



## Effects of L1 prosodic structure on narrow focus realizations in an L2: Evidence from Hungarian learners of German

*Susanne Beinrucker<sup>1</sup>, Felicitas Kleber<sup>1</sup>, Katalin Mády<sup>2</sup>*

Ludwig-Maximilians-Universität München, Institute of Phonetics and Speech Processing, Germany<sup>1</sup>  
Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, Hungary<sup>2</sup>

susanne.beinrucker@campus.lmu.de, kleber@phonetik.uni-muenchen.de, mady@nytud.hu

### Abstract

This study investigates the realization of prosodic structure by L2 learners assuming that prosodic features of the L1 are transferred into the L2. Hungarian and German are prosodically diverse, differing both on the word stress and sentence accent level. While Hungarian has (1) fixed word stress on the initial syllable and (2) a syntactically fixed focus position (before the finite verb), German has variable word stress, and words in any sentence position can act as a narrow focus (i.e. without a shift in word order). Additionally, narrow focus is typically produced with rising accents in German while they tend to be realized with falling accents in Hungarian. Five Hungarian learners of German and five German control speakers read various repetitions of two German target words differing in word stress (initial vs. medial) that occurred either before the finite verb or after it. Hungarians produced more falling accents than German speakers and tended to produce non-focused elements with strong prominence when they appeared before the finite verb. Word stress errors occurred more often in tokens with stress on the first syllable – presumably due to overgeneralization. Differences in pitch accents are discussed with respect to language acquisition models.

**Index Terms:** foreign language acquisition, prosody, focus, stress, German, Hungarian

### 1. Introduction

The aim of this study was to investigate the extent to which the prosodic structure of the first language (L1) transfers to the prosodic realizations of utterances in a second language (L2). The overall aim is to relate the findings to theories of second language acquisition and to the prosodic structure of a learner's L1. To this end we analyzed lexical stress and narrow focus in two sentence positions in Hungarian learners of German because the two languages differ both on the lexical stress and sentence accent level.

In Hungarian, lexical stress is fixed to the first syllable of a word [1]; German, on the other hand, has variable word stress, but with a tendency to occur on the penultimate syllable [2]. Evidence shows that Hungarians and listeners of other languages with fixed lexical stress tend to be stress 'deaf', i.e. they have difficulties to discriminate words differing in lexical stress [3]. Thus, one of the aims of this study was to test how well stress 'deaf' speakers realize lexical stress in unfamiliar position in a foreign language with variable stress.

German and Hungarian also differ in the way with which focus is realized on the phrase level. Both languages use focus, which is determined by the context, to emphasize the semantics of a word and to amplify the pragmatics of an utterance. To do so Hungarian again uses a fixed position in word order, i.e., focus is predominantly defined as a syntactic category. The word that is to be in focus needs to be placed immediately in front of the finite verb [4]. This positional constraint is related to the claim that Hungarian (again in contrast to German) is considered a left-headed language [5]. In German, focus can be placed anywhere in the sentence [6] and the words that act as focus are marked as such by pitch accents. L2 learners need to acquire the appropriate use of prosodic parameters (e.g. pitch accents) of the foreign language just as they have to learn the segmental structure. In order to minimize foreign accent and to communicate the intended pragmatics of an utterance correctly, L2 speakers must learn to produce the correct pitch accents and to place them as required by the context. [7] observed that Hungarians tended to produce at least one additional accent on a word closer to the left phrase boundary when narrow focus should be produced on the verb in sentence final position. This accords with broad focus in German. Thus, a second goal was to test whether Hungarian learners of German tend to realize sentences in which the focus occurs on a word after the finite verb with broad focus, i.e., with prosodic focus on the verb before and after the finite verb.

