

## Accentual phrases in Slovak and Hungarian

Katalin Mády<sup>1</sup>, Uwe D. Reichel<sup>2</sup>, and Štefan Beňuš<sup>3,4</sup>

<sup>1</sup>Institute for Linguistics, Hungarian Academy of Sciences, Budapest

<sup>2</sup>Institute for Phonetics and Speech Processing, University of Munich

<sup>3</sup>Constantine the Philosopher University, Nitra

<sup>4</sup> Institute of Informatics, Slovak Academy of Sciences, Bratislava

mady@nytud.hu, reichelu@phonetik.uni-muenchen.de, sbenus@ukf.sk

### Abstract

Languages with primarily delimitative function of word stress commonly make use of accentual phrases (APs) in their intonational phonology (e.g. Tamil or French). Slovak and Hungarian are genetically unrelated but geographically close languages with word-initial lexical stress. In this paper we compared the stylised f<sub>0</sub> of single accent groups (AGs) with the f<sub>0</sub> level pattern of the entire intonational phrase (IP) to test if AGs are relevant for the intonational phonology of Slovak and Hungarian. Steep f<sub>0</sub> slopes with a recurring pattern (rising or falling) and large deviations from IP level patterns were interpreted as evidence for the autonomy of the AG in the given language. The results suggest that in Hungarian, accent groups form an independent unit with a falling pitch contour. Such evidence was not found for Slovak.

**Index Terms:** prosodic phrasing, accentual phrase, intonation modelling, Hungarian, Slovak.

### 1. Introduction

Most prosodic models [1, 2, 3] assume that the highest prosodic unit is the intonational phrase (IP) that is marked by final and potentially initial boundary tones (or edge tones), phrase-final lengthening and an optional pause following the phrase. Other models subordinate the IP to a higher-level unit, such as the utterance, e.g. [4]. The smallest prosodic phrase is the prosodic word (PW). Two other units, the intermediate phrase (ip) and the accentual phrase (AP) are located between the intonational phrase and the prosodic word [5]. Whereas IPs and PWs are present in all languages, ips and APs are optional. (See [6, p. 444] for an overview of the presence or absence of certain prosodic units for a set of languages.)

Both Hungarian and Slovak have lexical stress on the left-most syllable of a prosodic word (apart from Eastern Slovak dialects with penultimate stress). In Hungarian, lexical stress is fixed to the first syllable, i.e. *Intonáció* ‘intonation’ (the “accent” sign marks long vowels and not stress). In Slovak, stress can be optionally moved from the word-initial syllable if the word is preceded by a preposition, e.g. *HOry* ‘hills’, *DO hory* ‘to the woods’. Hence, Slovak lexical words may combine with preceding monosyllabic prepositions and possible clitics following this word into a prosodic word with left-most stress.

In Hungarian, pitch accents are typically realised within the initial syllable of the accented word. In Slovak, pitch targets might be aligned within or after the accented syllable, but it is not clear if this is due to phonological differences or phonetic implementation.

In languages with fixed stress towards the left or right edge of the word, stress is often used as the onset or offset of a prosodic phrase [5], while the other edge of the phrase is marked by a boundary tone. In languages in which sequences between two accents form an accentual phrase (AP), the edge tones often show a regular pattern. Tamil for example has phrase-initial L\* tones and AP-final H tones (Ha).

[5] describe accentual phrases in the following way: (1) an AP can contain only one pitch accent, (2) there is a boundary tone at the other edge of the prosodic unit, (3) the edge tone is insensitive of stress, i.e. its realisation is independent of stress location, (4) APs are characterised by a recurrent rising or falling pattern depending on the language.

Given the observed correlation between fixed lexical stress position and relevance of AGs for the intonational phonology of a particular language, we investigate if left-headed prosodic words in Slovak and Hungarian also constitute a prosodic unit that is marked by tonal (f<sub>0</sub>) means. While initial efforts in building a ToBI system for Slovak prosody do not propose units below intermediate phrases [7], previous work on Hungarian suggests that pitch accents indeed initiate APs within intonational phrases [8, 9].

