Prosodic (non-)realisation of broad, narrow and contrastive focus in Hungarian: a production and a perception study

Katalin Mády

Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, Hungary

mady@nytud.hu

Abstract

In languages with variable focus positions, prominent elements tend to be emphasised by prosodic cues (e.g. English). If a language prefers a given prosodic pattern, i.e. sentence-final nuclear accents, like Spanish, the prosodic realisation of broad focus might not differ from that of narrow and contrastive focus. The relevance of prosodic focus marking was tested in Hungarian were focus typically appears in front of the finite verb. Prosodic cues such as f0 maximum, f0 peak alignment, segment duration and post-verbal deaccentuation were tested in an experiment with read question and answer sequences. While narrow and contrastive focus triggered post-verbal deaccentuation, none of the gradual measures distinguished focus types consistently from each other. A subsequent perception experiment was conducted in which the same sentences without post-verbal units were to be judged for their naturalness. F0 maximum, f0 peak alignment and accent duration were manipulated. Naturalness scores revealed a sequence narrow > contrastive > broad focus, i.e. a preference for narrow focus contexts compared to contrastive and broad focus ones, while the manipulated prosodic parameters had no effect on the scores. It is concluded that prosodic focus marking in Hungarian is optional and pragmatic rather than grammatical and syntax-related.

Index Terms: prosody, focus, information structure, Hungarian

1. Introduction

Information that is relevant for communication is emphasised in some way in most languages. Prominence can be marked by prosody, syntax, morphology, or by their combination. The flexibility of word order and the variability of prosodic patterns is interrelated: languages with a relatively rigid word order like English tend to use prosodic prominence marking, whereas languages with a preference for a certain prosodic pattern such as Spanish [1] or Czech [2] move the focussed element to the position that bears strong prominence by default. While the default nuclear accent position is sentence-final both in English and most Romance languages, the usual way to express emphasis in English would be to shift the nuclear accent if the prominent element is non-final, whereas in Spanish the prominent unit would be shifted to the sentence-final position. Languages that prefer prosodic prominence marking are referred to as prosodically 'plastic', since they show variable prosodic patterns, whereas languages that tend to maintain a certain prosodic pattern but allow for syntactic flexibility are referred to as prosodically 'non-plastic' languages [3, 4].

Prosodic plasticity is gradual with respect to the rigidity of the maintenance of word order vs. the prosodic pattern. For example, [1] showed that word order shift is obligatory in Spanish or Catalan, whereas the tendency is weaker in languages with an intermediate status such as Neapolitan Italian. In Spanish that is prosodically non-plastic, prosodic focus marking is optional, whereas it is obligatory in Neapolitan Italian, both in the default (sentence-final) and the non-default focus position.

At the same time, prosodic focus marking in prosodically plastic languages is not necessarily consistent. According to a study by Baumann et al. [5] in which they studied focus realisations in German, speakers applied different strategies to mark the degree of prominence: they used a combination of parametric means such as lengthening, f0 range and maximum and its position and of categorical means such as different pitch accents. In their study, only one out of 6 speakers made use of all available strategies.

Similarly to Spanish, Hungarian is a prosodically non-plastic language with syntactical focus marking, and it is prosodically left-headed [4]. Word order reflects a logical structure: according to [6], Hungarian sentences can be divided into a topic and a predicate part. The topic position includes units about which a prediction is made in the predicate part of the sentence. É. Kiss claims that the first major constituent of the predicate has to be accented, and it bears the heaviest accent in the sentence. According to her view, the nuclear accent is located at the left edge in Hungarian, but its domain is the predicate and not the entire sentence. As a consequence, the topic has to be treated as a separate intonational phrase (IP) as was suggested by [7, 8, 9]. However, this view is problematic since an IP boundary is not necessarily present in each sentence which contains an accented topic and a predicate.