German speakers realize prosodic focus predominantly with pitch accents in which  $f_0$  rises throughout the stressed vowel resulting in a late alignment of the  $f_0$  maximum (i.e., L\*+H according to the GToBI system [8]). Focused elements may also be realized with the default H\* or L+H\* pattern, in which the  $f_0$  peak is aligned with the stressed vowel and which signals new information [9]. Falling accents (i.e., H+L\*), on the other hand, should occur rarely on words in focus position given that they signal old information in German [9]. Despite their syntactic marking, words in focus position are also accompanied by pitch accents in Hungarian. In contrast to German, however, these are usually falling accents [10, 11], although the amount of speaker variability is very high, possibly because the focused unit is sufficiently marked by the syntactically defined preverbal position [11]. Moreover, the falling accents in Hungarian are characterized by a steep fall from the high pitch accent to a low tone (possibly a trailing tone), whereas in German the fall is the result of an interpolation between a high pitch accent and a low phrase accent [12]. A third goal of this study, therefore, was to

analyze aspects of the form of the pitch accents on words in focus position.

Because of the prosodic differences between Hungarian and German outlined above (variable vs. fixed word stress, variable vs. fixed focus, falling vs. rising focus accents), we expect interferences between the two languages when Hungarians learn German as a foreign language.

The majority of theories on second language learning focus on the acquisition of segmental features but in principal their predictions should also hold for prosodic features. According to the Speech Learning Model (SLM) [13], phonetic differences between the L1 and L2 phonemes are important for the ability to perceive them, i.e., the more different a phoneme is, the more likely it will be established as a new phoneme category. Whenever the L1 and the L2 phonemes are too similar, they should not be perceived as two different sounds and – according to the Perceptual Assimilation Model (PAM) [14] – learners should assimilate these phonemes. With respect to the present investigation, the question arises as to whether the presence of lexical stress in different positions and the (form of the falling) pitch accents in Hungarian and German (which are phonemic in the sense that they carry a functional load and that they are semantically and pragmatically contrastive) are such cases of prosodic similarity which are not clearly perceptible. More specifically, we were interested whether the L2 speakers orientate on their L1 and assimilate to it or whether they establish new prosodic categories for the L2 when these are too different from their L1.

## 2. Hypotheses

Three main hypotheses were investigated:

- H1. *Word stress*: Hungarians deviate more often from the correct lexical stress pattern when it diverges from their native pattern. That is, more errors are expected in German words with medial stress.
- H2. *Narrow vs. broad focus*: Hungarians produce less narrow focus in medial sentence position after the finite verb. In these instances they realize an additional pitch accent on the word preceding the finite verb.
- H3. *Pitch accent*: Hungarian learners of German transfer their native accent patterns into the foreign language by producing more falling accents on focused constituents than Germans. In addition, these accents are realized with a steep fall.

## 3. Method

In order to test the hypotheses German native speakers and Hungarian learners of German were compared in a production experiment.

### 3.1. Materials

The stimuli were short German verb-second clauses with one of two target words (/ˈhɪm.bɛ.rən/, ‘raspberries’, /meˈlɔ.nən/, ‘melons’) which acted always as narrow focus but which occurred in either one of two positions: sentence initial before the finite verb or sentence medial between the finite verb and the non-finite verb. Lexical stress was either on the initial syllable (*Himbeeren*) or on the penultimate (*Melonen*). To elicitate narrow focus on the target word, we embedded these sentences into a short dialogue. In this dialogue, the stimulus sentence was always the answer to the question “*Was hat*

*{Melanie | Verena} gegessen?*” (“What did {Melanie | Verena} eat?”), which triggers narrow focus on the target word in the answer. The name of the acting person varied to render the dialogues more variable. The two answers were either (1) “*{Melanie | Verena} hat {Himbeeren | Melonen} gegessen*” (“{Melanie | Verena} ate {Himbeeren | Melonen}”) or (2) “*{Himbeeren | Melonen} hat {Melanie | Verena} gegessen*” (fig. “*It was {Himbeeren | Melonen} that {Melanie | Verena} ate*”). The sentence structure of the first sentence is the default and unmarked structure in German. The sentence structure with the object in initial position (which enhances the emphasis on the word in focus) conforms to the Hungarian sentence structure where the focus position is before the finite verb. Thus, one set of stimuli consisted of 8 tokens (2 names x 2 fruits x 2 positions). The stimuli were repeated three times (24 tokens per subject) and in total there were 240 recordings.

