Focus and Prosodic Cues in Hungarian Noun Phrases

Corinna Langer[1], *Frank Kügler*[1]

[1]Goethe-University Frankfurt, Germany

langer@lingua.uni-frankfurt.de, kuegler@em.uni-frankfurt.de

Abstract

This study investigates the prosodic patterns of complex noun phrases in Hungarian in order to disentangle the roles of structural (syntactic) and prosodic prominence. While Hungarian has fixed syntactic positions for (narrow, exclusive) focus and topic, the structural marking of information structure cannot disambiguate multiple possible interpretations when the focus position is filled with complex phrases. We conducted a production study with complex noun phrases in the syntactic focus position, where contexts elicited either focus on the whole phrase or only on parts of it (e.g. only the noun). Our results show that while there is predominantly consistent prosodic marking of the left-edge of the syntactic focus position, there are different prosodic patterns in the NP depending on the position and the domain of focus. Thus, despite the structurally prominent syntactic focus position in Hungarian, prosodic prominence marking plays an independent role in focus marking.

Index Terms: information structure, prosody, prominence, modification, Hungarian

1. Introduction

Informations structure (IS) or information packaging (see, e.g. [1]) is the discourse relevant structuring of information within a sentence or utterance. Well studied notions of IS are topic and focus, which are also frequently discussed in work on Hungarian, especially on Hungarian syntax (e.g., [2], [3]). Following [4, 247] (which is based on *Alternative Semantics* [5]) we define focus as emphasizing the selection of one element out of a set of possible alternatives:

1. 'Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions.'

Focus is typically marked by prominence, which can be expressed through prosody, syntax or a combination of the two (see, e.g., [6]). For example, in intonation languages like English, focus is often marked by prosodic prominence, which is expressed through higher pitch (fundamental frequency; f0), longer duration of segments or syllables, and/or higher intensity (see, e.g. [7] for an overview). In languages with relatively free word order, such as Hungarian, IS is often marked via syntax.

Focus can have different domains or scopes in a sentence. According to [8], *broad* focus is focus that scopes over a whole sentence, while *narrow* focus only emphasizes constituents (of various sizes) of a sentence. English examples for broad and narrow focus with their often assumed explicit or implicit *question under discussion* [9] or *current question* [10] can be seen in 2 and 3. Focus domain is indicated by square brakets and 'F'.

2. Broad focus: *What happened?*
   [The girl ate the tasty muffin.]$_F$
3. Narrow focus:
   (a) *What did the girl do?*
   The girl [ate the tasty muffin]$_F$.
   (b) *What did the girl eat?*
   The girl ate [the tasty muffin]$_F$.
   (c) *Which muffin did the girl eat?*
   The girl ate the [tasty]$_F$ muffin.

Hungarian is a discourse-configurational language with fixed positions for (narrow, exclusive) focus in the immediate pre-verbal position and topic in the sentence-initial position. In sentences without a narrow focus, there are pitch accents on every content word in a downstep-relation, i.e. every accent is lower than the preceding one. Prosodic words in Hungarian are left-headed, thus, all accents are word initial (see, e.g. [11], [12]). If a constituent, e.g. a noun phrase (NP), is focused, it moves to the syntactic focus position and is assigned the main accent of the sentence, 'erasing' or 'reducing' the following accents (see, e.g., [13], [14], [11]).

If only parts of a phrase are focused (as in 3c), it is still the whole phrase that moves to the focus position, because elements such as, e.g., a modifier in an NP cannot move out of the phrase they are part of in order to be focused. In these cases, an ambiguitiy arises between an interpretation, where the whole phrase is focused, and interpretations, where only parts are in focus. Thus, the sentence in 4 could be the answer to each of the questions in 4a-4c.

4. *'The journalist invited the old writer.'*[1]
   Az idős író-t hívta meg az újságíró.
   the old writer-ACC invided VPRT the journalist
   (a) Who did the journalist invite?
   (b) Which writer did the journalist invite?
   (c) Which old person did the journalist invite?

