Shortening of long high vowels in Hungarian: a perceptual loss?

Katalin Mády

Institute of Phonetics and Speech Processing, University of Munich, Germany
mady@phonetik.uni-muenchen.de

Abstract

Vowel quantity is distinctive in Hungarian for all vowels. However, the quantity opposition in high vowels has become unstable in colloquial speech. Possible reasons are the low functional load of this distinctive feature and the lack of supporting acoustic cues such as spectral differences between long and short high vowels. According to Ohala's model of sound change, the quantity neutralisation process in the production of high vowels might go back to a perceptual loss. This was tested in a forced-choice categorisation experiment for high, mid, and low vowels with different age groups. Our results show a decreased sensibility for high and also mid vowels, but only in unstressed position.

1. Introduction

Hungarian is a language with distinctive vowel quantity in both stressed and unstressed syllables. However, the phonetic, phonological and morphophonological behaviour of the seven quantity-based vowel pairs /i iː y yː u uː o oː ø øː ɛ eː ɒ aː/ is not identical. First, there is a substantial spectral difference between short and long /a/, and also /e/ (called low vowels by [1]), whereas the short high vowels /i y u/ do not differ systematically from their long counterparts with respect to their spectral cues [2]. Second, vowels differ in their functional load: there is a large amount of quantity-based minimal pairs with low vowels such as *bal* /bal/ 'left' *vs* *bál* /baːl/ 'party, ball', but there are hardly any with high vowels (e.g. *zug* /zug/ 'niche' *vs* *zúg* /zuːg/ 'to sweep'). Third, there is a difference in the morphophonological behaviour of high and low vowels: while all vowels participate in systematic vowel shortening processes such as the Final Stem Vowel Shortening rule, only low vowels are involved in phonological lengthening processes [1]. Thus, quantity as a distinctive feature is functionally more crucial, but also more easily perceivable for low vowels than for high ones. The mid vowels /o ø/ show an intermediate behaviour: short vowels tend to be more centralised, thus the quantity distinction is connected to quality differences to some extent. There is also a larger amount of minimal pairs with these vowels than with high ones. Investigations such as [3] show a gradually growing stability of vowel length perception from high over mid to low vowels.

In the past decades, an increasing loss of this feature has been observed for high vowels, especially in unstressed syllables and in colloquial speech [1]. The neutralisation tendency was described as an ongoing sound change process in [4]. In this study, the duration of stressed high vowels was compared to that of unstressed high vowels in different age groups (N.B. stress is always word-initial in Hungarian). Durations of long vowels were close to or identical with that of the short ones in speakers younger than 30 years. No such tendency was found for the group older than 70 years, whereas speakers around 50 showed a smaller degree of shortening and a substantial variation.

Ohala [5] claims that sound change can arise if listeners fail to distinguish between two phonological categories perceptually. Although the original model is based on coarticulatory compensation, Hungarian high vowels might be another example for a sound change process that is triggered by perceptual loss. Given that high vowels cannot be reliably distinguished by their spectral characteristics, it is possible that at some point, listeners began to misinterpret short vowels as long ones or vice versa, leading to a quantity neutralisation in perception.

If the perceptual loss is still an ongoing process, it is to be expected that old speakers can distinguish better between long and short high vowels than younger ones. Besides, it is possible that the quantity loss observed for high unstressed vowels has spread over to other vowels such as stressed high or unstressed mid ones. Therefore a categorisation task was performed with speakers of different age.

2. Material and methods

2.1. Experimental design

Due to the small number of quantity-based minimal pairs involving high vowels, categorisation had to be triggered in a different way.

Twelve trisyllabic target words were chosen in which the target vowel /a aː o oː u uː/ was followed by /d/ and preceded by another alveolar consonant (/d/, /s/, or /l/). The preceding consonants for a vowel pair such as /o/ and /oː/ were always identical, thus, target syllables differed only by the quantity of the vowel.Target words had the structure CVCVCVC and contained only singleton consonants (no geminates or clusters) and only short vowels apart from the target vowels. Short and long equivalents were matched for word frequency, and none of the words was extremely common or seldom.

Stressed target syllables were word-initial, unstressed ones word-final. Target words were embedded in short meaningful sentences with low semantic weight so that none of the target words could be favoured above the other. Stressed syllables were preceded by exactly one content word that did not contain the target vowel (e.g. *A kertben* ***dud****orok.* *vs* *A kertben* ***dúd****olok.* 'In the garden, there are humps' *vs* 'In the garden, I am humming.'). Identically, unstressed targets were followed by an unstressed function word plus a content word whose meaning was loosely associated with both target words (e.g. *A fona****lad*** *egy*

