Subsegmental differences between accented and unaccented vowels in Hungarian

Alexandra Markó[1,2], Márton Bartók[2], Tamás Gábor Csapó[2,3], Tekla Etelka Gráczi[2,4], Andrea Deme[1,2]

[1]*ELTE Eötvös Loránd University, Budapest, Hungary*
[2]*MTA-ELTE "Lendület" Lingual Articulation Research Group, Budapest, Hungary*
[3]*Budapest University of Technology and Economics, Budapest, Hungary*
[4]*Research Institute for Linguistics, Budapest, Hungary*

marko.alexandra@btk.elte.hu, bartokmarton@gmail.com, csapot@tmit.bme.hu, graczi.tekla.etelka@nytud.hu, deme.andrea@btk.elte.hu

Abstract

In the present study we searched for an answer to the question if in Hungarian, similarly to the so far investigated Germanic languages, accent results in sonority expansion and/or localized hyperarticulation. The analysis was performed by EMA, with the participation of 9 speakers. Four vowels /i, u, ɒ, aː/ *were tested in nonsense* $pV_1pV_1pV_1pV_1$ *sequences uttered as sentences. Accented (first) and unaccented (second) syllables were compared. In terms of lip aperture index and duration, sonority expansion was detected in general, but post hoc tests did not show the same effect for each vowel as a function of accent. Tongue positions did not show differences, however the Euclidean distance from the acoustic vowel space centroid differed in accented and unaccented vowels, which might be traced back to the* F_1 *differences, probably due to the lip aperture enhancement in* /aː/. *All considered, effects of sonority expansion were observed, however, data did not show localized hyperarticulation patterns.*

Keywords: accent, sonority expansion, localized hyperarticulation, EMA, Hungarian

1. Introduction

A tendency can be observed for making accented segments more salient at various levels of human communication, not only in a prosodic way. In prominent positions, articulation is strengthened, i.e., articulatory gestures in these segments tend to be realized with increased spatio-temporal magnitude. Expression of contrast between prominent and non-prominent segments might be reached by sonority expansion and/or local hyperarticulation (see an overview in Mücke & Grice 2014).

Beckman et al. (1992) defined 'sonority expansion strategy' as an intention of the speaker to produce a louder vowel in the accented syllable. In articulation, this intention leads to a more open oral passage over a longer period. Based on Lindblom's (1990) H&H model, de Jong (1995) proposed the strategy of 'localized hyperarticulation' which suggests that in the accented syllables speakers are prone to utter more peripheric vowels (in terms of tongue position) in order to differentiate vowel qualities to a larger extent. Interpretation of hyperarticulation in the case of vowels obviously depends on the place of the target within the vowel space. Accordingly, if hyperarticulated, low vowels are produced with a lower tongue position, front vowels in general show a more fronted tongue position, while non-low front vowels exhibit a higher tongue body posture, and back vowels are pronounced with a more retracted tongue position (de Jong 1995; Harrington *et al.* 2000; Cho 2005). All considered, in the case of low vowels, both sonority expansion and localized hyperarticulation elicit lower tongue position and more open oral cavity, therefore the two strategies strengthen each other. However, in non-low vowels, the two goals are in conflict, as e.g. in the case of /i/ sonority expansion (mouth opening) contradicts with the reaching of the target more precisely (localized hyperarticulation). In this case, two contradictory effects may operate: either opening of the lips/mandible or tongue raise and fronting (Harrington *et al.* 2000).

In the prosodic focus marking languages (like German, English or French) evidence was found both for localized hyperarticulation and sonority expansion as well as segmental lengthening in the case of prominence (e.g., Cho 2005; Baumann *et al.* 2007). Harrington *et al.* (2000) analyzed Australian English /i æ a/ in two speakers. The authors observed large individual differences in the ratio of hyperarticulation and sonority expansion. For accented /i/ they found higher tongue position in one of the speakers, while more fronted tongue in the other one.

In Hungarian, these tendencies have not been analyzed so far. Since Hungarian is an obligatory syntactic focus marking language (Genzel *et al.* 2015), accentuation plays a limited role in expressing focus (Mády & Kleber 2010) compared to the widely studied (mostly Germanic and Romance) prosodic focus marking languages. In prosody, Hungarian (sentence) accent is mainly indicated by f0 (e.g., Szalontai *et al.* 2016). Additionally, due to the fixed first syllable word stress, the position of word level prominence is also constrained and predictable.