The ambiguities shown in 4 cannot be syntactically disambiguated. Based on experimental reading data, we argue that this disambiguation can be achieved by prosodically marking the focused word(s) to signal focus prominence within the structurally prominent pre-verbal syntactic position. This is in line with, e.g., [15] who found that even in unambiguous sentences with only a simple noun in the focus position, the prosodic pattern changes depending on the information status of the background and the type of focus (contrastive vs. non-contrastive).

Cross-lingustically, there are multiple prosodic accentuation strategies to disambiguate focus within NPs. Languages can, for example, (i) accentuate the focused elements of the NP (see, e.g. [16] for Dutch), (ii) accentuate every element of the NP (see, e.g. [16] for Italian) or (iii) accentuate the element at one of the edges of the focus domain (see, e.g. [17] for Dutch & German). For Hungarian, one could also hypothesize - contrary to [15] - (iv) that focus in NPs is not prosodically marked since the focused NP appears in a structurally prominent, syntactic position. Structural prominence has been shown to be a cross-linguistic expectation-based predictor for prominence [18], and

[1]Abbreviations used in this paper: ACC = acusative; Adj = adjective; Det = determiner; N = noun; V = verb; VPRT = verbal particle

listeners are more likely to interpret a word as being the most prominent one in such a structural position (see, e.g., [19]).

In this study, we investigate the prosodic patterns of sentences like 4 where the structural focus position is filled with a complex NP in order to establish if and how Hungarian uses prosody to disambiguate different possible focus domains when syntax cannot be used. Thus we try to disentangle the roles of structural (syntactic) focus marking and prosodic prominence marking and examine which (if any) of the possible accentuation strategies Hungarian uses.

2. Production Experiment

Our study adresses two main questions: (Q1) Does Hungarian disambiguate different focus positions and domains in complex NPs through prosody? (Q2) If yes, which of the aforementioned accentuation strategies (i)-(iv) does Hungarian use? To answer these questions, we conducted a production experiment with complex NPs in the syntactic focus position.

2.1. Method

The complex NPs used in our experiment consisted of one adjective and a noun. They were embedded in a context that elicited different focus domains in the target sentence, with focus either on the adjective, the noun, or the whole NP. The linearization of the target sentence was always as follows:

5. Linearization of the target sentence:
   [Det Adj N]$_{\text{target NP}}$ V VPRT [...]

We chose particle verbs for the target sentence, because the verbal particle only moves out of its default position in front of the verb if the immediate pre-verbal focus position is filled (see, e.g., [20]). Thus, the inverted order of verb and particle is a clear syntactic indicator for a filled focus position, which, in our case, was occupied by a complex NP. The material after the verbal particle differed in length but always consisted of at least one NP.

There were five target sentences, each embedded into three different contexts eliciting the 3 focus conditions (focus on the adjective (Adj_F), on the noun (N_F), or on the whole NP (NP_F)). Each context ended in a question that the target sentence was the answer to. All elements of the target sentence were textually given in the context to avoid influences of given vs. new information. Example 6 shows a context with Adj_F.

6. Context: *'In Germany, the journalist met a young writer and an old writer. He found one of the two so kind that he invited him for his birthday party. Which writer did the journalist invite?*'
   Az idős író-t hívta meg az újságíró.
   the [old]$_{\text{F}}$ writer-ACC invited VPRT the journalist
   Target: *'The journalist invited the OLD writer.'*

All contexts were recorded by a female native speaker of Hungarian. In the experiment, the contexts were presented both auditorily and visually in a randomized order. The participants were asked to read the visually presented target sentence as a reply to the question in the context. The whole experiment was recorded using a Zoom H4n Pro recorder and a Rode Lavalier condenser microphone in a quiet environment.[2] 20 native speakers of Hungarian participated in the study (16 female (18-72 years old); 4 male (18-55 years old)) resulting in 300 datapoints (5 target sentences x 3 contexts x 20 participants).