In Hungarian, foci are located immediately before the finite verb. If the sentence has broad focus, word order is relatively free, also with respect to the pre-verbal position. It is assumed that focus is mainly expressed by the obligatory deaccentuation of the finite verb and of post-verbal elements [10].

The goal of the present paper is to investigate the role of prosodic cues in focus marking in Hungarian. Given the strict syntactic marking of foci, it is expected that prosodic cues are not necessary to express focus prominence. First, a production experiment is presented (see [11] for an outline in Hungarian) which serves as a basis for a perception study. The research questions are the following: (1) is focus marking supported by prosodic cues in Hungarian? (2) Are there differences in the distribution of pitch accents? (3) Do certain pitch accent types contribute to the distinction? The study includes the investigation of the focus, the topic, the verb, and the post-verbal part of the comment. Three focus types are compared: sentences with broad focus are all-new sentences, i.e. not a single constituent, but the entire sentence is focussed. The term narrow focus is used for discourse-new constituents that answer a *wh*-question, while contrastive focus expresses explicit contrast.

2. Production experiment

Two target sentences were created with a so-called verbal modifier (VM). This unit appears immediately before the finite verb even if it is not focussed, thus it is possible to use string-identical sentences with broad, narrow and contrastive focus depending on the context.

Sentence 1:

| *A lányom* | *Németországba* | *ment* | *munkát keresni.* |
| the daughter-my | Germany-to | went | job-acc search-to |

'My daughter went to Germany to look for a job.'

Sentence 2:

| *Marianna* | *Máltán* | *nyaral* | *Mónival.* |
| Marianna | Malta-on | has-holiday | Móni-with |

'Marianna is on holiday in Malta with Móni.'

Since lexical stress is word-initial in Hungarian, the first syllable of each content word was the potential domain for accentuation. Sentences were displayed to participants on a screen using SpeechRecorder [12] preceded by a question that elicits broad, narrow or contrastive focus. Speakers were asked to read the target sentence as an answer to the following questions:

1. **broad focus:** (1): *How is it going? How is Erika?* (2): *How are you (plural)? And how is your colleague?*

2. **narrow focus:** (1): *So where is Erika looking for a job?* (2): *So where did your colleague and her friend decide to go for holiday?*

3. **contrastive focus:** (1): *Erika is looking for a job in Pest, right?* (2): *Your colleague and her friend went to Sweden in the end, didn't they?*

7 students (6 females, 1 male, mean age 21 years) participated in the experiment that was carried out in a sound-treated room at the Department of Phonetics, RIL HAS. Each speaker read each target sentence in three focus conditions with 6 repetitions in randomised order, totalling in a stimulus set with 252 items. The contexts were merged with the stimuli for another experiment serving as distractors here.

The following parameters were analysed in the pre-verbal part of the sentence, i.e. on the topic and the focus:

Categorical cues:

- occurrence of deaccented content words,
- distribution of pitch accent types.

Parametric cues within word-initial CVC sequences (irrespective of the syllable structure):

- f0 maximum,
- f0 minimum,
- f0 range,
- f0 slope,
- duration of the f0 rise or fall,
- alignment of the target tone within the CVC sequence,
- duration of the CVC sequence.

The analysis of the post-focal part was restricted to the categorical cues listed above.

Pitch accents were labelled by two labellers using tonal labels such as L*, H*, and their combinations. F0 was measured in semitones with a reference value of 100 Hz. The duration of the accented syllable was normalised to the entire sentence length.

Categorical cues were analysed by means of χ^2 tests. For parametric cues, generalised linear mixed models (`lme4` package in R) were used with focus type as a fixed effect and speaker, sentence and repetition as random effects. Significance level was set to $\alpha = 0.05$.

