



ARTICLE

The perception of post-focal prominence is (also) a matter of linguistic structure

Martine Grice^{1,*} , Michelina Savino², Maria Lialiou¹, Aviad Albert¹ and Petra Schumacher³

¹IfL-Phonetics, University of Cologne, Cologne, Germany, ²Department of Education, Psychology, Communication, University of Bari, Bari, Italy, and ³Department of German Language and Literature I, Linguistics, University of Cologne, Cologne, Germany

*Corresponding author. Email: martine.grice@uni-koeln.de

(Received 13 June 2024; revised 27 March 2025; accepted 9 April 2025)

Abstract

This paper reports on a web-based experiment to investigate the perception of prominence in words with different focus structures in Italian. In the experiment, native listeners of the Bari variety of Italian, and German learners of Italian, rated the perceived prominence of object nouns in broad and narrow focus and post-focally using a visual analogue scale. Although both groups of listeners rated words in narrow focus as higher in prominence than words in broad focus, there were differences between the two groups in rating post-focal words. While German learners rated post-focal words as less prominent than those in broad focus, native Italian listeners perceived words in both of these conditions as equally prominent. The Italian ratings are particularly striking, as post-focal words had flat pitch and were weaker in terms of periodic energy mass than words in broad focus, leading to the conclusion that the native listeners were rating the words by taking their knowledge of the prosodic system of Italian into account. Our results confirm phonological accounts of Italian as having post-focal accents, even when the pitch is flat.

Keywords: deaccentuation; post-focal accent; Italian; German; intonation

1. Introduction

In many languages, speakers employ prosody to highlight focussed information or to attenuate information that is out of focus, in particular when this information follows the focus. In Italian, as in West Germanic languages such as English and German, highlighting is achieved by means of accentuation, which is reflected in acoustic parameters such as a dynamic f_0 contour, local enhancement of spectral properties, and increases in periodic energy and duration. Conversely, attenuation can be achieved with deaccentuation, with a concomitant reduction in these parameters. However, Italian has been reported to have accents in post-focal position (Cruttenden 1993; Ladd 2008). It is this position that we concentrate on in the current study. In particular, we are concerned with how listeners rate the prominence of words that follow the focus of a sentence and how far they are guided by the linguistic system of their native language. To tap into the factors that affect native speaker perception of prominence, we collected ratings from native Italian listeners and compared

them with ratings of the same utterances by learners of Italian with German as their first language.

1.1 Post-focal prominence in Italian and German

Ladd (2008) describes Italian as strongly resistant to ‘contextual deaccenting’, the context being repeated words or phrases, or more generally given materials (Ladd 2008: 231). Experimental confirmation of this tendency in noun phrases (NPs) is provided by the study by Swerts, Krahmer & Avesani (2002), who found that in noun–adjective pairs (*triangolo rosso*, lit. triangle red), Tuscan Italian native speakers tended to produce a pitch accent on the adjective (*rosso*), i.e. the rightmost element in the phrase, regardless of whether the noun or the adjective was in focus. The resulting conclusion – that Italian does not mark focus structure in such NPs – has been called into question by Sbranna *et al.* (2023), who have shown that prosodic marking of post-focal given material is realised through the modulation of phonetic parameters in the f_0 trajectory and the amount of periodic energy (see section 1.2 below), rather than via a categorical presence versus absence of a pitch accent. Thus, Italian speakers do mark information status, but, crucially, they do not do it through deaccentuation.

Despite the ascribed general tendency for Italian to block deaccenting, several studies have shown that in sentence-length utterances, typically in narrow and/or contrastive focus declaratives, post-focal constituents can, in fact, be highly attenuated (this is also reported in Ladd (2008: 233) as ‘metalinguistic corrections’). In this case, post-focal elements are characterised by a low flat f_0 trace, yet they are still phonologically described as pitch accents, either low (L^*) or low-falling ($H+L^*$) in a greatly compressed pitch range (Grice *et al.*, 2005; Cangemi & D’Imperio, 2013; Gili Fivela *et al.*, 2015, among others), as they reportedly bear some prominence cued by aspects of the signal other than f_0 , namely duration, formant trajectories and spectral emphasis (Bocci & Avesani 2011, 2015).

In contrast to Italian, German has been described as a language that deaccents given materials (Ladd 2008), where deaccenting is acoustically marked by a low flat f_0 (Féry & Kügler 2008). Interestingly, Kügler and Féry (2017) detected a subtle f_0 movement on post-focal material, which they accounted for as a pitch accent in a greatly compressed pitch range. In this respect, their analysis is similar to that of post-focal accents in Italian declaratives. Moreover, Kügler (2008) found that words in post-focal position are longer than the same word in broad focus (all new), where the investigated word is the focus exponent. These results seem to suggest that, despite the fact that Italian and German are generally classified as having different post-focal structures, material in this position appears to have a similar degree of acoustic prominence in both languages. We return to this issue in the Discussion and Conclusion sections.

Although we are concerned with statements in our study, it is worth looking at the intonation of questions, as, in Italian, post-focal material can be acoustically more salient in questions than in statements. In fact, in polar questions in Bari Italian, a post-focal lexical item bears a clearly identifiable rising-falling pitch (analysed as $L+H^*$, followed by a low edge tone; Grice & Savino 2003; Grice *et al.* 2005; Savino 2012; Gili Fivela *et al.* 2015). This post-focal accent is linguistically important, as it can be used to distinguish between statements and questions, a feature shared with other (but not all) varieties of Italian (see Grice 1995; Grice *et al.* 2005; Savino 2012; Cangemi & D’Imperio 2013; Gili Fivela *et al.* 2015, among others). This feature in Bari Italian is exemplified in Figure 1 (a–b), showing the f_0 contour of the sentence *Mamma è andata a ballare da Lalla* (‘Mum went to dance at Lalla’s’) as a contrastive narrow focus declarative in (a), and as a narrow focus polar question in (b), with *ballare* as the focussed constituent in both cases (audio files and annotations taken from

Grice et al. 2005, and re-visualised with Praat (Boersma & Weenink 2018)). Note that in the narrow focus declarative, the post-focal accent on *Lalla* is characterised by a flat f_0 (here, the downstepped ‘!H+L*’ label is meant by the authors to indicate the post-focal H+L* within a very compressed pitch range), whereas in the narrow focus question, the post-focal accent on *Lalla* has a rise-fall f_0 shape, similar to the focal accent on *ballare*, (analysed as L+!H*’). This post-focal rise-fall is crucial for signalling interrogativity.

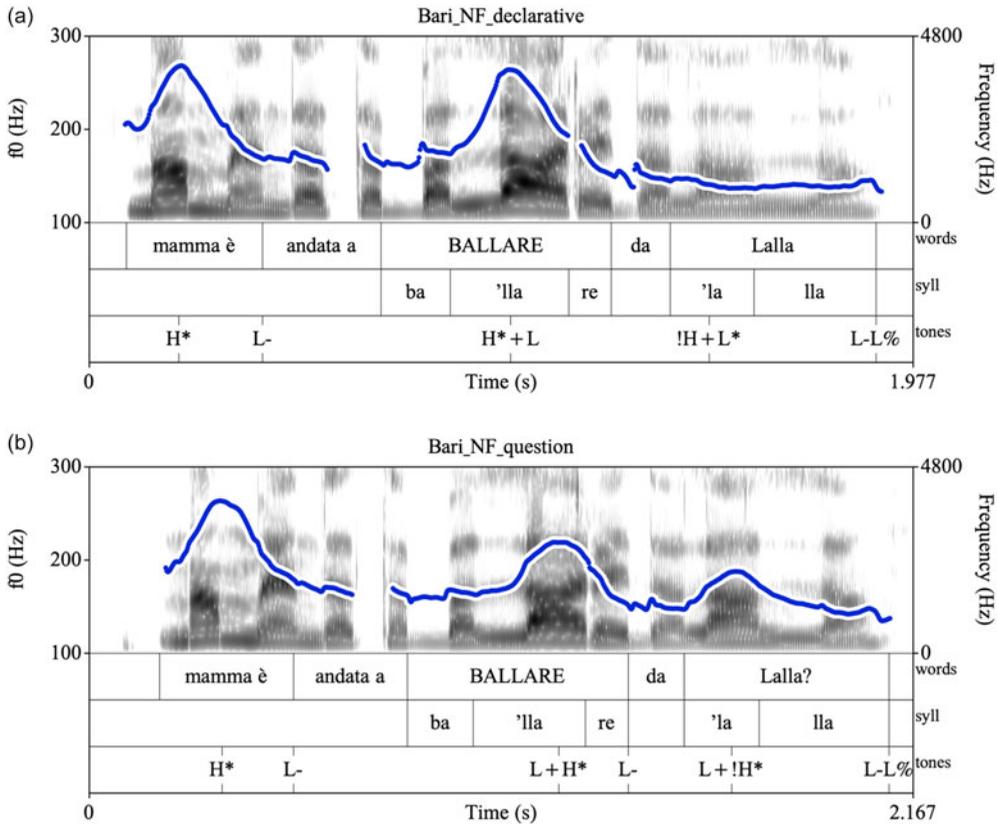


Figure 1 (a–b). Two renditions of the sentence *Mamma è andata a ballare da Lalla* ('Mum went to dance at Lalla's'), (a) as a statement, and (b) as a polar question, in both cases with narrow focus on *ballare* (from Grice et al. 2005).