As vowel quantity is distinctive in Hungarian, it was argued that lengthening cannot extensively be used for accentuation (see e.g., Mády & Kleber 2010). Recent studies, however, have found that longer vowel duration plays a role in the expression of prominence (e.g., Szalontai et al. 2016; Markó et al. 2018; 2019).

With respect to vowel quality in terms of prominence, Hungarian does not exhibit phonological vowel reduction (Gósy 1997), therefore segmental features of prominent vs. non-prominent vowels have hardly been analyzed systematically in Hungarian. Non phonological nature vowel centralization has been documented in spontaneous speech (e.g., Gósy 1997) but in these studies, prominence effect was not controlled. In a preliminary study (on a material which was recorded with another purpose and therefore was not perfectly controlled in terms of prominence and vowel quality) we performed a formant analysis of /u/, /i/, /ɒ/ vowels uttered in accented + stressed vs. unaccented + unstressed syllables (Markó *et al.*

2018). F_2 was lower in vowels bearing prominence than in vowels realized without prominence, although Euclidean distance from the vowel space centroid was not influenced by prominence. Variability of both formants was considerably smaller in /u/ and /ɒ/ if the vowel bore prominence, while /i/ did not show the same pattern. In another study (Markó *et al.* 2019), acoustic comparison of /u/, /i/, /ɒ/ and /ɛ/ was carried out between sentence-initial focus (higher prominence) and sentence-initial topic (lower prominence) positions, where the target vowels appeared in the first syllable (bearing word stress). In this case, variance of F_2 was also found to be smaller in focus position compared to that of topic position. In sum, in the previous studies on Hungarian vowels in accented and non-accented syllables, second formant values appeared to tend to be less variable in more prominent positions, which might be the result of localized hyperarticulation.

As we have seen, only a few acoustic data are available on Hungarian vowel realization differences as a function of prominence. With respect to the articulatory domain, nevertheless, there is no data on supraglottal articulatory characteristics of prominent and non-prominent vowels in Hungarian at all. The aim of the present study, therefore, was to compare segmentally identical and consecutively produced accented + stressed (first) vs. unaccented + unstressed (second) syllables, where the difference between these syllables is supposed to be maximal, both in the acoustic and the articulatory domains.

Our research question is twofold: whether any acoustic and/or articulatory evidence of either localized hyperarticulation or sonority expansion as well as segmental lengthening can be observed in Hungarian, as a function of prominence. Based on the above-mentioned previous results, we expected differences between the vowels appearing in accented and in unaccented syllables in both localized hyperarticulation (in terms of tongue position and acoustic vowel space) and sonority expansion (in terms of lip aperture and vowel duration).

2. Method

2.1. Participants, material, and data recording

Nine healthy native speakers of Hungarian participated in the study (all females, aged 25.2±5.9 years). They produced four-syllable nonsense words as single utterances with the structure of $pV_1pV_1pV_1pV_1$, where the V-to-V coarticulation effect was eliminated using the same vowel qualities over the syllables. Minimum 6 occurrences of each target word were recorded, in a random order, among fillers. The words were displayed on a computer screen.

Four edge vowels of the Hungarian vowel space were selected for the experiment: back and high /u/, front and high /i/, back and low /ɒ/ and medial and low /aː/. (NB: /aː/ phonologically behaves as a back vowel, and its quantity is always long, as it functions as the long phonological counterpart of /ɒ/. /i/ and /u/ have long phonological pairs with the same vowel qualities: /iː/ and /uː/, but these latter occurrences were not tested here.)

Data recordings were carried out in a sound treated room using a Carstens EMA AG501 system. We recorded the upper and lower lip movements, and the tongue movements at tongue tip (TT), tongue blade (TBL), and two points on the tongue dorsum (TBO1, TBO2). Synchronously, we recorded the acoustic signal with head-mounted omnidirectional condenser microphone, at 44.1 kHz.

2.2. Data processing and analysis

Head movement and bite plane corrections were done by the Carstens software, while further post-processing (3D-2D conversion, and production of Emu-compatible ssff tracks) was carried out by the custom-made converter of the IfL Phonetik, University of Cologne. Segmental labelling of the audio signal was carried out semi-automatically using the BAS web services G2P (Reichel 2012) and MAUS (Schiel 1999). For gestural labelling we used Emu (Winkelmann *et al.* 2018).

The following acoustic data were obtained from the audio signal in the case of the first and second syllables' vowel: duration, as well as f_0, F_1 and F_2 at the temporal midpoint of the target vowels. On the basis of F_1 and F_2 data, the Euclidean distance of the acoustic vowel space centroid and each token was calculated.