### 3.2. Recordings

The subjects were instructed in their L1. The Hungarian speakers were given a short explanation of how prosodic focus functions in German. To clarify this further one complete set of the stimulus dialogues (8 per subject) was presented both auditorily (using a prepared recording of the dialogue by two phonetically trained L1-German speakers) and orthographically. In this set the subjects were firstly asked to listen to the short dialogue and then read out loud the answer to the question that was played a second time. Subsequent to this training block, two repetitions of each stimulus set, which were then only orthographically presented in randomized order, were recorded with a sampling rate of 44.1 kHz using the SpeechRecorder software [15].

### 3.3. Speakers

Five native German speakers (age: Ø 25,6 years; two male speakers) and five female Hungarian learners of German (age: Ø 26,0 years) were recorded. The Hungarian speakers had learnt German in school. The mean duration of learning German was eight years with a range between four and 13 years. One additional condition for the Hungarian learners of German was that they should not have had a longer residence in a German speaking country. Three of them had previously stayed for a few weeks in one of those countries. Also three of them stated that they had got pronunciation training in German.

### 3.4. Labeling and data analysis

The data was automatically segmented into words and segments with WebMAUS [16]. Each boundary was then checked in praat [17] and corrected whenever necessary. On a third praat tier we annotated all accented words as such and each accented target word was additionally labeled prosodically on a fourth tier using a simplified ToBI system. The f0 maximum in the accented target word was marked with H, and L was placed before or after the H to quantify the pitch trajectory over a word. In words with a rising-falling pattern we labeled the second low tone as L3. Deviations from the correct lexical stress pattern were labeled as such by adding an asterisk to the tone that was aligned with the stressed syllable (i.e., this simplified system does not make use of starred tones, cf. [12]). In the subsequent analyses we analyzed falling (i.e., HL sequences) and rising (LH patterns) accents as well as monotonal high tones (H) and rising falling patterns (LHL3). The steepness of the fall was only analyzed in falling accents.

A total of 234 recordings of Hungarian learners of German and German control speakers were converted to an EMU format and analyzed within an EMU database [18] (six recordings were excluded because of hesitations or prosodically not classifiable structures). All repetitions including the block where listeners answered the question after they listened to the dialogue were included in the analysis.

The proportions of false lexical stress (H1) and broad focus (H2) were analyzed using Chi-square tests. Sentence position (initial vs. medial) was entered as a fixed factor to predict the proportion of broad focus. Hypothesis H3 was analyzed by means of ANOVAs with the f0 decrease from H to L in semitones per second in HL sequences as dependent variable and either speaker group (L2 vs. L1) or word (*Himbeeren* vs. *Melonen*) as independent variables.

#### 4. Results

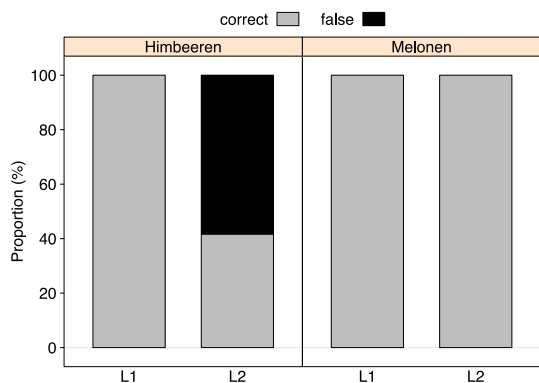


Figure 1: Proportion of correct stress patterns in initially (*Himbeeren*) and medially stressed syllables (*Melonen*) shown separately for Hungarian (L2) and German (L1) speakers.

In the first analysis, we included only the 212 accented target words from sentences with only one accented word. As can be seen in Figure 1, Hungarian speakers (L2) varied to a greater extent in the placement of lexical stress than German speakers (L1) ( $\chi^2[1] = 37.94$ ,  $p < 0.001$ ), but all of these word stress errors (13.2 % in total) occurred in *Himbeeren* where the lexical stress falls on the initial syllable. *Himbeeren* was produced significantly more often with a false lexical stress pattern (58.3 % of all *Himbeeren*-Tokens) than *Melonen* ( $\chi^2[1] = 35.48$ ,  $p < 0.001$ ). These findings do not support H1 which predicted more errors in *Melonen* given that stress on the second syllable does not occur in Hungarian.