1.2 Acoustic cues to post-focal prominence in Italian and German

In the previous section, we reported that, in the absence of f_0 movements, other acoustic features can cue post-focal prominence, primarily duration (Bocci & Avesani 2011, 2015 for Italian; Kügler 2008 for German). However, it has been recently shown that periodic energy mass (henceforth 'Mass'), a measure selectively reflecting the intensity of pitch-bearing components in the acoustic signal, is a better cue to post-focal prominence than duration, as it takes both duration and intensity into account, employing a single scale for measuring prosodic strength (Albert et al. 2022; Albert 2023).

Evidence for the role of Mass as a reliable acoustic cue to prominence has been provided for both Italian and German. In the above-mentioned study by Sbranna et al. (2023) on prosodic marking of information structure in Neapolitan Italian, it was found that in

two-word NPs, under different information structure conditions, the stressed syllable of a given element in final position was produced with reduced Mass in comparison to that of a new element. A comparable trend has also been reported for German by Lorenzen & Baumann (2024), who found that Mass is a consistent cue to lexical and referential givenness, with higher Mass on new information than on given information.

1.3 Effects of linguistic structure on the perception of prominence

Experimental evidence has shown that what listeners perceive in the acoustic signal is not exclusively ‘signal-driven’. Instead, it could be partially influenced by their knowledge of the utterance’s linguistic structure (Eriksson, Thunberg, & Traunmüller 2001; Cole, Mo, & Hasegawa-Johnson 2010; Bishop 2012, 2016; Calhoun, Wollum, & Kruse Va’ai 2021, among others). For example, Bishop (2012) found that prominence ratings of objects in SVO constructions by American English listeners were influenced by their language-based expectations of prominence patterns, and Calhoun *et al.* (2021) found effects of native linguistic knowledge on perceptual rating of prominence in New Zealand English and Samoan.

Notably, these ‘expectation-driven’ effects in Bari Italian have been recently investigated in an experiment based on Event Related Potentials (ERPs) by Ventura *et al.* (2020), who tested to what extent attentional resources are allocated to prosodically very prominent words marked by a narrow focal accent, compared to words in post-focal position in statements and polar questions. They found that attentional resources were directed not only to words in focal position, but also to words in post-focal position; this was attributed to the linguistic structure of the language, according to which prominent information is usually placed in final position in an utterance. Interestingly, despite the fact that post-focal target words are acoustically more prominent in questions than in statements (having a rise–fall f_0 movement in questions as opposed to flat f_0 in statements, as shown in Figure 1 above), attentional resources were reduced in the former context, since the rise–fall consumed attentional resources for processing sentence modality information, *i.e.* questioning, again as predicted by Bari Italian linguistic structure.

2. Motivation of this study and hypotheses

We have discussed that in both Italian and German statements, though post-focal prominence is not cued by f_0 (post-focal elements having low flat f_0), other acoustic cues appear to be available, such as duration and, more recently, Mass. Moreover, since the perception of prominence is not only ‘signal-driven’ but also ‘expectation-driven’, we ask whether the structure of listeners’ native language affects the perception of post-focal prominence, and how far Italian listeners take f_0 and Mass into account in their prominence judgements.

To answer this question, we built a set of stimuli in Bari Italian with target words in statements produced under different focus conditions – (contrastive) narrow focus (NF), broad focus (BF) and post-focal (PF) – and we instructed two groups of listeners to rate their prominence, namely: (1) a group of Bari Italian native speakers, and (2) a group of German learners of Italian.

If prominence judgements by native listeners of Italian are based not only on the acoustic signal – bottom-up processing – but also on their knowledge of the language’s phonology and their expectations as a result of that knowledge – top-down processing, we might expect them to rate post-focal words differently from German learners. For the learners,

the signal is expected to play a more important role than expectations resulting from a knowledge of the phonological distribution of accents in the language as a whole, and more importantly, of this variety in particular.

There are a number of reasons why learners of Italian might have less knowledge of the phonological distribution of accents in the language. First, our experiment uses a regional variety of Italian that is unlikely to be taken as a model for learners. Learners will be familiar with the lexical items and their phonological makeup (the consonants, vowels and lexical stress patterns) as well as the intonation pattern used in declaratives, which is strikingly similar across varieties (Gili Fivela et al. 2015; Cangemi & Grice 2016). However, they will be unfamiliar with the intonation pattern used in polar interrogatives in the Bari variety and are unlikely to have been exposed to acoustically salient rising-falling post-focal accents in such questions. The mere presence of this functional distinction post-focally may affect the prominence ratings of native speakers as compared to the learners, who are likely to rely more on the signal.

3. Materials and methods

3.1 Participants

Two groups of listeners took part in the experiment: 16 native speakers of Bari Italian (13 F, 3 M; mean age: 26.47) and 18 German intermediate and upper-intermediate learners of Italian (14 F, 4 M; mean age: 24.72) who were not familiar with the variety spoken in Bari. Participants were all either university students or recent graduates. None of them had a background in phonetics or prosody. They participated in the experiment on a voluntary basis and gave written informed consent.

3.2 Stimuli

Stimuli consisted of 30 utterances (10 utterances \times 3 conditions), all declaratives, elicited in a reading task as answers to questions. Stimuli were selected from recorded declarative sets produced by five different (female) native speakers of the Bari variety of Italian, all students at the local university without any background knowledge in prosody. In recording sessions, speakers were instructed to produce each target sentence as a response to a question read aloud by a confederate, with the text of each question/answer pair presented on a computer screen. Speakers were encouraged to read in a natural way, and to repeat any sentence if they were not satisfied with their production (in these cases, only the second rendition was considered as a potential stimulus).

One speaker's set of utterances was selected based on fluency and naturalness, as judged by two Italian phoneticians. Fluency and naturalness were also taken into account when selecting 10 from 20 candidate sentences produced by that speaker for each condition. The preceding context consisted of questions, which elicited three different focal conditions in which the target word occurs as broad focus (henceforth BF), contrastive narrow focus (henceforth NF) or post-focally, i.e. after a contrastive narrow focus (henceforth PF). An example of each of the conditions preceded by context sentences are provided in examples (1)–(3), where the target word is indicated in bold and underlined, and the contrastive narrow focus condition is indicated with capital letters. Note that in (3) the verb preceding the target word is capitalised, signalling its contrastive narrow focus status, triggering a post-focal realisation of the target word.

(1) TARGET in broad focus

Cosa bisogna fare per entrare nella stanza?

(‘What should one do to enter the room?’)

[Bisogna girare_{verb} la maniglia_{TARGET} sulla porta_{noun in PP}]_{FOCUS}
 V.3.S INF DET N PREP-DET N
 Need turn the handle on the door

‘One needs to turn_{verb} the handle_{TARGET} on the door_{noun in PP}’

(2) TARGET in contrastive narrow focus

Per entrare nella stanza bisogna girare la maniglia o il chiavistello sulla porta?

(‘To enter the room should one turn the handle or the latch on the door?’)

Bisogna girare_{verb} [LA MANIGLIA_{TARGET}]_{FOCUS} sulla porta_{noun in PP}

(3) TARGET in post-focal position

Per entrare nella stanza bisogna girare o spingere la maniglia sulla porta?

(‘To enter the room should one turn or push the handle on the door?’)

Bisogna [GIRARE_{verb}]_{FOCUS} la maniglia_{TARGET} sulla porta_{noun in PP}

The target word was always a noun (e.g. *la maniglia*, ‘the handle’ in the example above), preceded by a verb (e.g. *girare*, ‘to turn’) and followed by a prepositional phrase (e.g. *sulla porta*, ‘on the door’). Target nouns were always trisyllabic words with penultimate stress, verbs were trisyllabic with lexical stress in either penultimate or antepenultimate position, whereas nouns in the prepositional phrase could be either trisyllabic words with lexical stress in either penultimate or antepenultimate position or disyllabic words with penultimate stress. The text of all target sentences and context questions is provided in Appendix A. The audio files of all stimuli are available at <https://osf.io/6kjsx/>.

The selected stimuli were intonationally analysed by two trained transcribers using the autosegmental-metrical model for Italian described in Grice et al. (2005) using Praat (Boersma & Weenink 2018). The target words were realised with a H+L* falling pitch accent in the BF condition, and with a H*+L (rising-)falling pitch accent in the NF condition, whereas in the PF conditions they were realised with a flat f0, analysed as L*. The verb preceding the object noun was realised with a falling H+L* pitch accent in BF, with either a falling H+L* or a peak H* pitch accent in NF, and with a H*+L (rising-)falling accent in PF, i.e. when the verb was marked by contrastive narrow focus.

The use of these pitch accents in BF and NF declaratives has been largely attested in the literature for most of the Italian varieties, including the one spoken in Bari (Grice & Savino 1997, 2003, 2004; Grice et al. 2005; Gili Fivela et al. 2015; Grice, Savino, & Roettger 2019, amongst others).

Finally, the noun in PP was realised with a low-falling compressed pitch in BF and with a flat pitch in NF and PF, i.e. when occurring in post-focal position.

An example of the intonational realisations of the stimuli across the three prosodic conditions is offered in Figure 2(a–c).