In left-headed accentual phrases, pitch accents are always preceded by a prosodic boundary and thus automatically function as an edge marker. This would explain the relatively frequent occurrence of pauses before accented words in unexpected positions in Hungarian, e.g. between a definite article and an emphasised noun [10]. In this case, the pause can be interpreted as an additional cue marking the phrase boundary and thus strengthening the emphasis on the pitch-accented word. Apart from our theoretical interest in the phrasal organisation of Slovak and Hungarian prosody, the function of the phrase types can help to gain a better understanding of underlying communicative intentions and also to enhance the effectiveness of speech recognition and the naturalness of speech synthesis.

The analysis presented here is based on the assumption that an accentual phrase has its own f<sub>0</sub> slope in form of a rising or falling pitch contour, and that this pattern is independent of the overall pitch contour of the intonational phrase. Second, since APs are characterised by the frequent occurrence of the same pattern (e.g. rising pattern for French and Tamil), the frequent occurrence of a given pattern is interpreted as a further evidence for the presence of APs in the language.

### 2. Material and methods

50 Slovak and Hungarian spontaneous utterances forming a single IP were selected from collaborative dialogues (5 utterances

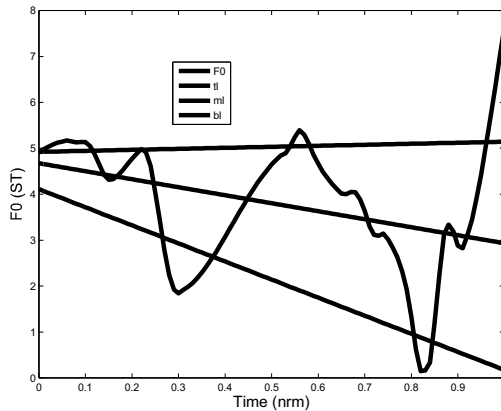


Figure 1: Stylised top-, mid- and baseline.

of 10 Slovak and 10 Hungarian speakers, respectively). All IPs had a low phrase-final boundary tone. Only IPs with at least two pitch accents were used for the analysis. Both IPs and pitch accents were identified manually by a phonetically trained native speaker (1st and 3rd author of this paper). Analysis was based on the f0 level and the range pattern throughout the IP and the accent group (AG), ranging from one accented syllable till the last unaccented syllable before the next pitch accent or the IP boundary. The Slovak material consisted of 157 accent groups and the Hungarian samples of 130 accent groups.

### 2.1. F0 extraction and preprocessing

F0 was extracted by autocorrelation (Praat 5.3, sample rate 100 Hz). Data were further processed in Matlab. Voiceless utterance parts and f0 outliers were interpolated by piecewise cubic splines [11]. The contour was then smoothed by Savitzky-Golay filtering using third order polynomials in 5 sample windows and transformed to semitones relative to a base value [12, 13]. This base value was set to the f0 median below the 5th percentile of an utterance and served to normalise f0 with respect to its overall level.

### 2.2. Stylisation

To capture the f0 register in terms of its level and range [14] we fitted a base-, a mid-, and a topline separately for the IP and all AGs within this IP (Fig. 1). The midline represents the f0 level, whereas the base- and topline provide the f0 range information. The robust fitting procedure that is motivated and explained in more detail in [15] consists of the following steps:

- A window of 200 ms length is shifted along the f0 vector with the step size of 10 ms.
- Within each window the f0 median is calculated
  - of the values below the 10th percentile for the baseline,
  - of the values above the 90th percentile for the topline, and

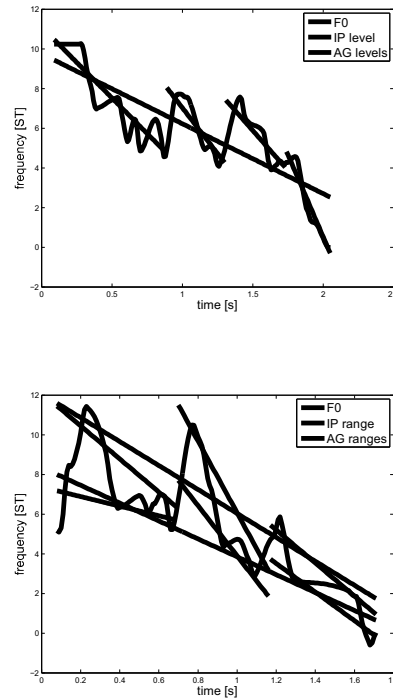


Figure 2: IP and AG level (top) and range (bottom) patterns for Hungarian.