2.1. Results

2.1.1. Categorical cues

Topics were deaccented in 71% of all cases, and if a pitch accent was present, it was usually rising (71%). According to a pairwise comparison, the distribution of pitch accent categories showed no effect of focus type ($p > 0.2$). Foci were accented in all but 2 cases and carried mostly falling accents (H+L*: 82%, H*+L: 8%, L*: 6%). Their distribution did not differ significantly across focus types ($p > 0.1$). The verb was deaccented in all but 6 cases (93%). post-verbal units were deaccented in about 50% of the cases in the narrow and contrastive focus condition and in 32% in the broad focus condition. If there was an accent, it usually had an H+L* pattern.

These results show a uniform prosodic pattern for all sentences. Pitch accents did not differ according to focus type. The only difference was that the deaccentuation of post-verbal elements was more frequent after narrow and contrastive focus. This shows that [10]'s claim about the obligatory accent deletion after a focus can be observed as a tendency, but not as a strict rule.

2.1.2. Parametric cues

Separate analyses were run for the tonal cues of the topic, focus and the post-focal units. Only comparisons with a significant difference are reported here.

The f0 minimum of the topic was aligned earlier and the maximum aligned later before contrastive focus. These CVC sequences were also significantly longer than those with broad and narrow focus, thus the difference in tonal alignment might be a result of lengthening. The f0 peak on the focus was aligned later if the focus was contrastive, but no lengthening of the accented CVC sequence was observed. Fig. 1 shows the f0 maximum alignment for each speaker where the percentage value specifies the point at which f0 starts to fall within the focus-initial CVC (a higher percentage means a later maximum alignment). Interestingly, the alignment is later in contrastive focus than in broad focus, while the marking of narrow focus is highly speaker-dependent: alignment is earlier than in broad focus in some speakers' production and later than in contrastive focus in other speakers' data.

The results show that broad, narrow, and contrastive focus were not distinguished consistently. Broad focus differed from narrow and contrastive focus by the more frequent accentuation of the post-verbal constituent. Neither pitch accent categories nor the occurrence of deaccentuation of the topic revealed a significant difference between focus types. There were two parameters that distinguished contrastive focus from the other two: they were preceded by longer topic CVC duration and a less steeper f0 rise on the accented syllable, and the slope for the fall on the focus started later leading to a longer plateau between the topic and the focus. Again, there was a large inter-speaker variation.

In the second experiment it was investigated whether tonal cues can enhance focus identification. Tonal alignment was tested along with two parameters that had been found to trigger the perception of emphasis in several languages.
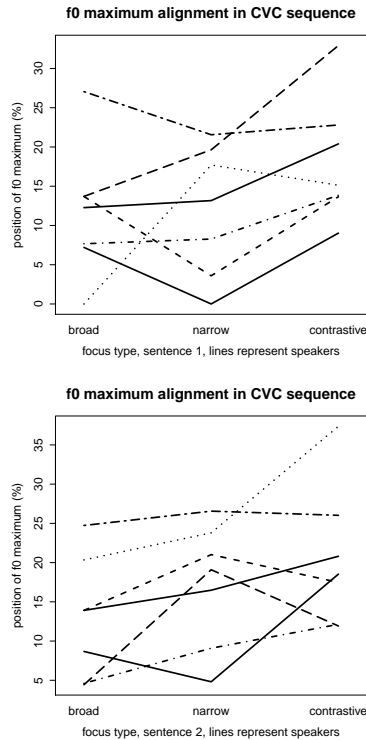
f0 maximum alignment in CVC sequence

f0 maximum alignment in CVC sequence

Figure 1: *Position of the f0 maximum within the accented CVC sequence of the focus. 0%: beginning of CVC duration, 100%: end of CVC duration. Top: sentence 1, bottom: sentence 2.*

3. Perception experiment

3.1. Material and methods

Similarly to the production experiment, target sentences were presented as question and answer pairs in which the verbal modifier could be interpreted as broad, narrow or contrastive focus. The target sentences differed slightly from those in the production experiment because they did not involve any post-verbal units, in order to avoid an influence of their (de)accentuation. Thus, sentences were *A lányom Németországba ment* 'My daughter went to Germany', and *Marianna Máltán nyaral* 'Marianna is on holiday in Malta'. The preceding questions were adjusted accordingly.