To test H2, all 234 recordings, i.e., including those with a deviant stress pattern, were analyzed. While L1 speakers showed narrow focus realizations in all sentences, L2 speakers produced 20.7 % of all sentences with broad focus (cf. Figure 2). These broad focus realizations were mainly caused by an additional accent on the initial word. Commensurate with Figure 2, broad focus realizations occurred significantly more often in sentences with the target word in medial (37.9 %) as opposed to initial (3.4 %) position ( $\chi^2[1] = 18.97$ ,  $p < 0.001$ ). This result supports hypotheses H2.

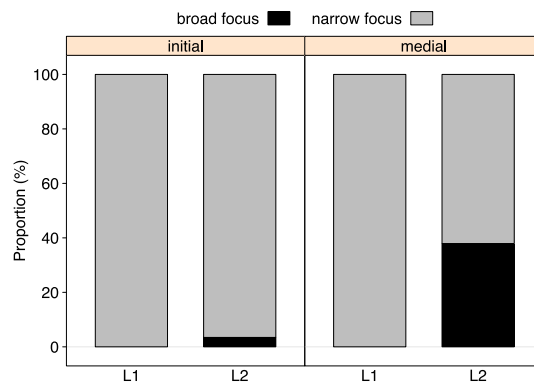


Figure 2: Proportion of broad focus realizations in sentences with the target word in initial (left) and medial (right) position shown separately for Hungarian (L2) and German (L1) speakers.

Speaker group	Pitch accents				Target word
	LH	HL	H	LHL3	
Hungarian	2	12	6	0	<i>Himbeeren</i>
	10	19	0	17	<i>Melonen</i>
German	13	9	12	26	<i>Himbeeren</i>
	49	0	0	9	<i>Melonen</i>
$\Sigma$	74	40	18	52	

Table 1: Distribution of pitch accents on the 184 target words with correctly realized lexical stress.

For the analysis that tested H3, we excluded all tokens with false lexical stress patterns. Within the remaining 184 tokens, Hungarians and Germans differed in the distribution of pitch accents realized on the target words (cf. Table 1). Compared to Germans, Hungarians realized indeed more target words in focus position with falling (HL) pitch accents (7.6 % vs. 47.0 %), but overall the number of rising (LH, 18.2 %) and rising-falling (LHL3, 34.8 %) accents (which are appropriate in this context according to the use of prosodic focus in German) was quite high (53.0 %). This finding supports the first prediction of hypothesis H3. Germans, on the other hand, produced in 92.4 % of all instances accents where a high tone is aligned with the stressed rhyme (i.e., rising, rising-falling or monotonal high accents) and which signal new information. The form of the pitch accent was highly dependent on the position of the lexical stress within a word.

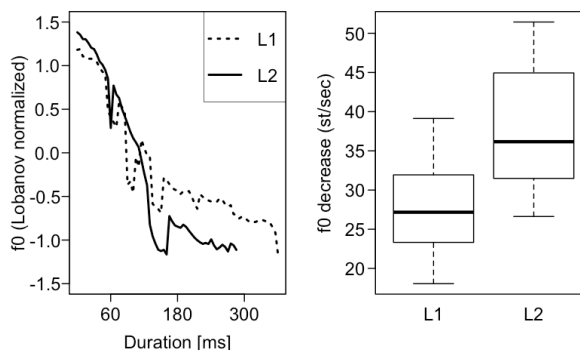


Figure 3: f0 trajectory between H and L in *Himbeeren* (left) and the decrease in f0 in semitones per second (right) shown separately for German (L1) and Hungarian (L2) speakers.

With respect to the second prediction of hypothesis H3, we compared the falling accents in *Himbeeren* produced by Hungarians and Germans (see Figure 3). The decrease in  $f_0$  in semitones per second was significantly steeper for Hungarian than for German speakers ( $F[1,18] = 8.4, p < 0.01$ ). Figure 4 shows that the form of the falling accents in Hungarian speakers is the same in *Melonen* and *Himbeeren*, i.e., the position of the lexical stress did not affect the steepness of the  $f_0$  fall ( $F[1,28] = 2.0, p = 0.17$ ).