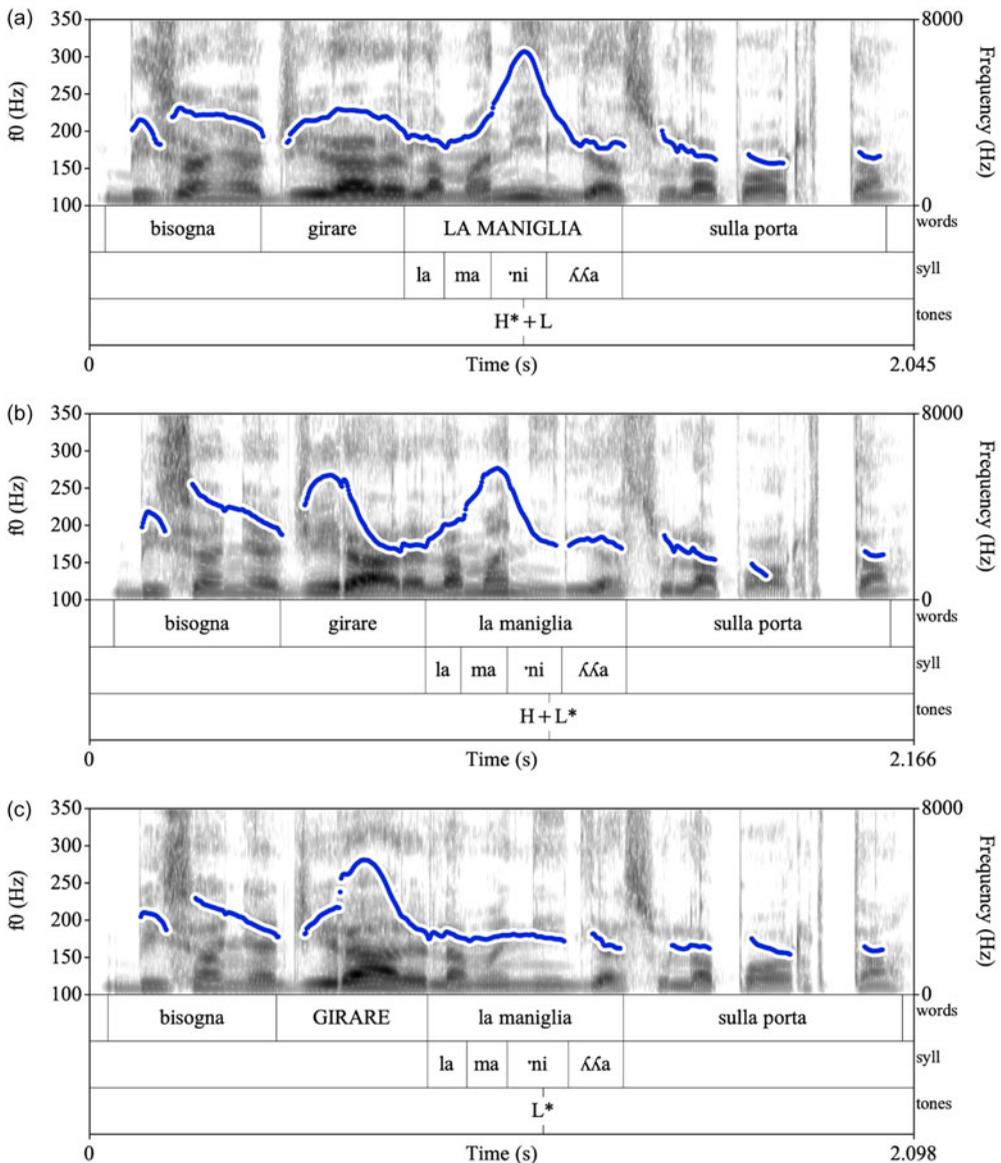


Figure 2 (a–c). F0 contours of examples with the target phrase *Bisogna girare la maniglia sulla porta* ('One needs to turn the handle on the door') where *la maniglia* ('the handle') is the target word, in (a) narrow focus, (b) broad focus and (c) post-focal positions.

Note that, given the position of the target before a prepositional phrase, the mapping between syntactic structure and prosodic structure generally assigns an intermediate phrase boundary after the target (see Selkirk, 1984, 1995; Nespor & Vogel 1986; Dehé & Samek-Lodovici, 2009). In our stimuli, such an intermediate phrase boundary was always realised as an L tone in both NF and BF conditions. In the PF condition, we could not directly observe cues for an intermediate phrase boundary but assumed the analysis of Bocci and Avesani (2011, 2015), who provided acoustic and perceptual evidence for a boundary based

on final lengthening cues¹. Therefore, we assumed the target word was realised, in all conditions, as the metrical head of the intermediate phrase in which it was located.

3.2.1 Acoustic measurements of the stimuli

We measured Mass on the stressed syllable of target words along with the f₀ span between L and H tones associated with that word, and compared Mass and f₀ values across BF, NF and PF conditions. Mass is calculated as the integral of duration and intensity, considering these two cues together in one variable. In particular, we measured relative Mass, i.e. the prosodic strength in terms of the area under the periodic energy curve of each syllable relative to the other syllables in the same utterance. The average relative Mass centres at zero, thus values lower than zero indicate a weak syllable (in terms both of duration and intensity), whereas values higher than zero indicate a strong syllable (weak < 0 < strong). The f₀ span was calculated as the difference between the maximum and the minimum f₀ values corresponding to the H and L tones (H* and L for NF; H and L* for BF, and the highest and lowest point within the stressed syllable in PF).

Mass and f₀ span were analysed using two separate hierarchical linear regression models in the Bayesian framework (for details see section 3.4). Specifically, we modelled relative mass and f₀ values (dependent variables) as a function of the treatment contrast coded predictor FOCUS ('Narrow', 'Broad', 'Post-focal'), with broad focus serving as the reference level. Random intercepts for ITEM² were included. Random effects for PARTICIPANTS were not needed as only one speaker produced the speech stimuli. We used regularising (i.e. minimally informative) priors for all our parameters, allowing for a wide range of effect sizes, while discouraging implausibly large effects. The full prior and model specification can be found in the code provided in the OSF platform (<https://osf.io/6kjsx/>). Inferences were based on the posterior distributions of the parameters, reported in relation to the posterior mean, the 90% credible interval (CrI), and the posterior probability that the true parameter value lies on one side of zero and is close to one (Franke & Roettger 2019).

Four MCMC sampling chains were run for 8,000 iterations with a warm-up period of 2,000 iterations, yielding a total of 24,000 posterior samples. There was no indication for convergence issues of the model (all *Rhat* = 1.00). Model fit was also assessed by visual inspection of the predictive posterior checks.

3.2.1.1 Mass. Figure 3 illustrates the distribution of relative Mass on the stressed syllable of target words in narrow focus (NF), broad focus (BF) and post-focal (PF) condition. On the right side of the figure, the data is depicted in half violin plots, visualising the distributions as a kernel density plot. On the left side of the figure the data is depicted in dot plots, resembling a histogram as the data points are stacked into one bin. Lastly, the black dots indicate the mean of each distribution. Visual inspection of the distributions of relative mass values in Figure 3 shows a reduction in the proportion of stressed syllables that exhibit greater mass than utterances' average strength (mass > 0) from NF to BF to PF, causing an increase of the overall mean from PF to BF to NF.

Estimated condition Mass means are presented in Figure 4, and posterior distribution of the estimated parameters are depicted in Figure 5. The model estimated that Mass on the stressed syllable of target words in the (baseline) BF condition was $\hat{\beta} = 0.36$ (90%-CrI: [0.06,

¹ We cannot check for final lengthening in our stimuli as phrase finality would be confounded with focus on the noun.

² Random effects for items initially included a full variance-covariance matrix over focus in the model (e.g. Barr *et al.* 2013), but this led to divergent transitions (convergence issues) that were not possible to fix by increasing either *adapt_delta*, or *max_treedepth* parameters of the brms model. Yet, the non-convergent version of the model with (1 + focus | item) showed the same qualitative results.

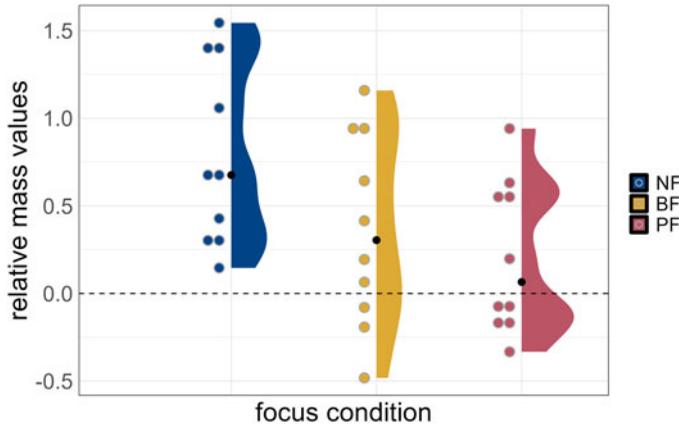


Figure 3 Distribution of relative Mass values associated with the stressed syllable of target words in NF (blue), BF (yellow) and PF (pink). For each condition, the data is depicted in two ways: to the right, in half violin plots, visualising the distributions as a kernel density plot; to the left the data is illustrated in dot plots, resembling a histogram as the data points are stacked. Black dots indicate the mean of each distribution.

