonably expect a stronger effect in a more interactive task. Second, we cannot exclude the possibility for participants to have reconstructed a context in which by uttering (2b) (i.e., without the color in the second clause), they implicitly intended (2.a) and eluded the final color adjective. This is felicitous in a context where there are several miradors (‘watchtowers’), but only one monument. Indeed, in this context, the restrictive modifier bleu (‘blue’) is required for mirador in order not to violate the maxim of quantity but is redundant if there is no other monument than the blue monument in the alternative space, that they constructed to felicitously produce (2b). If this interpretation like (2b) was ever made – under our hypothesis that the Initial Rise is preferably associated with the narrower focus, then this can only result in more Initial Rise in the broad focus condition than it should. It would dilute the effect. Thus, since we still observe an effect of the focus condition on the proportion of the Initial Rise demonstrating a stronger association of the intonational marking with the Noun focus condition, it suggests that the effect can only be stronger.

As with previous findings, the present results suggest how much the modeling of the relationship between the French Initial Rise and contrastive focus first requires considering a thorough study of the several sources of variability as probabilistic inputs. More studies at the production level are needed to shed light on the weight of all contributing factors influencing the occurrence of this anticipatory marking of the narrow focus span. Nevertheless, the hypothesis according to which French listeners can take advantage of using of the Initial Rise combined with other parameters to identify the left edge of a contrastive focus remains a valid hypothesis at this step and will be the topic of future studies. The present finding establishes that French speakers prosodically encode an upcoming contrast and do so as a function of the focus span.

5. Acknowledgements

This work, carried out within the Institut Convergence ILCB (ANR-16-CONV-0002), has benefited from support from the French government, managed by the French National Agency for Research (ANR) and the Excellence Initiative of Aix-Marseille University (A*MIDEX).

6. References