[2]Due to the Covid-19 pandemic, it was not possible to record in the sound proof booth in our laboratory.

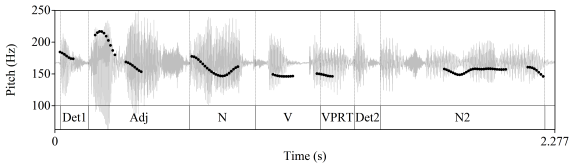

Figure 1: *Waveform and pitch track of sentence 6 with focus on the adjective produced by a female participant*

2.2. Data analysis

Using Praat [21], all sentences were extracted from the sound files and manually labeled for words, syllables (only in the NP) and vowels. A Praat script then detected and labeled the highest and lowest f0 per vowel. These points where checked and corrected manually by the first author to avoid microprosodic influences and calculation errors. At the same time, voice quality, i.e. the presence of creaky voice, and pauses were labeled.

For the target NPs, we extracted f0-minimum and -maximum in semitones (ST) with speaker-dependent baselines (see, e.g., [22]), their position in relation to the vowel, the syllable and the word, and the duration of the vowels and syllables. We also measured f0 (in ST) at ten points per word and extracted the number (and position) of creaky voiced syllables and pauses for the whole sentences.

From these points, we calculated f0 trajectories of the whole sentence, contour type on all words in the NP and steepness of the slopes for the rising and falling contours using R scripts [23]. The trajectories were later used for generalized additive mixed models (GAMM - R package mgcv [24]) identifying time-windows of significant differences between conditions. For the calculation of the contour types, the R script took 1ST as the just noticeable difference, classifying data points with less than a 1ST difference as a plateau contour. This procedure led to the closest match between calculated contours and perceived contours as perceived by the first author and a Hungarian native speaker consultant. Words with creaky voiced syllables (16 data points) were manually checked for perceived contour type and excluded from the slope calculations to prevent skewed results based on these extreme measurements.

While there were no significant results for syllable duration and position of f0-minimum and -maximum, the GAMMs of the f0-contour, the distribution of contour types (Chi-Square test) and the slopes of the falling contours (ANOVA) showed significant differences between the three focus domains. There were not enough pauses in the data for a statistical analysis and there was no clear pattern based on voice quality differences alone (see 2.3 for voice quality in combination with slope).

2.3. Results

2.3.1. Prosodic structure of the whole sentence

Figure 1 shows the pitch track of sentence 6 in Adj_F spoken by a female participant. It has the typical and expected prosodic pattern of a narrowly focused sentence in Hungarian, with the main accent on the focus position and deaccenting/reduced stress on the constituents after it (see Section 1). The speaker in this example produced two falling contours on both elements of the NP occupying the focus position (with a rise at the end of the noun marking the NP boundary). Following the NP, the speaker reduced stress on the verb and used creaky voice starting at the verbal particle until the end of the

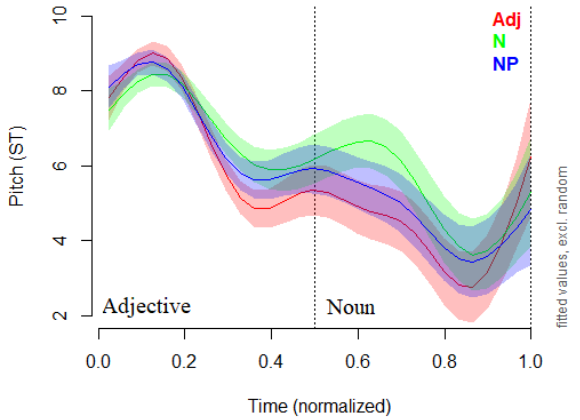

Figure 2: *GAMM model of the three mean-f0 trajectories (in semitones) for adjective, noun and NP focus in the NP*

sentence. All speakers in our data set used this pattern of main accent on the focus position and reduced or erased accents after the focus position. This often led to creaky voice in the post-verbal domain, especially with female speakers (72.1% of sentences vs. 50% with male speakers), presumably further enhancing the perceived difference between prominence on the focus position and post-focal elements. The pattern on the complex NP, however, differed depending on the focus domain and was subject to inter- and intra-speaker variation.