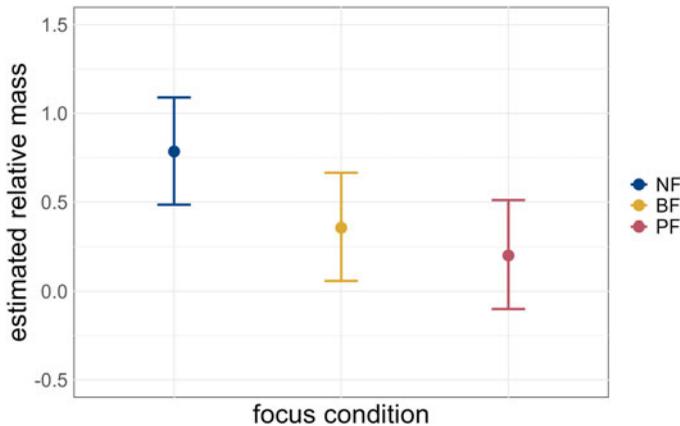


Figure 4 Estimated condition Mass means with 90% credible per focus condition (NF in blue, BF in yellow, PF in pink).

0.67]), in PF $\hat{\beta} = 0.20$ (90%-CrI: $[-0.10, 0.51]$) and in NF $\hat{\beta} = 0.79$ (90%-CrI: $[0.49, 1.09]$). The model provided strong evidence for the stressed syllable of target words showing higher Mass values in NF than in PF (0.79 vs. 0.20, $\hat{\beta} = 0.58$, 90%-CrI: $[0.48, 0.69]$, $P(\beta > 0) = 1$), as well as than in BF (0.79 vs. 0.36, $\hat{\beta} = 0.43$, 90%-CrI: $[0.32, 0.53]$, $P(\beta > 0) = 1$), indicating that the stressed syllable of target words in NF is prosodically stronger (in terms of both duration and intensity) than in BF and PF conditions. In turn, Mass on the stressed syllable of target words exhibited lower Mass values in PF than in BF conditions (0.20 vs 0.36, $\hat{\beta} = -0.16$, 90%-CrI: $[-0.26, -0.05]$, $P(\beta < 0) = 0.99$), indicating that the stressed syllables of target words in PF are prosodically weaker in PF than BF conditions.

These results suggest that the prosodic strength on target words increases from PF to BF and to NF, thus reflecting the different degrees of prosodic prominence of the target word in the different conditions.

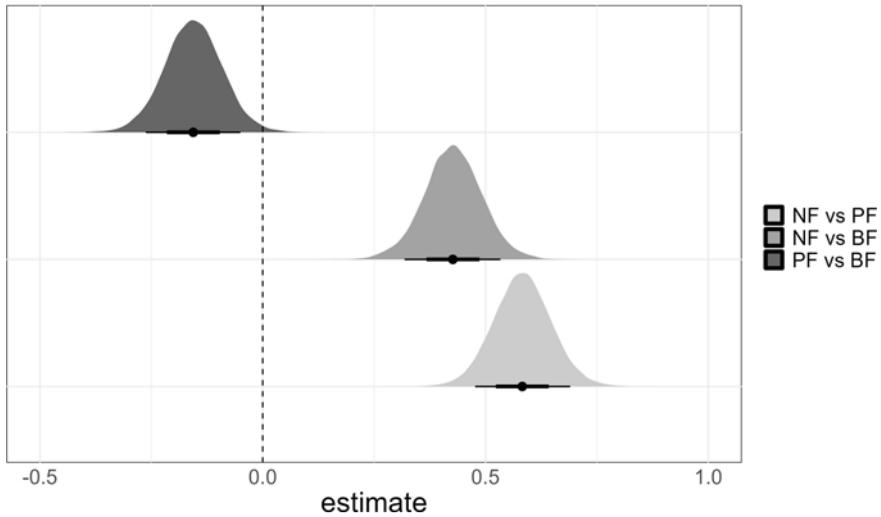


Figure 5 Posterior distributions for the focus effects on Mass. Error bars around the posterior means illustrate 66% (thick line) and 90% (light line) credible intervals.

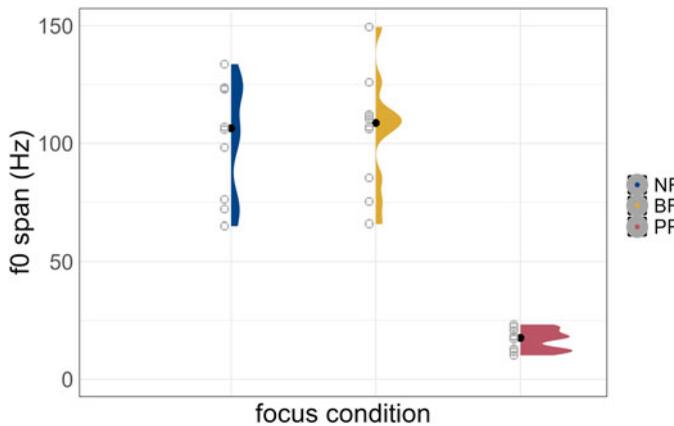


Figure 6 Distribution of *f0* span on the stressed syllable of target words in NF (in blue), BF (in yellow) and PF (in pink). On the right side of the figure, the data is depicted in half violin plots, visualising the distributions as a kernel density plot. On the left side of the figure the data is illustrated in dot plots, resembling a histogram as the data points are stacked into one bin. Black dots indicate the mean of each distribution.

3.2.1.2 *F0 span.* Figure 6 illustrates the distribution of *f0* span associated with the stressed syllable of target words in narrow focus (NF), broad focus (BF) and post-focal (PF). On the right side of the figure, the data is depicted in half violin plots, visualising the distributions as a kernel density plot. On the left side of the figure the data is depicted in dot plots, resembling a histogram as the data points are stacked into one bin. Lastly, the black dots indicate the mean of each distribution. Visual inspection of the distributions of *f0* span in Figure 6 shows that NF and BF exhibit *f0* span of similar height whereas, as expected, it is considerably lower in PF.

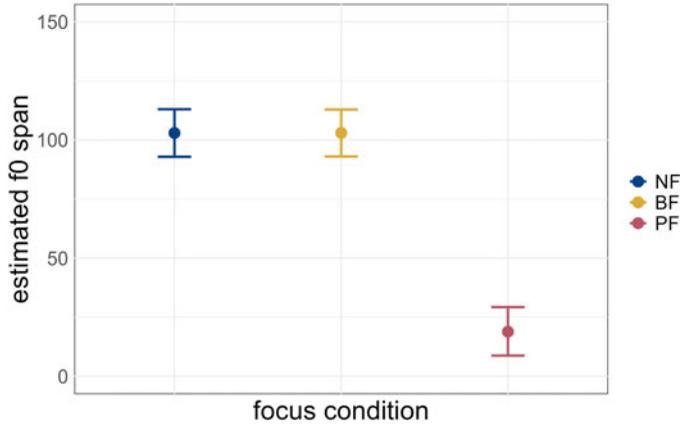


Figure 7 Estimated condition means for f0 span with 90% credible per focus condition (NF in blue, BF in yellow, PF in pink).

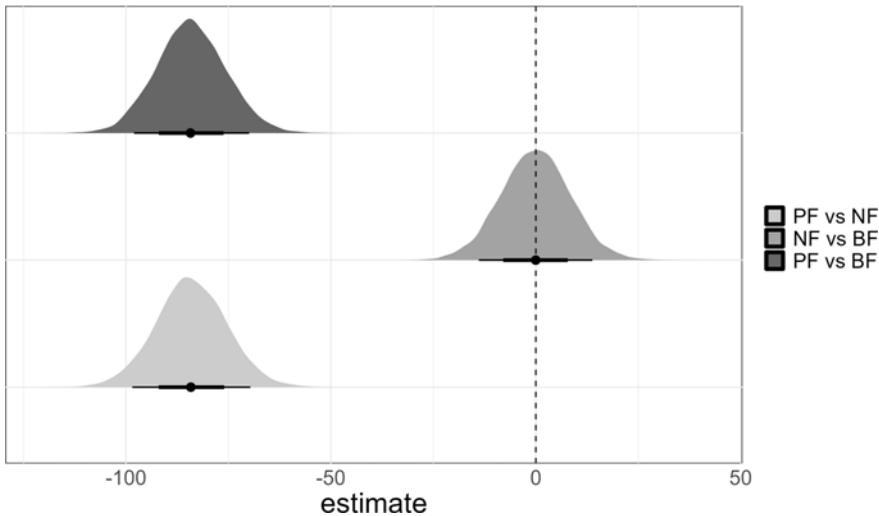


Figure 8 Posterior distributions for the focus effects on f0 span. Error bars around the posterior means illustrate 66% (thick line) and 90% (light line) credible intervals.

Estimated condition f0 span means are presented in Figure 7, and posterior distribution of the estimated parameters are depicted in Figure 8. The model provided strong evidence for the stressed syllable of target words showing lower f0 span values in PF than in NF ($\hat{\beta} = -84.07$, 90%-CrI: $[-98.47, -69.64]$, $P(\beta < 0) = 1$) as well as than in BF ($\hat{\beta} = -84.12$, 90%-CrI: $[-98.02, -69.95]$, $P(\beta < 0) = 1$). In turn, the model did not provide evidence for an f0 span difference between NF and BF conditions ($\hat{\beta} = -0.05$, 90%-CrI: $[-13.91, 13.79]$, $P(\beta > 0) = 0.5$). These results show not only the relative absence of f0 movement on the stressed syllable of words in the PF condition, but also that the f0 span in BF is similar to that in the NF condition.