The difference in prominence between the first and second syllable vowel was controlled by the difference of the f0 values (**Figure 1**). The values showed significant difference ($F(1, 418.03) = 180.32$, $p < 0.001$), and the post hoc tests showed $p < 0.001$ significance in all vowel qualities. Thus, the first syllable was considered as accented, while the second one as unaccented syllable.

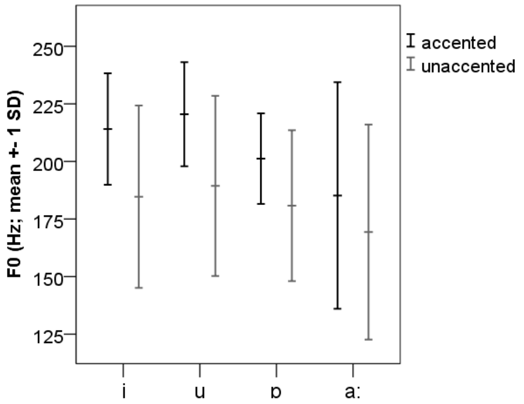

Figure 1: *F0 (mean and SD) as a function of prominence.*

From the articulatory data, the horizontal and vertical displacement of the lip sensors and the first tongue dorsum (TBO1) sensor (obtained also at the temporal midpoint) were used in the present analysis. The horizontal (x) and the vertical (y) positions of the TBO1 (3 cm from the tongue tip) receiver coil were measured and normalized based on Cho (2004). In order to get compatible data across speakers, the minimum and maximum of x and y values were obtained for each speaker. In terms of horizontal axis, the minimum x value was the extreme case of /u/, i.e. the backmost tongue position for /u/ (0%), and the maximum x value was the extreme case of /i/, i.e. the most fronted tongue position for /i/ (100%) (**Figure 2**).

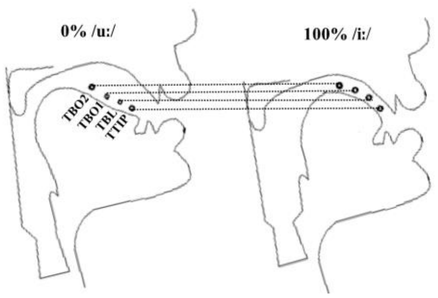

Figure 2: *Reference points used for the normalization of sensor positions:* /u/ *= 0%),* /i/ *= 100%), and the scale defined by these extremes on the x axis (based on Cho 2004).*

Likewise, the minimum *y* value was the extreme case of /aː/ (the lowest tongue position for /a/, 0%) and the maximum *y* value was the extreme case of /i/ (the highest position of it, 100%). Based on the horizontal (*x*) and the vertical (*y*) positions of the TBO1 sensor, the Euclidean distance of the articulatory vowel space centroid and each token was calculated.

Vertical (*y*) positions for the upper lip and the lower lip transducers were examined. These positions were used to calculate the distance between these receivers, called lip aperture *y* (lip aperture y = upper lip y – lower lip y; Byrd 2000), then normalized for each speaker. Lip aperture *y* maximum per participant provided 100%.

GLMMs were run on the data in R (R Core Team 2018), using the lme4 package (Bates *et al.* 2015). *p*-values were obtained via the Satterthwaite approximation available in lmerTest package (Kuznetsova *et al.* 2017). We included random intercepts for speakers. Post hoc analysis (Tukey test) was carried out by lsmeans package.(Lenth 2016).

3. Results

Localized hyperarticulation was operationalized by the differences of tongue position and acoustic vowel space, while sonority expansion was measured by lip aperture and duration differences between accented (first) and unaccented (second) syllable vowels.

Tongue positions did not show differences (**Figure 3**), neither in horizontal nor in vertical dimension. However, Euclidean distance from the acoustic vowel space centroid differed between accented and unaccented vowels ($F(1, 445.92) = 8.92$, $p < 0.01$) (**Figure 4**).