Question and answer pairs were recorded with a female speaker five times in the broad, narrow and contrastive focus setting. The parameters for manipulation were the following: the f0 maximum in the initial (= accented) CVC sequence of the VM, the localisation of the f0 maximum, and the duration of the accented vowel.

One realisation of each sentence with narrow focus was selected as the basis for manipulation. Both sentences had a rising accent on the topic and a falling accent on the focus, as was observed in the production data.

The parameters to be manipulated were analysed in the data of the model speaker, and the measured minimum and maximum values were used for manipulation. Since the accented CVC sequences in the two sentences were different (/neːm/ vs. /maːl/), the edge values were adapted to the sentence-specific realisations. The following parameters were manipulated using Praat's PSOLA synthesis module:

- the value of f0 maximum on the accented syllable (sentence 1: 250 and 270 Hz, sentence 2: 240 and 260 Hz), the f0 minimum being constant through all sentences (sentence 1: 185 Hz, sentence 2: 150 Hz),

- the alignment of the f0 peak, i.e. the time point where the pitch fall in the H+L* accent started (20 ms before and after the onset of the accented vowel in both sentences),

- the duration of the accented vowel (sentence 1: 85 and 125 ms, sentence 2: 145 and 185 ms).

Thus, a set of $2 \times 2 \times 2 = 8$ combinations was created for each sentence. Stimuli were presented according to the following design: each manipulated sentence was preceded by a question that elicited either broad, narrow or contrastive focus. Participants heard a question and answer pair, and they were asked to rate how natural the answer sounded to the question, using a discrete scale between 1 and 7. In order to exclude the influence of potential other factors such as the quality of the synthesis or semantic unnaturalness, participants were instructed to pay attention to accent strength that varied across sentences. Each combination (3 questions $\times$ 2 f0 maxima $\times$ 2 f0 slope alignments $\times$ 2 vowel durations $\times$ 2 sentences) was played twice to each participant, totalling in a set of 96 question and answer pairs. The same amount of distractors was added to the stimulus material. 28 listeners participated in the experiment (18 males, 10 females, mean age 31 years).

Analysis was performed using generalised linear mixed models with naturalness scores as the dependent variable, focus type, f0 maximum, f0 slope alignment and vowel duration as fixed effects and speaker, sentence and repetition as random effects. Interactions were tested by the `Anova` function provided in the `car` package in R. Significance level was set to $\alpha = 0.05$.

3.2. Results

Repetition was not significant as a random effect and was thus ignored in further analysis. The effect of focus type was highly significant ($p < 0.001$). Participants judged answers to questions that elicited narrow focus as most natural (mean score: 5.2, median: 6), and broad focus as least natural (mean: 4.17, median: 4), while contrastive focus answers received intermediate judgements (mean: 4.82, median: 5). Means for each participant are shown in Fig. 2.

The preference for the narrow focus context can have two reasons: first, the original utterances that underwent manipulation were both answers to a *wh*-question that elicited a narrow focus reply. It is possible that there were parameters other than the analysed ones that crucially influenced the choice between focus types. The other explanation is a pragmatic one: an utterance that contains discourse-new information such as the answer to a *wh*-question "Where did your colleague go on holiday?" is a natural way of information seeking in everyday communication, whereas the same reply to an all-new question like "How is your colleague?" might sound less natural. Again, explicit contrast is relatively rare in spontaneous communication. The reason for the preference of narrow focus answers is currently being tested in a follow-up experiment.