3.3 Procedure

To run the perceptual rating task, we developed a web-based tool using a visual-analogue scale along the line of several studies dealing with similar types of tasks (e.g. Terken 1997; Gussenhoven & Rietveld 1998; Arnold, Wagner, & Möbius 2011). In our case, we represented degrees of prominence iconically by associating it with thickness and darkness of the scale (thicker and darker line indicating higher prominence, see Figure 9), as in Baumann & Röhr (2015). Our web-based rating task tool was implemented with the ‘SoSci Survey’ software (Leiner 2014); a screenshot of the Italian version is shown in Figure 9.

In the rating task, participants were asked to listen to each sentence and to rate the perceived prominence of three words contained in the sentence and selectively presented on the screen, by clicking on the corresponding position on the scale. These three words always corresponded to the verb preceding the target, the target noun and the noun in the following PP. So as not to ask the listeners to focus exclusively on the target word, they also rated the preceding verb and the following noun in the prepositional phrase (e.g. in *Bisogna girare la maniglia sulla porta*, the underlined words were rated).

For Italian listeners, the left pole of the visual analogue scale was labelled as ‘*per niente saliente*’ (‘not at all salient’) while the right pole was labelled as ‘*al massimo saliente*’ (‘maximally salient’). For German listeners, the labelling was ‘*gar nicht hervorgehoben*’ (‘not at all highlighted’) and ‘*sehr stark hervorgehoben*’ (‘very strongly highlighted’). We decided to use the Italian ‘*saliente*’ (‘salient’) and the German ‘*hervorgehoben*’ (‘highlighted’), as these terms were deemed less technical and thus more appropriate for listeners who are naïve to the purposes of the experiment.

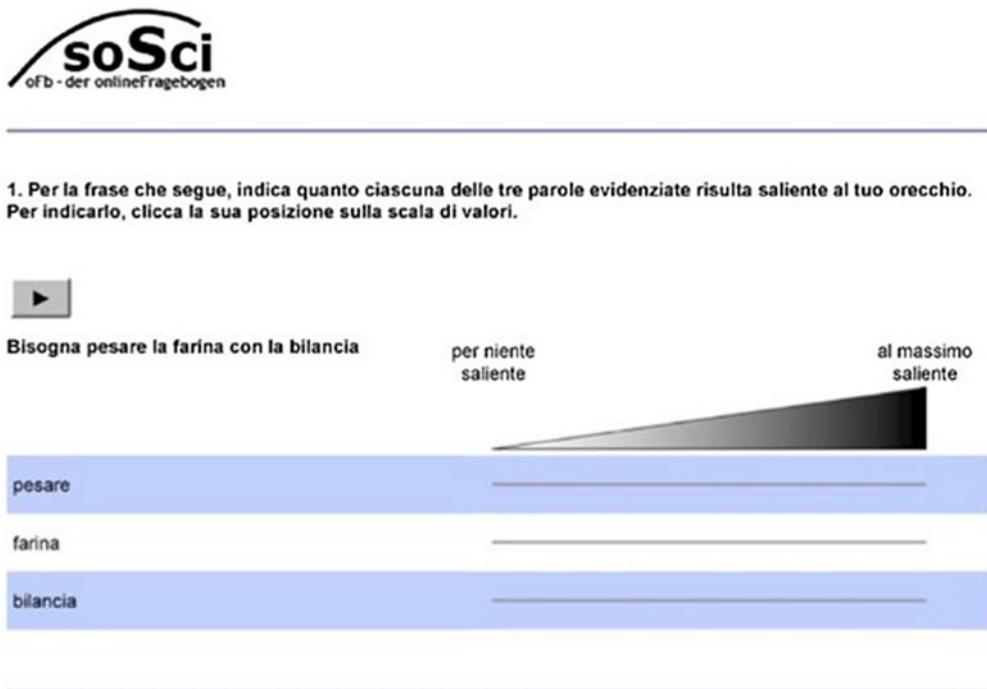


Figure 9 Screenshot of the web-based prominence rating task for the Italian participants for the sentence *Bisogna pesare la farina con la bilancia* (‘One needs to weigh the flour with the scales’). The target word is in the white row (here: *farina*, ‘flour’).

Instructions were designed to direct listeners towards an acoustically based interpretation, specified by the formulation ‘how salient/highlighted the words sound **to your ear**’. The instructions in each language are provided in Appendix B.

In the task, stimuli were presented both acoustically and orthographically, and participants were allowed to listen to each stimulus as many times as they wanted. The task was preceded by a short training session, during which no feedback was given to participants.

Stimuli were presented in a pseudo-randomised order, avoiding consecutive or close repetitions of the same utterance (independent of the prosodic condition), by ensuring repetitions were at least six stimuli apart. The same prosodic condition in different items never occurred consecutively, with a minimum of one other condition in-between. Different pseudo-randomised order-of-presentation lists were created and evenly distributed among participants.

As mentioned before, participants were allowed to listen to the stimuli as many times as they wanted before rating. However, to ensure homogeneity of the data, trials in which participants listened to a stimulus more than five times were excluded from further analyses, along with the corresponding items in the other two conditions. Moreover, one participant from the learner group was excluded from further analyses as this person listened to five trials more than five times, ranging from six to 13 times (the majority of participants listened to the stimuli less than five times: the median of the learner group = 2; the median of the native group = 1).

3.4 Statistical data analysis

All analyses (including those on Mass and f_0 span measurements of target words across BF, NF and PF conditions as described in sections 3.2.1.1 and 3.2.1.2) were conducted in R 4.1.2 (R Development Core Team 2021) using the *brms* 2.17.0 package (Bürkner 2018). For data processing we used the R package *tidyverse* 1.3.1 (Wickham et al., 2019) and for visualisations we used the R packages *ggplot2* 3.3.5 (Wickham 2016) and *tidybayes* 3.0.2 (Kay 2021). All data tables as well as scripts for the analyses are available at <https://osf.io/6kjxr/> on the OSF platform.

4. Results

As mentioned in section 3.3 above, our prominence rating data were collected using a visual-analogue rating scale, known also as a slider scale, allowing participants to assign continuous degrees of prominence ratings. Yet, slider scales are unlikely to provide normally distributed responses as the scale consists of lower and upper bounds (Vuorre 2018). Ratings were encoded as interval data from 0 (lower bound) to 100 (upper bound). For modelling purposes, the range was rescaled from 0–100 to 0–1 by dividing each value by 100.

Prominence ratings of the target words were analysed using a zero-one inflated Beta (ZOIB) hierarchical regression model in the Bayesian framework. The *brms* ZOIB family accounts for a continuous range between 0 and 1 as well as the actual presence of 0s and 1s in the data, i.e. data bounded at the end points of the scale (e.g. Liu & Kong 2015). Treatment contrast was used to code two categorical predictors: focus (‘Narrow’, ‘Broad’, ‘Post-focal’) with broad focus being the reference level, and group³ (‘Natives’, ‘L2 learners’) with natives being the reference level. We estimated both the effects of these factors

³ Proficiency level of the L2 learner group was included as predictor in a model, but it did not provide evidence for a reliable effect, thus we decided to consider the L2 group as having a homogeneous proficiency.

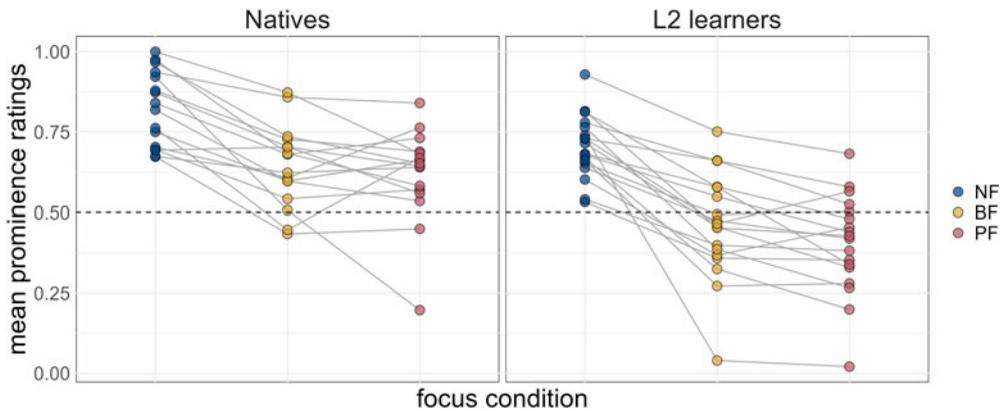


Figure 10 Mean prominence ratings on the target words across the three focal conditions (NF in blue, BF in yellow, PF in pink), as a function of group (left panel depicts natives, right panel depicts L2 learners), per participant.

and their interaction. The model also included random effects for participants and items with full variance-covariance matrices (e.g. Barr et al. 2013). We used regularising (i.e. minimally informative) priors for all our parameters, allowing for a wide range of effect sizes, discouraging though implausibly large effects. The full prior and model specification can be accessed in the code provided in the OSF platform (see section 3.4 for the link). Inferences were based on the posterior distributions of the parameters, reported in relation to the posterior mean, the 90% credible interval (CrI), and the posterior probability that the true parameter value lies on one side of zero and is close to one (Franke & Roettger 2019).

Four MCMC sampling chains were run for 8,000 iterations with a warm-up period of 2,000 iterations, yielding a total of 24,000 posterior samples. There was no indication for convergence issues in the model (all $Rhat = 1.00$). Model fit was also assessed by visual inspection of the predictive posterior checks.