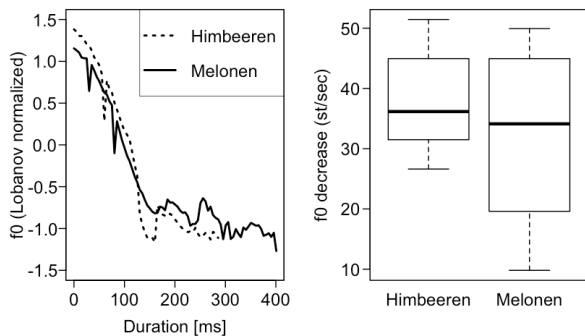


Figure 4:  $f_0$  trajectory between H and L in *Himbeeren* and *Melonen* (left) produced by L2 speakers and the decrease in  $f_0$  in semitones per second (right) shown separately for the two words.

## 5. Discussion

Three main findings arise from this study. First, Hungarian learners of German made errors in lexical stress, but, unexpectedly, all these errors occurred in *Himbeeren*, where the lexical stress falls on the first syllable and which therefore should not cause any problems for L1 speakers of a language with fixed stress on the initial syllable. One possible explanation for this finding could be that the learners in this study overgeneralized the frequent pattern in which lexical stress is (1) on the penultimate syllable and (2) co-occurs with long vowels, which they may have learnt in the foreign language pronunciation training. The results also point to a certain amount of stress deafness within Hungarian speakers, but this trend appeared less pronounced as suggested by [3], given that they produced *Melonen* always correctly. Moreover, contrastive stress occurs only in very few German minimal pairs, which is why Hungarian learners of German possibly judge stress as being less important and vary to some extent on the lexical stress level.

The second finding was that Hungarian learners of German tend to realize utterances with the focused word in medial position with a second accent on the first word. This finding thus confirms hypothesis 2 and supports the claim that Hungarian is a left-headed language. Nevertheless, Hungarians' performance of narrow focus realizations in medial position was quite good, which may be a result of the unmarked German sentence structure which is likely to be frequently used in foreign language courses. Additionally, rising accents on the focused words occurred in the dialogues at the beginning of the experiment. However, we do not believe that the L2 speakers merely imitated the model speakers of the initially presented dialogues because then we should see rising accents in both positions, but Hungarians tended to produce falling accents in initial position. Another reason may have been the German sentence bracket which is formed when a sentence contains a complex predicate. The

finite verb in our test sentences (whose position predicts focus placement in Hungarian) was always in second position, but the last position in the sentence was filled with the infinite main verb which carries the content verb. Therefore, the medial position in the carrier phrase may have been interpreted as preverbal by some Hungarian speakers. The finding that Hungarians' narrow focus realizations were almost native-like in initial position can be explained by the fact that in this position of the sentence the target word occurred unambiguously in preverbal position. Note that the observed broad focus realizations may not be intended as such by the Hungarian learners of German given that they should accent phrase-initial units due to the left-headedness in Hungarian; but they conform to broad focus realizations in German. This finding also supports previous findings for Hungarian where an additional accent was found when the target focus occurred in sentence final position [7].

The third finding showed the transfer of native forms of pitch accents in the foreign language. More precisely, the target word *Melonen* was produced with falling accents only by Hungarians. Falling accents on focused constituents are less common in German, because they are associated with given information. The falling accents were predominantly produced by two Hungarian subjects who never stayed in a German speaking country. The target word *Himbeeren* was realized with falling accents by both groups, although the German control speakers (where this accent may have been a consequence of the lexical stress on the initial syllable) produced them less than the Hungarians. The comparison of the two speaker groups revealed a significant difference in the decrease of the  $f_0$  (measured in semitones per second). This finding supports the analysis of the low target as a trailing tone in Hungarian. The falling accents as produced by Hungarians had the same form in *Himbeeren* and in *Melonen*. Another observation was that when the Hungarian learners of German produced a rising accent pattern in *Himbeeren*, then they shifted the lexical stress to the penultimate syllable. Thus, Hungarians may have learned that in German focus is associated with rising accents, but the realization of these accents is accompanied by stress errors (this observation could also explain the unexpected finding of more stress errors in *Himbeeren*).