Subsequently, the effect of the tonal manipulation was investigated within each focus type. In the broad focus condition, none of the fixed effects reached significance level ($p > 0.2$). In the narrow focus condition, targets with a higher f0 maximum triggered significantly higher naturalness scores ($p = 0.02$).
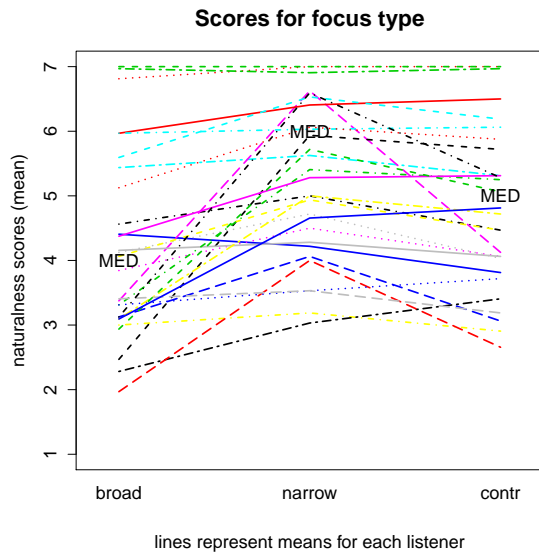
Scores for focus type

Scores for narrow focus

Figure 2: *Means of naturalness scores for the broad, narrow and contrastive focus conditions for each participant. MED refers to the overall medians.*

Figure 3: *Means of naturalness scores for high and low f0 maxima in the narrow focus condition for each participant. MED refers to the overall medians.*

However, there was a large variation among participants: sentences with high f0 received higher scores from 17 listeners ($= 0.46$, $sd = 0.43$) and lower scores from 8 participants ($mean = -0.37$, $sd = 0.31$), while responses of 3 participants did not differ from each other (see Fig. 3). If there was a bias in the preference, it was not larger than 0.9 scores, apart from one participant (1.5). In other words, narrow focus replies were judged as relatively natural almost independently of tonal manipulations. In the contrastive focus condition, there was a slight increase of naturalness scores if f0 maximum was high and the vowel duration was long, signalised by an interaction between f0 maximum and vowel duration $p = 0.09$. No other effect was significant.

4. Discussion

In terms of information structure, both narrow and contrastive focus bear emphasis. Contrastive focus is explicitly contrastive, while narrow focus carries implicit contrast, and both are assumed to be exhaustive in Hungarian. It is expected that if prosody plays a role in focus marking, then the difference between broad focus and the other two types will be greater, since broad focus does not express contrast. However, the results of the production experiment do not support this assumption. Apart from a strong inter-speaker variation, only contrastive focus is marked by tonal alignment, whereas narrow focus is less marked than broad focus in many speakers' production.

The perception experiment reveals that later tonal alignment does not enhance contrastive interpretation: samples with a later f0 fall did not get higher naturalness scores than those with an earlier one. At the same time, higher f0 lead to higher naturalness scores for narrow, but not for contrastive focus, whereas a corresponding tendency was not present in the production experiment.

A recent study by [13] which is under publication leads to contradicting conclusions. In this study, focus types along with the givenness of the background information were investigated.
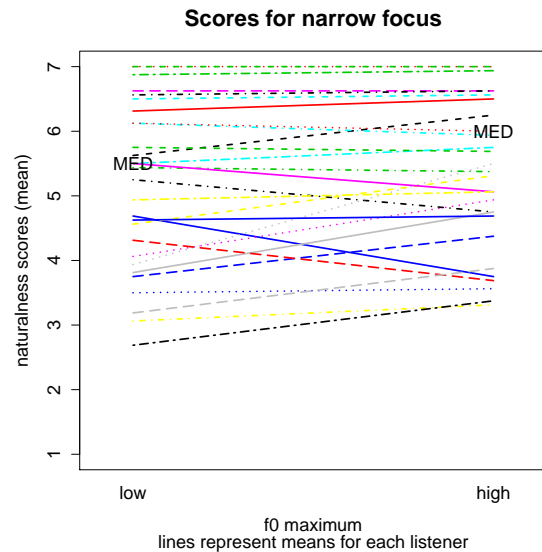
Here, a significant increase of f0 about 1 semitone was found in focussed syllables. Besides, more topics with rising pitch accents were found when followed by contrastive foci. This seems surprising at the first glance, because in a study on contrastive topics in Hungarian, [14] found that rising signalled the contrastivity of the topic. However, based on the demonstration of rising topics in the paper of [13] it is to be assumed that this kind of topic realisation would rather be categorised as deaccented according to our guidelines. Thus, the deaccentuation of the topic might trigger a stronger prominence of the following focus. However, this tendency is not in line with our data where topic deaccentuation was not more frequent before narrow and contrastive focus than before broad focus.