Figure 10 depicts mean prominence ratings of the two groups, natives and L2 learners, on the target words⁴ across the three focal conditions (NF, BF, PF) per participant. Visual inspection of this figure exhibits a pattern of higher ratings from NF to BF to PF conditions. Further, we can observe that target words in all three focal conditions have native prominence ratings of over 50% whereas many L2 learners' ratings of target words in the broad and post-focal conditions are below 50%. Let us now move to the results of our model.

The model estimated that prominence rating of target words by native listeners in BF (baseline condition) was 64% ($\hat{\beta} = 0.56$, 90%-CrI: [0.29, 0.83]), in PF was 61% ($\hat{\beta} = 0.45$, 90%-CrI: [0.16, 0.73]), and in NF was 80% ($\hat{\beta} = 1.42$, 90%-CrI: [1.11, 1.72]). Similarly, prominence rating of target words by L2 learners in BF was 47% ($\hat{\beta} = -0.12$, 90%-CrI: [-0.39, 0.15]), in PF was 41% ($\hat{\beta} = -0.37$, 90%-CrI: [-0.65, -0.08]), and in NF was 72% ($\hat{\beta} = 0.94$, 90%-CrI: [0.64, 1.23]). The estimated means of prominence ratings (back transformed from log-odds to proportions) in each condition are shown in Figure 11.

For the native group, the model does not provide robust evidence for target words in PF being rated much lower than target words in the baseline BF (61% versus 64%, $\hat{\beta} = -0.11$, 90%-CrI: [-0.31, 0.09], $P(\beta < 0) = 0.83$). Yet, ratings in NF were reliably higher than both

⁴ Although our main focus is on prominence ratings on the target word (the object noun, e.g. in the example given in (1) *maniglia* 'handle'), we provide in the OSF repository results from the ratings of the verb preceding the object noun (*girare* 'turn') and of the noun (*porta* 'door') in the following prepositional phrase (*sulla porta* 'on the door').

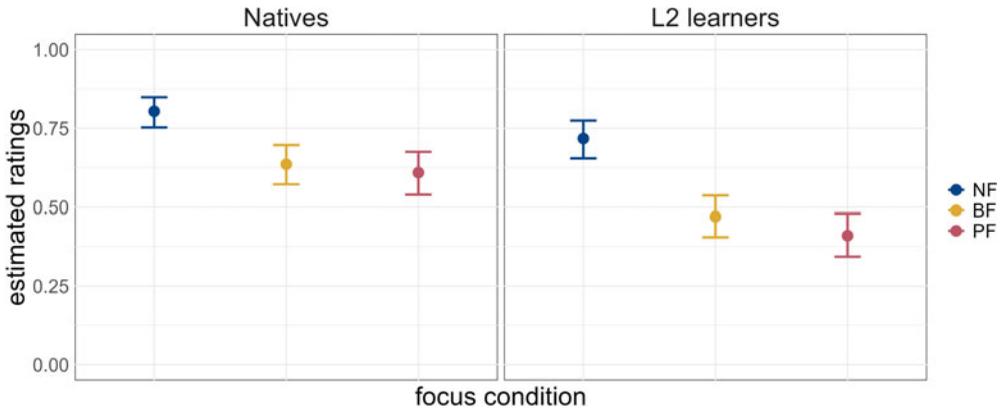


Figure 11 Estimated condition means of prominence ratings with 90% credible intervals per group (left panel: natives; right panel: L2 learners) back transformed from log-odds to the original scale (proportions).

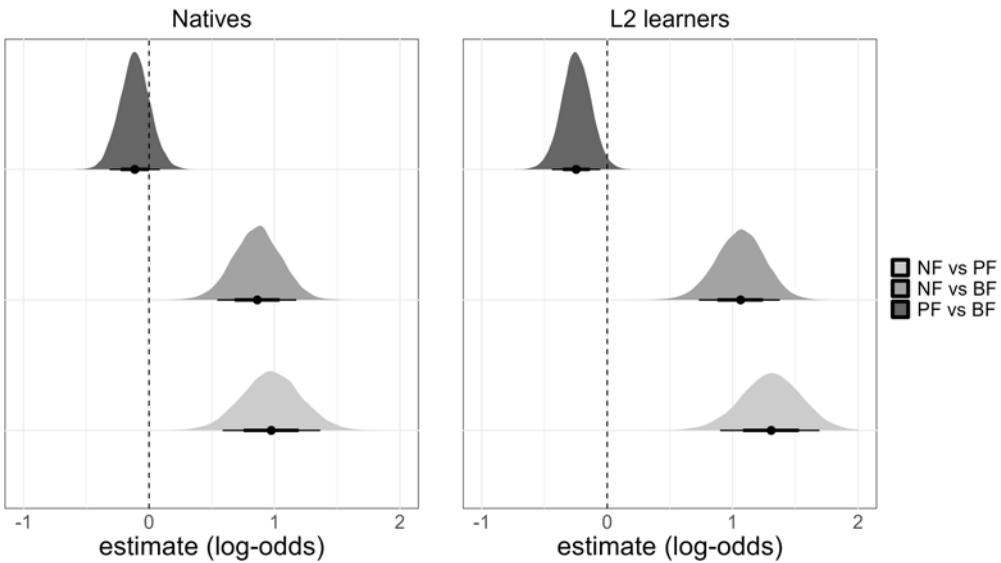


Figure 12 Posterior distributions for the focus effects on prominence ratings as a function of group. All parameters are on log-odds scale. Error bars around the posterior means illustrate 66% (thick line) and 90% (light line) credible intervals.

in BF (80% versus 64%, $\hat{\beta} = 0.86$, 90%-CrI: [0.54, 1.17], $P(\beta > 0) = 1$), and in PF (80% versus 61%, $\hat{\beta} = 0.97$, 90%-CrI: [0.59, 1.37], $P(\beta > 0) = 1$). These results indicate that target words were perceived by native listeners as more prominent in narrow focus than in the other two focal conditions. Yet, target words in broad and post-focal conditions were perceived as equally prominent. Posterior distributions of the parameters are illustrated in the left panel of Figure 12.

For the L2 group, we find compelling evidence for a rating difference between BF and PF, with target words in PF being rated lower than in BF (41% versus 47%, $\hat{\beta} = -0.25$, 90%-CrI: [-0.44, -0.05] $P(\beta < 0) = 0.98$). In the case of NF, target words were rated reliably higher than in BF (72% versus 47%, $\hat{\beta} = 1.06$, 90%-CrI: [0.73, 1.38], $P(\beta > 0) = 1$) as well as in PF

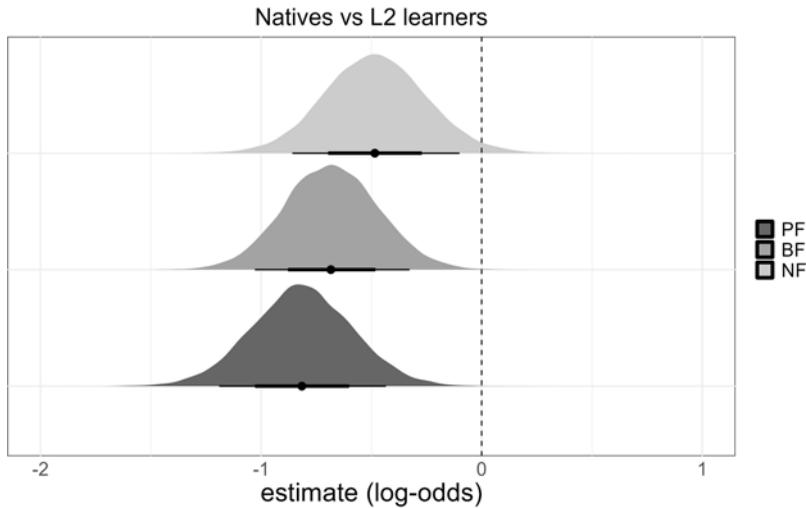


Figure 13 Posterior distributions for the between group differences of the focus effects on prominence ratings. All parameters are on log-odds scale. Error bars around the posterior means illustrate 66% (thick line) and 90% (light line) credible intervals.

(72% versus 41%, $\hat{\beta} = 1.31$, 90%-CrI: [0.9, 1.7], $P(\beta > 0) = 1$). These differences suggest that L2 learners considered target words in NF as more prominent than in BF and PF, in a similar way to natives. However, contrary to natives, target words appear to have been perceived as less prominent in post-focal position than in broad focus. Posterior distributions of the parameters, for the L2 group, are illustrated in the right panel of Figure 12.

Overall, we found that target words in NF were rated as more prominent than in BF and PF condition by both groups. However, although L2 learners rate target words as less prominent in PF than in BF, natives rate these two conditions as equally prominent. To assess group differences directly, we compare the effect of NF, BF and PF on prominence ratings between groups. The resulting posteriors suggest a strong difference between natives and L2 learners, with L2 learners assigning lower ratings across all conditions compared to natives. Additionally, the posteriors reveal a reliable decrease in prominence ratings from NF to BF to PF by the L2 group compared to the native group (NF, $\hat{\beta} = -0.48$, 90%-CrI: [-0.86, -0.1], $P(\beta < 0) = 0.98$; BF, $\hat{\beta} = -0.68$, 90%-CrI: [-1.03, -0.33], $P(\beta < 0) = 1$; PF, $\hat{\beta} = -0.81$, 90%-CrI: [-1.19, -0.43], $P(\beta < 0) = 1$). Posterior distributions are shown in Figure 13.