*élmény. vs A fuvo**lád** egy élmény.* 'Your wool is an experience.' *vs* 'Your flute is an experience.'). The sentence structure did not allow for a prosodic boundary before or after the target syllable. The sentence list was recorded in a sound-proof room with several repetitions, spoken by a trained female speaker of Hungarian. Target vowels were cut from the carrier sentence, and their duration was manipulated using Praat's PSOLA algorithm. The first 25% and the final 25% of the vowel duration was left unchanged (thus, formant slopes maintained their original shape), and the central 50% of the vowel was manipulated, resulting in 13 vowel segments in 10 ms steps. /u/ ranged from 50 to 170 ms, /o/ from 60 to 180 ms, and /a/ from 70 to 190 ms, reflecting the different intrinsic durations of these vowels. Durations were kept identical across stressed and unstressed vowels, since lengthening is usually not regarded as a relevant cue for stress in Hungarian [6, 7]. The procedure was applied to both originally short and long vowels, leaving their spectral characteristics unchanged.
It cannot be excluded that f0, at least on the accented syllable, provides cues about the quantity of the vowel. Therefore, manipulated short and long vowels were synthesised with identical f0 curves. Due to the prosodic structure of the sentences, both accented and unaccented syllables were produced with an f0 fall on the target syllable. (The last pitch accent in read sentences is normally a falling one, see [8].) Unstressed target syllables were never followed by a strong pitch accent or a phrase boundary in the sentences and were slightly falling. The fundamental frequency for each member of a vowel pair (stressed /o/ vs. stressed /oː/ etc.) was measured at the beginning and the end of the vowel. The f0 contour of the short vowel was modified such that both the first and the last f0 value fell between the original values for the two segments. Then f0 was decreased evenly between the two endpoints. The f0 contour was stretched to the length of the original long vowel, and the stepwise duration manipulation was performed for both the short and the long vowel. Thus, a vowel pair contained 13 vowels with the spectral structure of the short vowel, and 13 vowels with identical f0, but with the spectral characteristics of the long vowel. This procedure resulted in 156 manipulated vowels (three vowel types, two quantities, stressed *vs* unstressed condition, with 13 durations for each). The manipulated vowels were embedded into the original sentences that were interrupted by the noise of a waterfall, for stressed targets after the consonant following the target word, for unstressed ones before the preceding consonant. Thus, stimuli were either of the form *A kertben **dud*** + noise or noise + ***lad** egy élmény*. The waterfall noise was taken from a real waterfall sound and did not sound unpleasant or disturbing.

2.2. Methods

48 native Hungarian listeners without speech or hearing loss participated in a forced-choice categorisation task. They represented three groups with 16 participants, respectively: a young one (around 20 years), a mid one (around 35 years), and an old one (above 50 years).
The experiment run on a laptop in a Java environment, equipped with a high-quality headphone. Participants listened to the stimuli by clicking on a loudspeaker symbol and at the same time, they saw two sentences which contained the short and long equivalent of the target vowel (e.g. /u/ and /uː/ in stressed position). Their task was to decide which sentence they have heard. They could listen to each sentence twice. Stimuli were presented in five blocks, each of which included all stimuli in a fully randomised order.
Data were fitted to a logistic regression model, using generalised linear mixed models. Thus we obtained an intercept *k* and slope *m* for the S-shaped regression curves. The inflection point at 50% of each fitted curve was then calculated as $-k/m$.

3. Results

3.1. Vowel class

A series of repeated measures MANOVA tests was performed pairwise for all short and long vowels, both for the inflection point and the slope *m*. Significance level was set at p=0.05.
Inflection points and *m* values differed significantly for each short and long pair. After the Bonferroni correction was applied, the difference for stressed /o/ (p=0.010) and unstressed /u/ (p=0.043) did not reach significance level any more.
As shown in Figure 1, listeners relied mainly on the spectral characteristics when identifying long and short /a/. This is in line with [3] where a stable categorisation of low vowels was found. Interestingly, identification was influenced by duration for unstressed short /a/ to some extent, especially in the old group, whereas the identification of long /aː/ was stable throughout both conditions and all age groups.
Figure 2 shows the binomial regression curves for stressed and unstressed /o/. Towards the ends of the duration scale (close to steps 1 and 13), listeners rely mainly on durational cues. However, the difference between the inflection points for short and long vowels is relatively large (about 40 ms for stressed, 30 ms for unstressed vowels). This means that if the duration is ambiguous, spectral cues are utilised as additional quantity markers.
The difference between the inflection points for both stressed and unstressed /u/ is smaller than for /o/, indicating that spectral cues are even less relevant for quantity identification. While the relatively steep slope of the stressed vowel, comparable to that of /o/, indicates that the perception of quantity is categorical for this sound, the flat slope of the curve for unstressed /u/ is a sign of gradual rather than categorical perception of quantity for this sound.
In order to test the impact of stress on category boundaries, a repeated-measures MANOVA was performed on all data with stress, vowel class, and vowel quantity as independent variables and inflection point as dependent variable and participant as a within-subject factor. The interaction of vowel quantity and stress is significant (p=0.043). This indicates that stressed vowels have to be longer in order to be perceived as long than unstressed ones. This is far from being obvious in Hungarian where lengthening is usually not regarded as a consistent cue for stress, as mentioned above.