2.3.2. Prosodic structure of the NP

For the mean contours of the NPs, generalized additive mixed models (GAMM) were calculated (see Figure 2). The determiner of the target NP is excluded from the analysis, because it is very short and was often omitted by the speakers. The GAMMs showed a significant influence of FOCUS ($p < 0.001$ for all three conditions) and the pairwise comparisons showed windows of significant differences on the second half of the adjective, and the most part of the noun. In all three focus conditions, the highest mean f0 is on the first syllable of the adjective and does not differ between conditions. On the noun, mean f0 was highest in N_F, lowest in Adj_F, and in between in NP_F.

Table 1 shows the distribution of contour types on the adjective and the noun depending on the focus type. As described in 2.2, contours were automatically calculated by R using 1ST as the threshold between plateau and rising/falling contours. There were significant differences in the distributions of contour types depending on WORD and FOCUS, with more falling contours on the adjective than on the noun and more plateau contours on the noun than on the adjective. Figure 3 shows the correlations between CONTOUR TYPE and FOCUS on each word. For both words, the correlation was significant (Chi-Square test on the adjective: X-squared = 13.438, df = 4, $p < 0.01$; on the noun: X-squared = 20.139, df = 4, $p < 0.001$). On the adjective, there is a correlation between falling contours and rising contours with focus on the adjective and the noun/NP, respectively. On the noun, falling contours correlated with noun focus, and plateau contours with adjective focus.

In contrast to the distribution of contours, the position of f0-min and -max in relation to the vowel, syllable and word was not significant. It is thus unlikely that the timing of the f0-peak plays a role in our data. The height of the f0-max, however, showes significant differences, resulting in varying slopes of the contours (see, e.g., [15] for similar results with simple

Table 1: *Distribution of the contour types on adjective and noun*

Contours on the Adjective				Contours on the Noun			
Focus	Fall	Plateau	Rise	Focus	Fall	Plateau	Rise
Adj	92	3	5	Adj	58	31	11
N	72	9	19	N	82	9	9
NP	79	7	14	NP	76	13	11
Total	81%	6.33%	12.67%	Total	71%	17.67%	10.33%

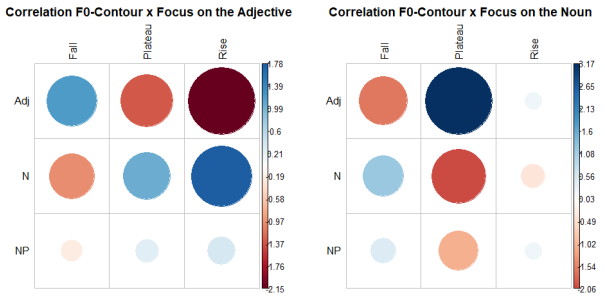

Figure 3: *Correlation plots for the correlations (blue) and anti-correlations (red) between focus condition and contour type on the adjective (left side) and the noun (right side)*

NPs). Because of the relatively low amount of rising contours on both the adjective and the noun, slopes of the rising contours did show a very high amount of variation and are thus not reported. Some speakers displayed a high usage of creaky voice on post-focal nouns and non-initial syllables most likely to further enhance the perceived difference between stressed and unstressed syllables. Even though these are valid tokens of the contour types and thus included in Table 1, they were excluded from the slope calculations (see 2.2).