- of all values for the midline.

This gives 3 sequences of medians, one for the base-, the mid-, and the topline, respectively.

- Within each median sequence outliers are replaced by linear interpolation.
- Finally, for all three median sequences linear polynomials are fitted.

### 2.3. Features

If accentual phrases play a role in the intonational phonology of a language, they should show a significant difference from the corresponding part of the intonational phrase in terms of local f0 level deviations and f0 ranges. In other words, the slope of the f0 midline of a given AP is supposed to be steeper than that of the IP, and the range larger than can be expected on the basis of the overall top- and baseline differences in the IP.

We extracted features to capture first the general AG register and second its deviation from the IP.

#### Level

For the f0 level, the AG pattern is represented by the slope of the fitted midline (*mlslope*), as shown in the left graph in Figs. 2 and 3. AG deviation was measured (1) with respect to shape in terms of the absolute slope difference of the AG and the IP midlines (*mlSlopeDiff*), (2) with respect to overall distance given by the root mean squared deviation of the AG line from the corresponding section of the IP line (*mlRms*, and (3) with respect to local differences by subtracting the corresponding values of the

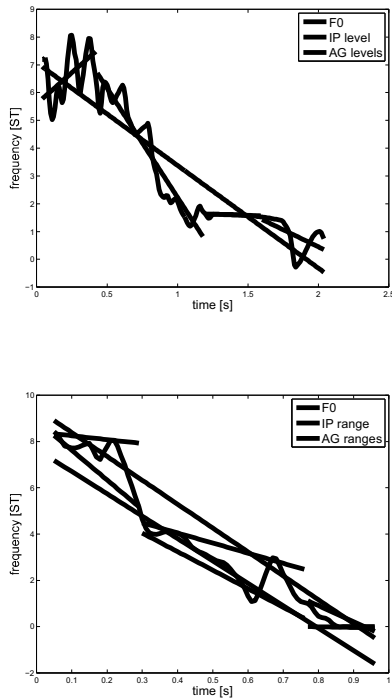


Figure 3: IP and AG level (top) and range (bottom) patterns for Slovak.

IP line from the initial and the final f0 values of the AG line ( $mlInitYDiff$ ,  $mlFinYDiff$ ).

### Range

Next to level deviations, AGs can be set apart from their underlying IPs by the AG range, see the bottom graphs in Figs. 2 and 3. We thus measured the range of each AG in terms of the root mean squared deviation between its base- and topline ( $rangeRms$ ).

The expected acoustic correlates for the presence of accented phrases are prominent f0 movements reflected in high AG range values ( $rangeRms$ ) as well as considerable local level deviations between the AG and the IP expressed in high values for the features  $mlSlopeDiff$ ,  $mlRms$ ,  $mlInitYDiff$ , and  $mlFinYDiff$ . Results were compared by means of  $t$ -tests. First, a one-sample  $t$ -test was carried out for the features for both languages, i.e. results were compared to a sample with  $mean = 0$ . Subsequently, Hungarian and Slovak samples were compared to each other. If the data were not normally distributed, the Mann-Whitney test was carried out instead. The equality of variances was tested by the Levene test that is also applicable to non-normally distributed data. Significance level was set to  $p = 0.05$ .

Languages with falling or rising APs are expected to have mostly negative or mostly positive slopes that differ considerably from the IP slope. At the same time, the onset and/or the offset of the AG slope is supposed to differ from the corresponding section of the IP. Additionally, larger overall distances (root

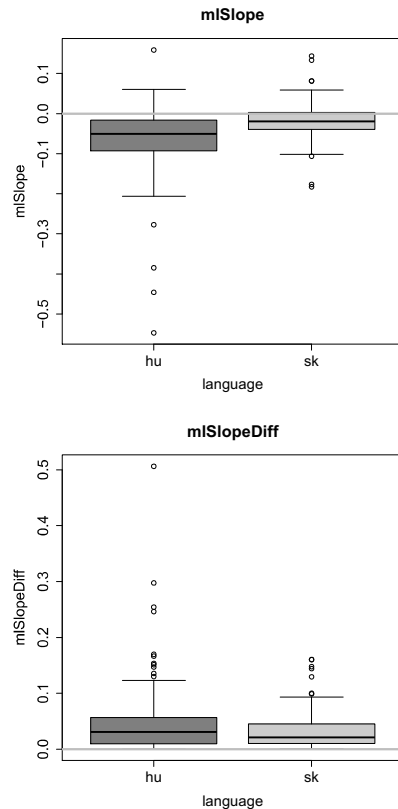


Figure 4: Slope ( $mlSlope$ ) and absolute slope difference ( $mlSlopeDiff$ ) for Hungarian and Slovak. The grey line indicates 0.

mean square distances) and larger ranges demonstrate a greater autonomy of the AG as a prosodic phrase.