5. Discussion and conclusions

The results generally indicate that when assigning prominence values to words in sentence context, native listeners may take both the acoustic signal and structural expectations into account, whereas learners rely more on the acoustic signal.

Prominence ratings of the target words in narrow focus were consistently higher than for broad focus across the two groups. However, although learners rated words in broad focus as more prominent than those in post-focal position, native speakers of Italian did not rate these differently. This result is particularly striking, given the differences in the acoustic realisation of the target words in the two conditions. The broad focus words have

a greater Mass (indicating that they were longer and louder) and a considerably greater f_0 span. The fact that Italian listeners did not take this acoustic difference into account indicates that they were using a top-down strategy in their perception of prominence. By contrast, the German learners appeared to be guided by the acoustic saliency of the broad focus words compared to the reduced Mass and f_0 span in post-focal position, leading to lower prominence ratings in post-focal position. In this case their judgements could be based on a bottom-up strategy, which is understandable, given their learner status.

Moreover, our results confirm that the preceding prosodic context may also play a role in building up expectations of prominence, in the sense that prominence perception may rely on a language-specific distribution of accents. Learners appear to interpret the low flat pitch of post-focal targets as low in prominence, both because they have less recourse to distributional properties of accents in the language, and possibly also because this is how flat pitch would be in their native language (Röhr & Baumann 2010). By contrast, Italian listeners interpret low flat pitch as similarly prominent to dynamic pitch and appear to rely more on their native language-driven expectations to find cues for prominence in post-focal position (Bocci & Avesani 2011, 2015). This tendency could be explained by the higher-level structures of the language that play a role in the generation of expectations for the realisation of upcoming words and by the fact that learners seem to prioritise bottom-up inferences (being guided more by the signal).

In line with the study by Ventura et al. (2020), where attention to post-focal words is enhanced by flat intonation post-focally (or blocked by a post-focal dynamic pitch accent), it appears that flat pitch has a different status in Italian to German, reflecting the different prosodic systems of the two languages. Since native speakers of Italian perceive these flat stretches as relatively prominent, these results support the analysis of Bocci & Avesani (2011, 2015) who, despite the flat pitch, treat words in this position as accented. Consequently, the analysis of German as a deaccenting language and Italian as resistant to deaccenting appears to hold.

Acknowledgements We would like to thank Caterina Ventura for her valuable contribution in an earlier stage of this study. We also gratefully acknowledge funding from the German Research Foundation (Project ID 281511265 ‘Prominence in Language’). Preliminary results from this dataset were presented at the third Conference on Phonetics and Phonology in Europe (PaPE 2019).

Appendix A

Stimuli and elicitation contexts

Below the sentences selected for the stimuli in the perception experiment are reported. The questions provided were used as contexts for the different focus structures (numbers refer to audio recordings, which can be found at <https://osf.io/6kjxr/>).

Target word is marked in bold. Elicited prosodic realisations are Broad Focus (BF), contrastive Narrow Focus (NF) and Post-Focal (PF).

Sentence 02

Bisogna prendere **gli stivali** per il fango

(One should take **the boots** for the mud)

BF

Cosa bisogna fare quando piove?

(What should one do when it rains?)

NF

Per il fango bisogna prendere gli stivali o i sandali?

(For the mud should one take the boots or the sandals?)

PF

Quando piove bisogna prendere o buttare gli stivali per il fango?

(When it rains should one take or throw away the boots?)

Sentence 03

Bisogna girare **la maniglia** sulla porta

*(One should turn **the handle** on the door)*

BF

Cosa bisogna fare per entrare nella stanza?

(What should one do to enter the room?)

NF

Per entrare nella stanza bisogna girare la maniglia o il chiavistello sulla porta?

(To enter the room should one turn the handle or the latch on the door?)

PF

Per entrare nella stanza bisogna girare o spingere la maniglia sulla porta?

(To enter the room should one turn or push the handle on the door?)

Sentence 04

Bisogna portare **il maglione** nella borsa

*(One should carry **the jumper** in the bag)*

BF

Cosa bisogna fare per non avere freddo?

(What should one do so as not to be cold?)

NF

Quando fa freddo bisogna portare il maglione o il cappello nella borsa?

(When it is cold should one carry the jumper or the cap in the bag?)

PF

Per non avere freddo bisogna portare o lasciare il maglione nella borsa?

(To not be cold, should one carry or leave the jumper in the bag?)

Sentence 05

Bisogna cuocere **le ciliegie** nella pentola

*(One should cook **the cherries** in the pot)*

BF

Cosa bisogna fare per preparare la marmellata?

(What should one do for making jam?)

NF

Per preparare la marmellata bisogna cuocere le ciliegie o i semi nella pentola?

(For making jam, should one cook the cherries or the seeds in the pot?)

PF

Per preparare la marmellata bisogna cuocere o schiacciare le ciliegie nella pentola?

(For making jam, should one cook or mash the cherries in the pot?)

Sentence 07

Bisogna avere **la cantina** con le botti

*(One should have **the cellar** with the barrels)*

BF

Cosa bisogna fare per produrre il vino?

(What should one do to produce wine?)

NF

Per produrre il vino bisogna avere la cantina o la cucina con le botti?
 (For producing wine, should one have the cellar or the kitchen with the barrels?)
 PF

Per produrre il vino bisogna avere o noleggiare la cantina con le botti?
 (For producing wine, should one own or rent the cellar with the barrels?)

Sentence 08

Bisogna usare **la pianola** con frequenza
 (One should use **the pianola** with frequently)

BF

Cosa bisogna fare per imparare a suonare?
 (What should one do to learn to play?)

NF

Per imparare a suonare bisogna usare la pianola o il computer con frequenza?
 (To learn to play should one use the pianola or the computer frequently?)

PF

Per imparare a suonare bisogna usare o osservare la pianola con frequenza?
 (To learn to play should one use or watch the pianola frequently?)

Sentence 14

Bisogna coprire **la ferita** con la pomata
 (One should cover **the wound** with the ointment)

BF

Cosa bisogna fare quando ci si taglia?
 (What should one do if one cuts oneself?)

NF

Bisogna coprire la ferita con la pomata o con la garza?
 (Should one cover the wound with the ointment or the gauze?)

PF

Bisogna coprire o grattare la ferita con la pomata?
 (Should one cover or scratch the wound with the ointment?)

Sentence 15

Bisogna comprare **un regalo** per gli amici
 (One should buy **a present** for one's friends)

BF

Cosa bisogna fare a Natale?
 (What should one do at Christmas?)

NF

Bisogna comprare un regalo o una pianta per gli amici?
 (Should one buy a present or a plant for one's friends?)

PF

Bisogna comprare o riciclare un regalo per gli amici?
 (Should one buy or recycle a present for one's friends?)

Sentence 19

Bisogna pesare **la farina** con la bilancia
 (One should weigh the flour with the scales)

BF

Cosa bisogna fare per preparare la torta?
 (What should one do to prepare a pie?)

NF

Bisogna pesare la farina o la vanillina con la bilancia?
(Should one weigh the flour or the vanillin with the scales?)

PF

Per fare una torta bisogna pesare o filtrare la farina con la bilancia?
(To prepare a pie, should one weigh or filter the flour with the scales?)

Sentence 20

Bisogna sbattere la tovaglia dalle briciole
(One should shake off the crumbs from the tablecloth)

BF

Cosa bisogna fare dopo il pranzo?
(What should one do after lunch?)

NF

Bisogna sbattere la tovaglia o i tovaglioli dalle briciole?
(Should one shake off the crumbs from the tablecloth or the napkin?)

PF

Bisogna sbattere o lavare la tovaglia dalle briciole?
(Should one shake off or wash the crumbs from the tablecloth?)

Appendix B

Instructions for the prominence rating task

For the Italian listeners:

Per la frase che segue, indica quanto ciascuna delle tre parole evidenziate risulta saliente al tuo orecchio. Per indicarlo, clicca la sua posizione sulla scala di valori.

For the German learners of Italian:

Geben Sie für den folgenden Satz an, wie hervorgehoben die drei Wörtern in Ihren Ohren klingen. Um dies anzuzeigen, klicken Sie jeweils auf eine Position der Werteskala.

Translation into English:

‘For the following sentence, indicate how each of the words sound highlighted to your ear. To indicate this, click a position on the scale.’

References

- Albert, Aviad. 2023. A model of sonority based on pitch intelligibility. *Studies in Laboratory Phonology* 13. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.7837176>
- Albert, Aviad, Francesco Cangemi, Timothy Mark Ellison, & Martine Grice. 2022. ProPer: PROsodic Analysis with PERiodic Energy. OSF. <https://doi.org/10.17605/OSF.IO/28EA5>
- Arnold, Denis, Petra Wagner, & Bernd Möbius. 2011. Evaluating different rating scales for obtaining judgments of syllable prominence from naïve listeners. In Wai Sum Lee & Eric Zee (eds.), *17th International Congress of Phonetic Sciences (ICPhS XVII)*, Hong Kong, 252–255.
- Barr, Dale J., Roger Levy, Christoph Scheepers, & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Baumann, Stefan & Christine Röhr. 2015. The perceptual prominence of pitch accent types in German. In: The Scottish Consortium for ICPhS 2015 (eds.), *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS XVIII)*, Glasgow, UK, 298–303.
- Bishop, Jason. 2012. Information structural expectations in the perception of prosodic prominence. In: Gorka Elordieta & Pilar Prieto (eds.), *Prosody and Meaning*, 239–270. Berlin: Walter de Gruyter.