The results of both experiments show weak evidence for the relevance of tonal cues for focus marking in Hungarian. There was an asymmetry between the results in production and perception: while only one of the (many) investigated parameters of the focus showed that contrastive foci are marked tonally, this effect could not be replicated in the perception experiment. At the same time, higher f0 maximum that lead to higher naturalness scores for narrow foci was not present in the production data. Thus it is concluded that prosodic focus marking in Hungarian is optional and is not utilised systematically. Since no evidence for a consistent distinction between broad focus on the one hand and narrow and contrastive focus on the other was found, prosodic cues for foci do not seem to reflect grammatical categories or information structure in the pre-verbal position in Hungarian.

5. Acknowledgements

6. References

[1] T. L. Face and M. D'Imperio, "Reconsidering a focal typology: evidence from Spanish and Italian," *Italian Journal of Linguistics*, vol. 17, no. 2, pp. 271–289, 2005.

[2] T. Duběda and K. Mády, "Nucleus position within the intonation phrase: a typological study of English, Czech and Hungarian," in *Proc. Interspeech*, Makuhari, Japan, 2010, pp. 126–129.

[3] E. Vallduví, "The role of plasticity in the association of focus and prominence," in *Proc. ESCOL 7*, 1991, pp. 295–306.

[4] D. R. Ladd, *Intonational phonology, 2nd ed.* Cambridge: Cambridge University Press, 2008.

[5] S. Baumann, M. Grice, and S. Steindamm, "Prosodic marking of focus domains – categorical or gradient?" in *Proc. 3rd Speech Prosody, Dresden, Germany*, 2006, pp. 301–304.

[6] K. É. Kiss, *The syntax of Hungarian*. Cambridge: Cambridge University Press, 2002.

[7] I. Vogel and I. Kenesei, "The interface betwen phonology and other components of grammar: the case of Hungarian," *Phonology Yearbook*, vol. 4, pp. 243–263, 1987.

[8] L. Varga, *Intonation and stress: evidence from Hungarian*. Basingstoke & New York: Palgrave Macmillan, 2002.

[9] L. Hunyadi, *Hungarian sentence prosody and universal grammar: on the phonology–syntax interface*. Frankfurt/Main: Lang, 2002.

[10] P. Siptár and M. Törkenczy, *The Phonology of Hungarian*. Oxford: University Press, 2000.

[11] K. Mády, "A fókusz prozódiai jelölése felolvasásban és spontán beszédben [Prosodic marking of focus in read and spontaneous speech]," in *Beszéd, adatbázis, kutatások*, M. Gósy, Ed. Budapest: Akadémiai Kiadó, 2012, pp. 91–107.

[12] C. Draxler and K. Jänsch, "SpeechRecorder – a universal platform independent multi-channel audio recording software," in *Proc. International Conference on Language Resources and Evaluation*, Lisbon, 2004, pp. 559–562.

[13] S. Genzel, S. Ishihara, and B. Surányi, "The prosodic expression of focus, contrast and givenness: A production study of Hungarian," *Lingua*, in press, http://dx.doi.org/10.1016/j.lingua.2014.07.010.

[14] B. Gyuris and K. Mády, "Contrastive topics between syntax and pragmatics in Hungarian: an experimental analysis," *Papers of Chicago Linguistic Society*, p. 15 p., 2014.