## 3. Results

### 3.1. Slope

The majority of slopes ( $mlSlope$ ) was negative in both languages. Absolute slope differences ( $mlSlopeDiff$ ) were larger for Hungarian (Fig. 4). Mean slopes and mean slope differences differed significantly from 0 in both languages, and also between the two languages. All differences were highly significant ( $p < 0.0001$ ).

The overall distance ( $mlRms$ ) between the AG midline and the corresponding IP section (Fig. 5) differed significantly from 0, but not between the two languages ( $p = 0.24$ ).

### 3.2. Local initial and final AG–IP difference

The distance between the AG and the IP midlines was calculated both for the onset ( $mlInitYDiff$ ) and the offset ( $mlFinYDiff$ ) of the AG. Larger distances refer to a larger deviation and thus a larger independence of the AG midline. The distance was expressed in absolute values.

Both AG-initial and AG-final distances were significantly larger than 0 in both languages ( $p < 0.01$ ) (Fig. 6). Onset

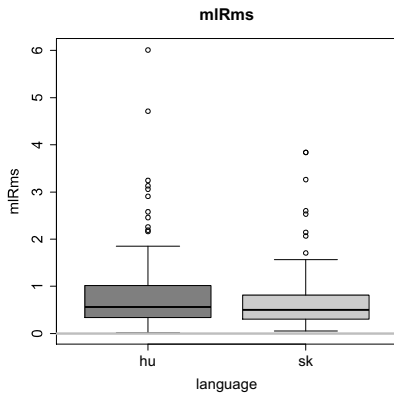


Figure 5: Root mean square between the AG and IP midline (mlRms for Hungarian and Slovak. The grey line indicates 0).

distances were significantly larger in Hungarian than in Slovak, whereas offset distances did not differ between the two languages.

**3.3. Range**

Finally, the range between the base- and the topline was compared for the AG. It was assumed that a larger range within the AG shows the autonomy of the AG as a separate prosodic unit.

Ranges differed significantly from 0, but not between languages ( $p = 0.67$ ).

**4. Conclusions**

AGs show a falling pattern in both languages, i.e. their f0 slope is steeper than the IP slope. This tendency is more clear-cut in Hungarian than in Slovak. The onset and the offset of the AG slope deviate from the IP slope for both languages, while the distance of the onset is greater in Hungarian than in Slovak. F0 slopes in Slovak are rather flat, and the tendency for uniform patterns is not present. In other words, the findings do support the hypothesis that Hungarian prosody involves accentual phrases, but the evidence is less clear for Slovak.

One possible reason for the flatness of Slovak f0 slopes is that the AGs might have a non-linear pattern. As was said before, the f0 maxima of pitch accents are often delayed in Slovak which might cause a rising-falling pattern or be triggered by it. This option needs further testing by non-linear stylisation.

A future large-scale study will involve more languages without fixed lexical stress. This would allow to specify the interrelation of f0 slopes and prosodic phrases in a more detailed way.

**5. Acknowledgements**

This work was supported by the Hungarian Scientific Research Fund (PD 101050) and the Momentum program of the Hungarian Academy of Sciences (first author) and by the ERDF Research & Development Operational Programme *Research and development of new information technologies for forecasting and mitigation of crisis situations and safety*, ITMS 26240220060 (third author).