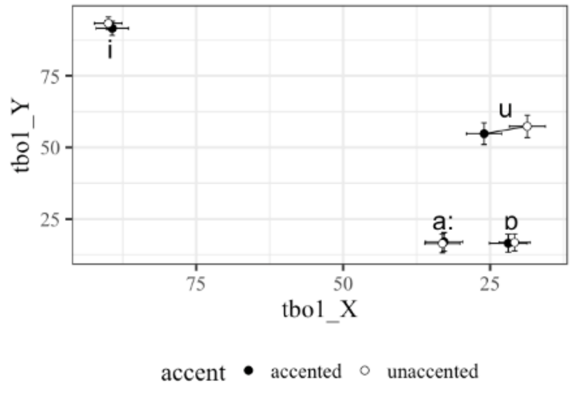

Figure 3: *Tongue position (%, mean and SD) of the analyzed vowels as a function of prominence.*

This is probably due to the difference detected in F_1 ($F(1, 411.14) = 4.69$, $p < 0.05$). Although, post hoc test was not significant in any of the vowels, a vowel-dependent tendency can be observed in F_1 (**Table 1**). In all vowel qualities, accented realizations appear to be 'lower', however vertical tongue position did not vary as we have seen above. F_2 did not differ between the target vowels as a function of accent.

Lip aperture *y* (**Figure 5**) showed significant differences between accented and unaccented vowels ($F(1, 445.96) = 3.87$, $p < 0.05$), however, post hoc test was significant only in the case of /aː/ ($p < 0.001$). Duration (**Figure 6**), in general, differed between the conditions ($F(1, 411.02) = 4.40$, $p < 0.05$) but post hoc tests did not replicate this effect within any of the vowel qualities.

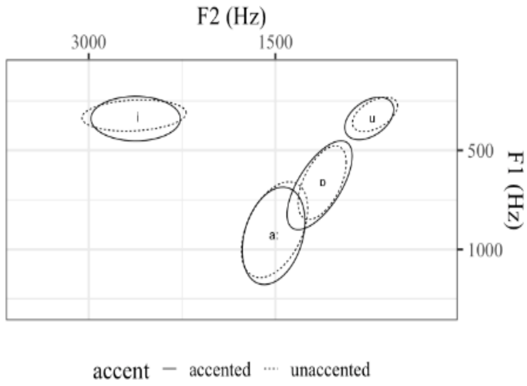

Figure 4: *F_1 × F_2 space of the analyzed vowels as a function of prominence.*

Table 1: *F_1 values as a function of the prominence (mean and SD).*

	accented	unaccented
/i/	334±49 Hz	324±40 Hz
/u/	338±45 Hz	326±57 Hz
/ɒ/	661±112 Hz	654±97 Hz
/aː/	939±112 Hz	906±106 Hz

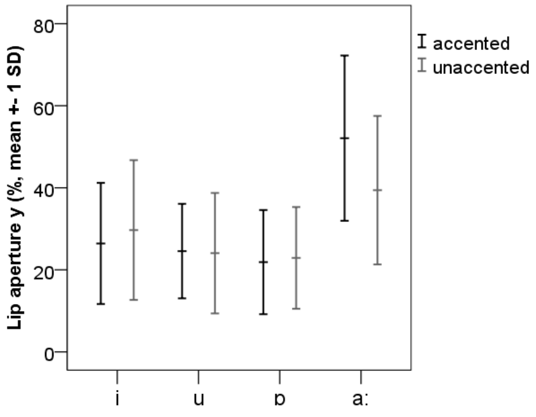

Figure 5: *Lip aperture y values (mean and SD) as a function of prominence.*

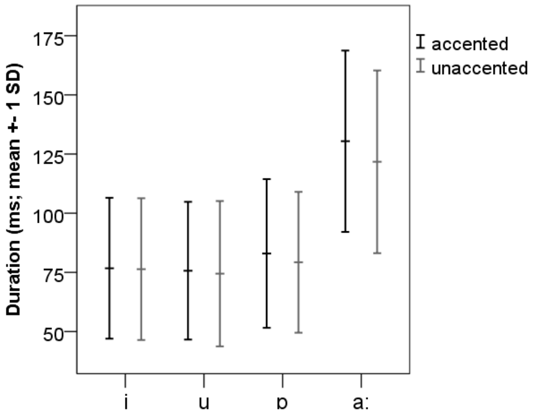

Figure 6: *Tongue position of the analyzed vowels as a function of prominence.*

4. Discussion and conclusion

The present study searched for an answer to the question if in Hungarian, similarly to the so far investigated Germanic languages, accent results in sonority expansion and/or localized hyperarticulation. Tongue position data suggest that in accented syllables localized hyperarticulation did not occur in the analyzed vowels. On the basis of lip aperture measurement, sonority expansion appears to be an existing strategy in expressing accent, especially in the case of the phonologically long /aː/. The acoustic data showed differences in F_1, which might be traced back to the differences of lip aperture, since tongue positions did not vary as a function of accent. Differences of F_1 might influence the Euclidean distance from the acoustic vowel space centroid leading to differences between accented and unaccented syllables. Vowel duration data also correspond to the sonority expansion strategy.