The difference in rising vs. falling accents may be perceptually salient to some learners and should therefore be produced according to the SLM. The L2 speaker produced indeed quite a large number of rising focus accents despite the fact that these are typically falling in Hungarian. The difference in the form of falling accents, on the other hand, is presumably less salient in perception. Phonemic categories that are phonetically similar in the L2 and the L1 should be assimilated according to PAM. The steeper fall produced by Hungarians may be just such an instance of perceptual assimilation in the domain of prosodic categories.

## 6. Acknowledgements

This research was carried out within the project "Form and function of prosodic structure in Hungarian and German" supported by DAAD grant no. 57143585 to F. Kleber & MÖB grant no. 65032 to K. Mády. The authors thank U. Reichel & B. Surányi for feedback on this project, Á. Szalontai for help with recordings of the Hungarian learners of German, & R. Winkelmann for technical support in the data analysis.

## 7. References

- [1] Varga, L. (2002). *Intonation and stress: evidence from Hungarian*. Palgrave Macmillan, New York.
- [2] Wiese, R. (2011). *Phonetik und Phonologie*. Wilhelm Fink UTB, Paderborn, 67-99.
- [3] Peperkamp, S., Dupoux, E. (2002). A typological study of stress 'deafness'. *Laboratory phonology*, 7, 203-240.
- [4] E. Kiss, K. (2004). *The syntax of Hungarian*. Cambridge University Press, Cambridge.
- [5] Ladd, D. R. (2008). *Intonational phonology*. Cambridge University Press.
- [6] Helfrich, H. (1985). *Satzmelodie und Sprachwahrnehmung: Psycholinguistische Untersuchungen zur Grundfrequenz*. Berlin, New York: de Gruyter, 3-152.
- [7] Mády, K., Kleber, F. (2010): Variation of pitch accent patterns in Hungarian. In *5th Speech Prosody Conference*, Chicago, USA.
- [8] Baumann, S., Becker, J., Grice, M., Mücke, D. (2007). Tonal and articulatory marking of focus in German. In *Proceedings 16th ICPHS*, 1029-1032.
- [9] Kohler, K. (1991). Terminal Intonation Patterns in Single-Accent Utterances of German: Phonetics, Phonology and Semantics. *AIPUK*, 25, 115-185.
- [10] Hunyadi, L. (2002). *Hungarian sentence prosody and Universal Grammar*. Peter Lang, Frankfurt am Main.
- [11] Mády, K. (2015). Prosodic (Non-) Realisation of Broad, Narrow and Contrastive Focus in Hungarian: A Production and a Perception Study. In *Proceedings Interspeech 2015*, Dresden, 948-952.
- [12] Grice, M., Ladd, D. R., Arvaniti, A. (2000). On the place of phrase accents in intonational phonology. *Phonology*, 17(02), 143-185.
- [13] Flege, J. (1995). Second-language speech learning: Theory, findings and problems, in Strange, W. (1995). *Speech Perception and Linguistic Experience*. York Press: Baltimore.
- [14] Best, C. T. (1995). A Direct Realist View of Cross-Language Speech Perception. *Speech perception and linguistic experience: Issues in cross-language research*, 171-204.
- [15] Draxler, C., Jansch, K. (2004). SpeechRecorder - a Universal Platform Independent Multi-Channel Audio Recording Software. In *Proceedings of LREC*, Lissabon, Portugal, 559-562. Version 2.10.16
- [16] Kisler, T., Schiel, F., Sloetjes, H. (2012). Signal processing via web services: the use case WebMAUS. In *Proceedings Digital Humanities 2012*, Hamburg, Germany, 30-34.
- [17] Boersman, P., Weenink, D. (2015). *Praat: doing phonetics by computer*. Version 5.4.06
- [18] Winkelmann, R. (2015). Managing Speech Database with emuR and the EMU-webApp. In *Sixteenth Annual Conference of the International Speech Communication Association*.