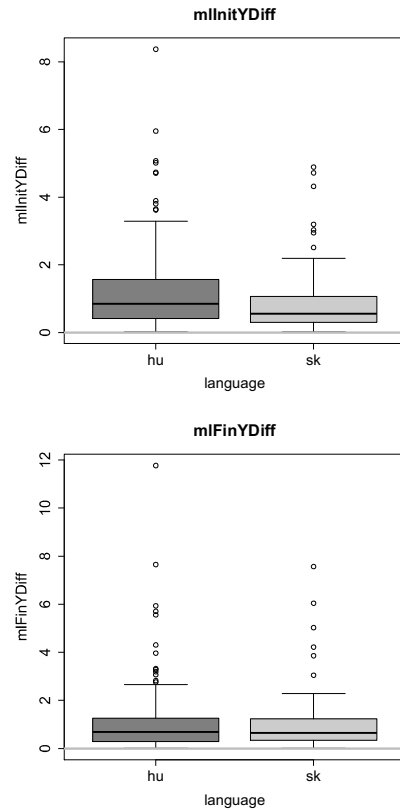


Figure 6: Local initial (mlInitYDiff) and final (mlFinYDiff) AG-IP differences for Hungarian and Slovak (absolute values). The grey line indicates 0.

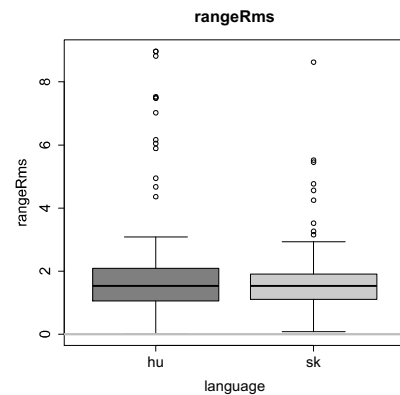


Figure 7: Range between base- and topline (rangeRms) for Hungarian and Slovak. The grey line indicates 0.

## 6. References

- [1] A. Cruttenden, *Intonation, 2nd ed.* Cambridge: University Press, 1997.
- [2] M. Beckman and J. Pierrehumbert, "Intonational structure in Japanese and English," *Phonology*, pp. 255–309, 1986.
- [3] M. Vogel and I. Nespó, *Prosodic phonology*. Dordrecht: Foris Publications, 1986.
- [4] C. Gussenhoven, "Transcription of Dutch intonation," in *Prosodic typology*. Oxford: University Press, 2005, pp. 118–145.
- [5] S.-A. Jun and J. Fletcher, "Methodology of studying intonation: From data collection to data analysis," in *Prosodic Typology II: the new development in the phonology of intonation and phrasing*. Oxford: University Press, 2014.
- [6] S.-A. Jun, Ed., *Prosodic typology*. Oxford: Oxford University Press, 2005.
- [7] M. Rusko, R. Sabo, and M. Džúr, "Sk-tobi scheme for phonological prosody annotation in Slovak," in *Text, speech and dialogue*, ser. Lecture notes in computer science, V. Matoušek and P. Mautner, Eds. Berlin & Heidelberg: Springer, 2007, vol. 4629, pp. 334–341.
- [8] L. Hunyadi, *Hungarian sentence prosody and universal grammar: on the phonology–syntax interface*. Frankfurt/Main: Lang, 2002.
- [9] K. Mády, A. Szalontai, A. Deme, and B. Surányi, "On the interdependence of prosodic phrasing and prosodic prominence in Hungarian," in *Proc. 11th International Conference on the Structure of Hungarian*, Piliscsaba, Hungary, 2013.
- [10] K. Mády and F. Kleber, "Variation of pitch accent patterns in Hungarian," in *Proc. 5th Speech Prosody Conference, Chicago*, 2010, pp. 100924:1–4.
- [11] C. de Boor, *A Practical Guide to Splines*, ser. Applied Mathematical Sciences. Springer, 1978, no. 27.
- [12] A. Savitzky and M. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.
- [13] P.-O. Persson and G. Strang, "Smoothing by Savitzky-Golay and Legendre Filters," in *Mathematical systems theory in biology, communications, computation, and finance*, D. Gilliam, Ed. Springer, 2003, pp. 301–315.
- [14] T. Rietveld and P. Vermillion, "Cues for Perceived Pitch Register," *Phonetica*, vol. 60, pp. 261–272, 2003.
- [15] U. D. Reichel and K. Mády, "Parameterization of F0 register and discontinuity to predict prosodic boundary strength in Hungarian spontaneous speech," in *Elektronische Sprachsignalverarbeitung*, ser. Studentexte zur Sprachkommunikation, P. Wagner, Ed., vol. 65. Bielefeld: TUDpress, 2013, pp. 223–230.