For the slopes of the falling contours, the ANOVA showed a significant effect of WORD ($df = 1$, $f = 26.18$, $p < 0.001$) and of the interaction of WORD and FOCUS ($df = 2$, $f = 10.08$, $p < 0.001$), while there was no significant effect of FOCUS alone ($df = 2$, $f = 0.22$, $p = 0.8$). Figure 4 shows the slopes of the two words in the different focus domains. The post-hoc Tukey-HSD test showed that in both the Adj_F and the NP_F condition, the slope on the adjective was significantly steeper than on the noun (Adj_F: $p < 0.001$; NP_F: $p < 0.01$). In the N_F condition, there were no differences in the slope of adjective and noun ($p = 0.99$). The slope on the adjective did not differ significantly between the three conditions. This might seem contradictory to the GAMM in Figure 2, where the red line (Adj_F) on the adjective ends lower than the other two conditions. However, this is due to the higher amount of rising and plateau contours in N_F and NP_F that rises the mean f0 in the GAMM at the end of the adjective. There was a significant difference between the slope on the noun in the N_F condition and in the Adj_F condition with it being steeper in the former ($p < 0.01$), and there was a tendency for the slope on the noun to be steeper in the N_F than in the NP_F condition ($p = 0.06$). There was no significant difference in the slopes of the narrow focused elements, i.e. the adjective in Adj_F and the noun in N_F ($p = 0.43$).

3. Discussion

Regarding our two research questions, the experimental data show in response to Q1 that prosody disambiguates different

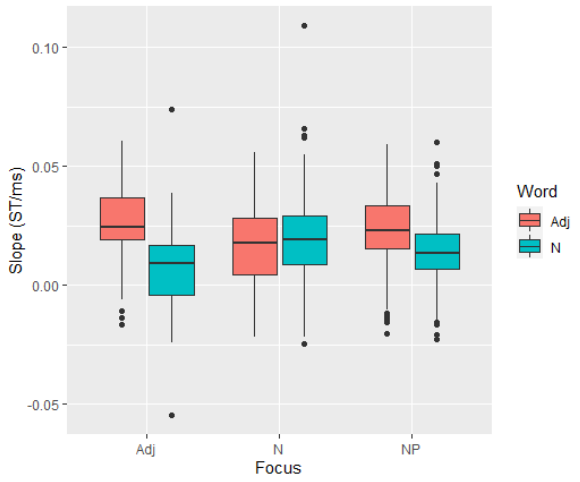

Figure 4: *Slopes of the falling contours on the words in the target NP per focus condition*

Table 2: *Distribution of the contours on the whole NP per focus condition; F = falling, P = plateau, R = rising*

Focus	FF	FP	FR	PF	PP	PR	RF	RP	RR
Adj	54	27	11	2	1	0	2	3	0
N	57	7	8	8	1	0	17	1	1
NP	58	11	10	6	0	1	12	2	0
Total	56.3%	15%	9.6%	5.3%	0.6%	0.3%	10.3%	2%	0.3%

focus positions and domains in complex NPs. In response to Q2, we observe a distinct accentuation strategy in Hungarian NPs that includes prosodic left edge marking of the focus domain and additional prosodic focus marking on the focused constituent if it is not the left edge element.

The results of the f0-analysis show that the highest f0 of the sentences is typically on the leftmost element of the NP occupying the focus position (see Figure 2 and Table 1). This is true for all sentences with a falling or plateau contour on the adjective which are the overwhelming majority in all three focus conditions and across speakers (87,33% of all data points; see Table 1). Thus, we assume that in the default case, Hungarian marks the left edge of the (noun) phrase containing the focus with the highest prosodic prominence (in line with, e.g., [11], [12]). The remaining 12.67% of the sentences seem to show an alternative pattern with a rising contour on the adjective that significantly correlates with conditions where the noun is part of the focus domain (N_F or NP_F). Table 2 shows that in these cases, the adjective is typically followed by a falling contour on the noun[3]. However, there are only five speakers in our data set that use this pattern more than once and half of the speakers do not use it at all. We will thus ascribe this to inter- and intra-speaker variation and concentrate on the default case.