- Bishop, Jason. 2016. Individual differences in top-down and bottom-up prominence perception. In Jonathan Barnes, Alejna Brugos, Stefanie Shattuck-Hufnagel, & Nanette Veilleux (eds.), *Proceedings of Speech Prosody 2016*, Boston, 668–672.
- Bocci, Giuliano & Cinzia Avesani. 2011. Phrasal prominences do not need pitch movements: Post-focal phrasal heads in Italian. In Piero Cosi & Renato De Mori (eds.), *Proceedings of the 12th Annual Conference of the International Speech Communication Association*, Florence, Italy, 1357–1360.
- Bocci, Giuliano & Cinzia Avesani. 2015. Can the metrical structure of Italian motivate focus fronting? In Ur Shlonsky (ed.), *Beyond functional sequence: The cartography of syntactic structures*, 23–41. New York, NY: Oxford University Press.
- Boersma, Paul & David Weenink. 2018. Praat: Doing phonetics by computer (version 6.0.43). <http://www.praat.org/> [last accessed 10 July 2025].
- Bürkner, Paul-Christian. 2018. Advanced Bayesian multilevel modelling with the R package brms. *The R Journal* 10(1), 395–411. <https://doi.org/10.32614/RJ-2018-017>
- Calhoun, Sasha, Emma Wollum, & Emma Kruse Va'ai, E. 2021. Prosodic prominence and focus: Expectation affects interpretation in Samoan and English. *Language and Speech* 64 (2), 346–380.
- Cangemi, Francesco & Mariapaola D'Imperio. 2013. Tempo and the perception of sentence modality. *Journal of the Association for Laboratory Phonology* 4(1), 191–219.
- Cangemi, Francesco & Martine Grice. 2016. The importance of a distributional approach to categoricity in autosegmental-metrical accounts of intonation. *Laboratory Phonology* 7(1), 9. <https://doi.org/10.5334/labphon.28>
- Cole, Jennifer, Yoonsook Mo, & Mark Hasegawa-Johnson. 2010. Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology* 1, 425–452.
- Cruttenden, Alan. 1993. The de-accenting and re-accenting of repeated lexical items. In: Paul Touati & David House (eds.), *Proceedings of the ESCA workshop on Prosody*, Lund, 16–19.
- Dehé, Nicole & Vieri Samek-Lodovici. 2009. On the prosody and syntax of DPs: Evidence from Italian noun adjective sequences. *Natural Language & Linguistic Theory* 27(1), 45–75.
- Eriksson, Anders, Gunilla C. Thunberg, & Hartmut Traunmüller. 2001. Syllable prominence: A matter of vocal effort, phonetic distinctness and top-down processing. In: Paul Dalsgaard, B. Lindberg, Henrik Benner, & Zheng-Hua Tan (eds.), *Proceedings of Eurospeech 2001*, Aalborg, 399–402.
- Féry, Caroline & Frank Kügler. 2008. Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics* 36(4), 680–703.
- Frank, Michael & Timo B. Roettger. 2019. Bayesian regression modeling (for factorial designs): A tutorial. PsyArXiv. <https://doi.org/10.31234/osf.io/cdxv3>
- Gili Fivela, Barbara, Cinzia Avesani, Marco Barone, Giuliano Bocci, Claudia Crocco, Mariapaola D'Imperio, Rosa Giordano, Giovanna Marotta, Michelina Savino, & Patrizia Soriano. 2015. Intonational phonology of the regional varieties of Italian. In: Sonia Frota & Pilar Prieto (eds.), *Intonation in Romance*, 140–197. Oxford: Oxford University Press.
- Grice, Martine. 1995. *The intonation of interrogation in Palermo Italian: Implications for intonation theory*. Tübingen: Niemeyer.
- Grice, Martine, Mariapaola D'Imperio, Michelina Savino, & Cinzia Avesani. 2005. Strategies for intonation labelling across varieties of Italian. In: Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, 362–389. Oxford: Oxford University Press.
- Grice, Martine & Michelina Savino. 1997. Can pitch accent type convey information-status in yes-no questions? In: Kai Alter, Hannes Pirker, & Wolfgang Finkler (eds.), *Proceedings of the ACL Workshop "Concept to Speech Generation Systems"*, Madrid, 29–38.
- Grice, Martine & Michelina Savino. 2003. Map Tasks in Italian: Asking questions about given, accessible and new information. *Catalan Journal of Linguistics* 2, 153–180.
- Grice, Martine & Michelina Savino. 2004. Information structure and questions – Evidence from task-oriented dialogues in a variety of Italian. In: Peter Gilles & Jörg Peters (eds.), *Regional variation in intonation*, 161–190. Tübingen: Niemeyer.
- Grice, Martine, Michelina Savino, & Timo B. Roettger. 2019. Tune-text negotiation: The effect of intonation on vowel duration. In: Sasha Calhoun, Paola Escudero, Marija Tabain, & Paul Warren (eds.), *Proceedings of the 19th International Congress on Phonetic Sciences (ICPhS XVIV)*, Melbourne, 562–566.
- Gussenhoven, Carlos & Toni Rietveld. 1998. On the speaker-dependence of the perceived prominence of F0 peaks. *Journal of Phonetics* 26(4), 371–380.
- Kay, Matthew. 2021. *tidybayes: Tidy Data and Geoms for Bayesian Models*. R package version 3.0.2. 10.5281/zenodo.1308151

- Kügler, Frank. 2008. The role of duration as a phonetic correlate of focus. In: Plinio A. Barbosa, Sandra Madureira, & Cesar Reis Barbosa (eds.), *Proceedings of Speech Prosody 2008*, Campinas, Brazil, 591–594.
- Kügler, Frank & Caroline Féry. 2017. Postfocal downstep in German. *Language and Speech* 60(2), 260–288.
- Ladd, Robert D. 2008. *Intonational phonology*. Cambridge: Cambridge University Press.
- Leiner, Dominik J. 2014. SoSci Survey (version 2.4.00-i) [Computer Software]. <https://www.sosicisurvey.de>
- Liu, Fang & Yunchuan Kong. 2015. zoib: An R package for Bayesian inference for Beta Regression and Zero/One Inflated Beta Regression. *The R Journal* 7(2), 34–51. <https://doi.org/10.32614/RJ-2015-019>
- Lorenzen, Janne & Stefan Baumann. 2024. Does communicative skill predict individual variability in the prosodic encoding of lexical and referential givenness? In Yiya Chen, Aojun Chen, & Amalia Arvaniti (eds.), *Proceedings of Speech Prosody 2024*, Leiden, 802–806.
- Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris Publications.
- R Development Core Team (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Röhr, Christine T. & Stefan Baumann. 2010. Prosodic Marking of Information Status in German. In *Proceedings of 5th International Conference on Speech Prosody*, vol. 100019, 1–4. Chicago, USA.
- Savino, Michelina. 2012. The intonation of polar questions in Italian: Where is the rise? *Journal of the International Phonetic Association* 42(1), 23–48.
- Sbranna, Simona, Caterina Ventura, Aviad Albert, & Martine Grice. 2023. Prosodic marking of information status in Italian. *Journal of Phonetics* 97, 101212. <https://doi.org/10.1016/j.wocn.2023.101212>.
- Selkirk, Elizabeth. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Selkirk, Elizabeth. 1995. Sentence prosody: Intonation, stress and phrasing. In: John A. Goldsmith (ed.), *The handbook of phonological theory*, 550–569. Oxford: Blackwell.
- Swerts, Marc, Emiel Krahmer, & Cinzia Avesani. 2002. Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics* 30(4), 629–654.
- Terken, Jacques. 1997. Variation of accent prominence within the phrase: Models and spontaneous speech data. In: Yoshinori Sagisaka, Nick Campbell, & Norio Higuchi (eds.), *Computing prosody: Computational models for processing spontaneous speech*, 95–116. New York, NY: Springer.
- Ventura, Caterina, Martine Grice, Michelina Savino, Diana Kolev, Ingmar Brilmayer, & Petra B. Schumacher. 2020. Attention allocation in a language with post-focal prominences. *Neuroreport* 31, 624–628.
- Vuorre, Matti. 2018. How to analyze visual analog (slider) scale data? Sometimes | R. <https://vuorre.com/posts/2019-02-18-analyze-analog-scale-ratings-with-zero-one-inflated-beta-models/>
- Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer-Verlag.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy McGowan, Romain François, Garrett Grolemund, Alex Hayes, Lionel Henry, Jim Hester, Mark Kuhn, Thomas Pedersen, Evan Miller, Stefan Bacher, Kirill Müller, Jeroen Ooms, David Robinson, Dana Seidel, Vitalie Spinu, et al. 2019. Welcome to the tidyverse. *Journal of Open Source Software* 4(43), 1686. <https://doi.org/10.21105/joss.01686>

Cite this article: Grice Martine, Savino Michelina, Lialiou Maria, Albert Aviad, and Schumacher Petra (2025). The perception of post-focal prominence is (also) a matter of linguistic structure. *Journal of the International Phonetic Association*. <https://doi.org/10.1017/S0025100325100479>