As for our presupposition, with respect to prominence these data confirm sonority expansion strategy in the case of the analyzed Hungarian vowels, but we did not find evidence of localized hyperarticulation.

Tendencies of prosodic strengthening, as in English, showed variable patterns, thus individual differences might have led to non-significant results. In the present experiment we used nonsense speech material, which made it possible to control various effects to a large extent, however, it might have resulted in a rather artificial output. Finally, the present analysis was carried out on static data (basically obtained from the temporal midpoint of the vowels), but kinematic characteristics have not been analyzed yet. Further studies are necessary to analyze for example kinematic data, real words, and individual differences.

5. Acknowledgements

The authors acknowledge the kind help of Zsófia Weidl, Zsófia Puzder, Valéria Krepsz, Ákos Buza, István Mrázik, Anne Hermes, Doris Mücke, and Theo Klinker. The research was funded by the Bolyai János Research Scholarschip of the Hungarian Academy of Sciences, the ÚNKP-20-5 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund, and the Thematic Excellence Program of the Ministry for Innovation and Technology.

6. References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software,* 67, 1-48.

Baumann, S., Becker, J., Grice, M., & Mücke, D. (2007). Tonal and articulatory marking of focus in German. *ICPhS XVI* ID 1378.

Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In: G. J. Docherty & D. R. Ladd (Eds.), Papers in laboratory phonology II: Segment, gesture, prosody. Cambridge: Cambridge University Press. 68-86.

Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica,* 57, 3–16.

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics,* 32, 141-176.

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /ɑ,i/ in English. *Journal of the Acoustical Society of America, 117,* 3867-3878.

Genzel, S., Ishihara, S., & Surányi, B. (2015). The prosodic expression of focus, contrast and givenness: A production study of Hungarian. *Lingua,* 165, 183–204.

Gósy, M. (1997). Semleges magánhangzók a magyar beszédben [Neutral vowels in Hungarian speech]. *Magyar Nyelvőr,* 121, 9-19.

Harrington, J., Fletcher, J., & Beckman, M. E. (2000). Manner and place conflicts in the articulation of accent in Australian English. In M. Broe (Ed.), Papers in Laboratory Phonology, 5. Cambridge: Cambridge University Press. 40-55.

de Jong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America,* 97, 491-504.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software,* 82, 1-26.

Lenth, R. V. (2016). Least-squares means: the R package lsmeans. *Journal of Statistical Software,* 69, 1-33.

Lindblom, B. (1990). Explaining Phonetic Variation: A Sketch of the H&H Theory. In W. Hardcastle & A. Marchal (Eds.), Speech production and speech modelling. Dordrecht: Kluwer Academic. 403-439.

Mády, K., & Kleber, F. (2010). Variation of pitch accent patterns in Hungarian. *Proceedings of Speech Prosody 2010.* http://speechprosody2010.illinois.edu/papers/100924.pdf

Markó, A., Bartók, M., Gráczi, T. E., Deme, A., & Csapó, T. G. (2018). Prominence effects on Hungarian vowels: A pilot study. *9th International Conference on Speech Prosody,* 868-872.

Markó, A., Bartók, M., Csapó, T. G., Deme, A., & Gráczi, T. E. (2019). The effect of focal accent on vowels in Hungarian: articulatory and acoustic data. In Proceedings of the 19th ICPhS, 2715-2719.

Mücke, D., & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation – Is it mediated by accentuation? *Journal of Phonetics,* 44, 47-61.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reichel, U. (2012). PermA and Balloon: Tools for string alignment and text processing. In Interspeech, 13th Annual Conference of the International Speech Communication Association, 9. - 13. September 2012, Portland, USA. Paper no. 346.

Schiel, F. (1999). Automatic phonetic transcription of non-prompted speech. In Proceedings of International Congress of Phonetic Sciences. 607–610.

Szalontai, Á., Wagner, P., Mády, K., & Windmann, A. (2016). Teasing apart lexical stress and sentence accent in Hungarian and German. In Tagungsband 12. Tagung Phonetik und Phonologie im deutschsprachigen Raum (P&P 12). 216–219.

Winkelmann, R., Klaus, J., Cassidy, S., & Harrington, J. (2018). *emuR: Main Package of the EMU Speech Database Management SystemR,* package version 1.1.1.