The pitch of the noun depends on the position and domain of the focus, with following f0-peaks being generally lower than preceding ones (see Figure 2). This correlates with a higher number of plateau and rising contours on the noun (see Table 1) where the phonetically rising contours are a product of a low first syllable and a high boundary tone at the right edge of the NP. If the focus is only on one element of the phrase, the f0 of this element is boosted (see, e.g. [25] for 'boosting by focus'), i.e., it is significantly higher and followed by a steeper fall (see also [15]) than a non-focused element (see e.g. Figure 2 the green line (N_F) on the noun vs. blue and red). Post-focal elements often have a plateau or rising contour or at least a significantly lower f0-peak than elements that are part of a (larger) focus domain.

We analyse the lower f0-peaks on the noun as downstepped accents. Based on the differences in contour distribution and the slopes of falling contours, we argue for a three-way distinction of accent types: (a) a neutral pitch accent on an element that is part of a broad(er) focus domain (identical to pitch accents in a 'neutral' sentence; see Section 1), (b) an enhanced accent that is boosted by focus, either because it is a phrase accent on the left-most element of the phrase containing the focus (in our case: the adjective of the NP) or because of narrow focus on the element itself (see also [15]), and (c) a reduced accent in post-focal position. This would explain all significant differences (and lacks thereoff) in the slopes (see e.g. Figure 4). In the Adj_F and the NP_F condition, there are significant differences between the enhanced accent on the adjective and the normal (NP_F) or reduced (Adj_F) accent on the noun, while there are no significant differences in N_F because both words bear an enhanced accent: the adjective, because of its phrase initial position, and the noun, because of the narrow focus. Between sentences of the different focus conditions, there are no significant differences between the accents on the adjective: it is always an enhanced accent. The only difference can be seen on the noun. The slope of the enhanced accent in N_F is significantly steeper than the slope of the reduced accent in Adj_F. The slope of the normal accent in NP_F however, does not differ significantly from either of the two because it is acoustically between them, leading to our analysis as a three way distinction between enhanced, normal, and reduced accentuation. If these subtle differences can be used in perception is a task for further research.

4. Conclusions

As expected, the highest f0 within the sentence typically falls on the leftmost element of the focus position, prosodically marking the edge of the focused NP. This is in line with various analyses of Hungarian prosody, such as for example [26]'s approach, in which the focused element moves to the first position of the intonational phrase to recieve the main accent that is always aligned with the left edge of this IP. However, leftmost accentuation cannot explain the prosodic variation found inside the syntactically focused NP. We argue in line with [15] that, despite the structurally prominent syntactic focus position in Hungarian, prosodic prominence marking plays an important, independent role in focus marking. Prosodically, Hungarian represents a further case of the NP accentuation strategies mentioned in Section 1 combining left-alignment of prosodic prominence with the phrase (iii) and additional marking of the focused element (i) by a downstepped accent. Hence, prosodic prominence marking adds to syntactic focus marking in Hungarian.

5. Acknowledgements

We are grateful to our student assistant Noémi Ecsedi for her help with creating stimuli, recording and data processing. This work was supported by the DFG (KU2323/4-1). Many thanks to the two anonymous reviwers and the audience of the 2021 'Tone and Intonation' for discussion and feedback.

[3]This pattern could be analysed as accentuation strategy (i) where (only) the focused elements of an NP are accentuated.

6. References

[1] W. Chafe, "Language and consciousness," *Language*, vol. 50.1, 1974.

[2] K. É. Kiss, "Topic and focus: Two structural positions associated with logical functions in the left periphery of the Hungarian sentence," *Acta Linguistica Hungarica*, vol. 55.3-4, pp. 287–296, 2008.

[3] L. Hunyadi, "Remarks on the syntax and semantics of topic and focus in Hungarian," *Acta Linguistica Academiae Scientiarum Hungaricae*, vol. 31.1/4, pp. 107–136, 1981.

[4] M. Krifka, "Basic notions of information structure," *Acta Linguistica Hungarica*, vol. 55.3-4, p. 243–276, 2008.

[5] M. Rooth, "Association with focus," Ph.D. dissertation, Amherst: University of Massachusetts dissertation, 1985.

[6] M. Zimmermann and C. Féry, Eds., *Information structure. Theoretical, typological, and experimental perspectives*. Oxford: Oxford University Press (Oxford linguistics), 2010.

[7] F. Kügler and S. Calhoun, "Prosodic Encoding of Information Structure. A typological perspective," in *The Oxford Handbook of Language Prosody*, C. Gussenhoven and A. Chen, Eds. Oxford: Oxford University Press, 2020, pp. 453–467.

[8] D. R. Ladd, *The Structure of Intonational Meaning: Evidence from English*. Bloomington: Indiana University Press, 1980.

[9] C. Roberts, "Information structure in discourse: Toward a unified theory of formal pragmatics," *Ohio State University Working Papers in Linguistics*, vol. 49, pp. 91–136, 1996.

[10] D. I. Beaver and B. Z. Clark, *Sense and Sensitivity: How Focus Determines Meaning*. John Wiley & Sons, 2009.

[11] L. Varga, *Intonation and Stress: Evidence from Hungarian*. New York: Palgrave Macmillan, 2002.

[12] G. Olaszy, "A beszéd szupraszegmentális szerkezete [Suprasegmental structure of speech]," in *A magyarbeszéd: Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek. [Hungarian speech: Speech research, speech technology and systems]*, G. Olaszy and G. Németh, Eds. Akadémiai Kiadó, 2010.

[13] K. É. Kiss, *Discourse Configurational Languages*. New York/Oxford: Oxford University Press, 1995.

[14] I. Kenesei, "Adjuncts and Arguments in VP-focus in Hungarian," *Acta Linguistica Hungarica*, vol. 45.1-2, p. 61–88, 1998.

[15] S. Genzel, I. Shinichiro, and B. Surányi, "The prosodic expression of focus, contrast and givenness: A production study of Hungarian," *Lingua*, vol. 165, p. 183–204, 2015.

[16] M. Swerts, E. Krahmer, and C. Avesani, "Prosodic marking of information status in Dutch and Italian: A comparative analysis," *Journal of Phonetics*, vol. 30.4, p. 629–654, 2002.

[17] C. Féry, "Focus as prosodic alignment," *Natural Language & Linguistic Theory*, vol. 31.3, p. 683–734, 2013.

[18] M. Grice and F. Kügler, "Prosodic Prominence - A Cross-Linguistic Perspective," *Language and Speech*, vol. 64.2, p. 253–260, 2021.

[19] T. Luchkina and J. Cole, "Perception of word-level prominence in free word order language discourse," *Language and Speech*, vol. 64.2, p. 381–412, 2021.

[20] B. Surányi, "Discourse-configurationality," in *The Oxford Handbook of Information Structure*, C. Féry and S. Ishihara, Eds. Oxford: Oxford University Press, 2015, p. 422–440.

[21] P. Boersma and D. Weenink, "Praat: doing phonetics by computer," URL: http://www.praat.org/, 2017.

[22] G. Fant, A. Kruckenberg, K. Gustafson, and J. Liljencrants, "A new approach to intonation analysis and synthesis of Swedish," *TMH-QPSR 44*, p. 161–164, 2002.

[23] R Core Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, URL: https://www.R-project.org/, 2017.

[24] S. Wood, "Generalized Additive Models: An Introduction with R," Chapman and Hall/CRC, 2017, 2. edition.

[25] C. Féry and I. Shinichiro, "How focus and givenness shape prosody," in *Information structure. Theoretical, typological, and experimental perspectives*, M. Zimmermann and C. Féry, Eds. Oxford: Oxford University Press (Oxford linguistics), 2010, pp. 36–65.

[26] K. Szendrői, "A stress-based approach to the syntax of Hungarian focus," *The Linguistic Review*, vol. 20.1, 2003.