3.2. Age

For the analysis of the differences between age groups, the difference between equivalent inflection points was calculated for each subject in each condition (inflection point for short unstressed /u/ minus inflection point for long unstressed /u/ for participant 1, etc.). A larger difference indicated that the spectral characteristics of the long and short vowel had a larger impact on quantity categorisation, a smaller difference pointed to

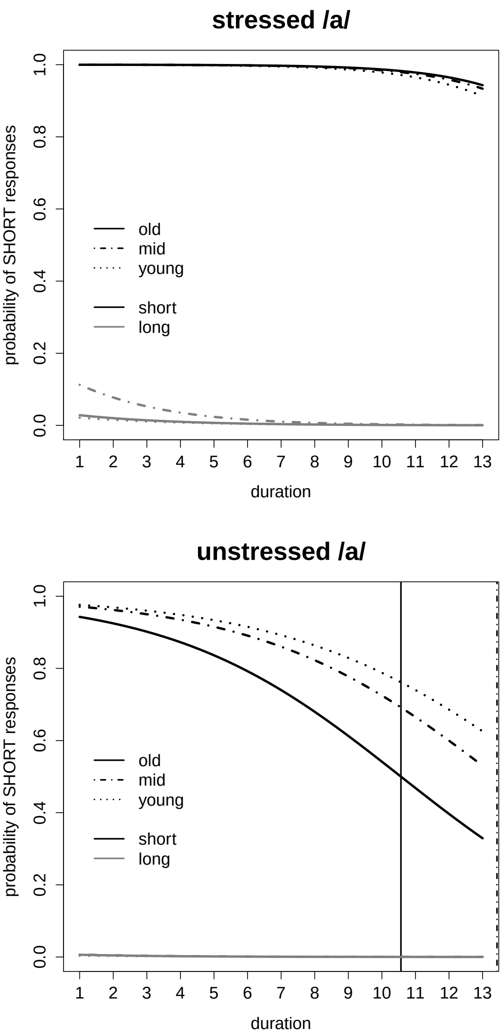

Figure 1: *Logistic regression curves for stressed and unstressed* /a/*, number of short responses over duration increasing in 13 steps. Vertical lines mark the inflection points for each group.*

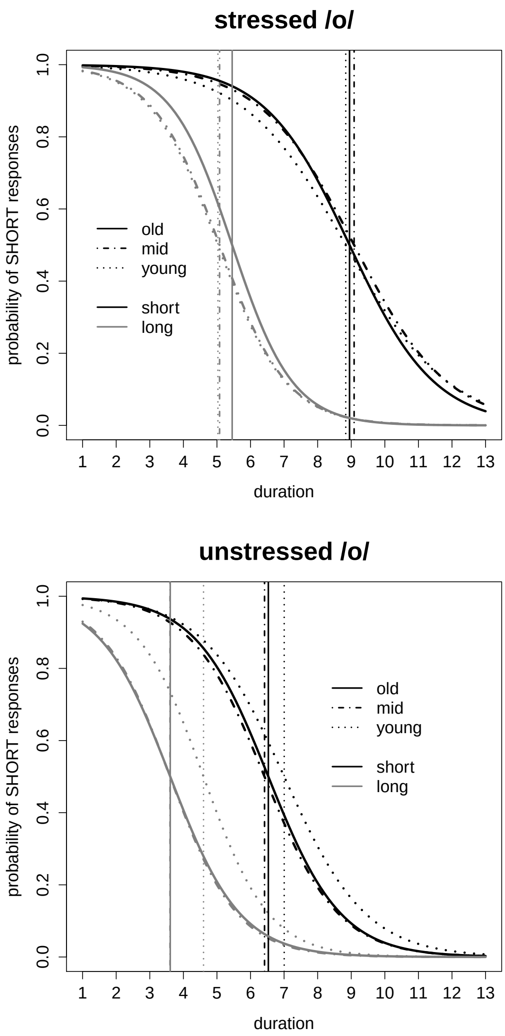

Figure 2: *Logistic regression curves for stressed and unstressed* /o/*, number of short responses over duration increasing in 13 steps. Vertical lines mark the inflection points for each group.*

a smaller influence of spectral cues on quantity categorisation.

For the /a/-vowels, inflection points fell outside of the range of 1 and 13, indicating that decisions cannot be modelled as a logistic regression curve. This means that quantity decisions were not substantially influenced by duration. Thus, /a/ vowels were excluded from further analysis. Results for /o/ and /u/ are shown in Figure 4.

Again, a repeated-measure MANOVA was conducted for the stressed and unstressed vowels, respectively. When all three age groups were involved, no significant difference was found for any condition. However, if the mid group was eliminated, the difference was significantly greater for the old group in the unstressed condition (p=0.012), but not for stressed ones (p=0.65).

4. Discussion

The results of the perception experiment presented above show that vowel quantity is perceived differently in low, mid, and high vowels in Hungarian. For the perception of low vowels, spectral cues are more salient than durational ones, while for high vowels, spectral cues play a very limited role. Mid vowels show an intermediate behaviour: listeners ignore spectral information if duration is non-ambiguous, but they rely on spectral characteristics when duration alone is insufficient for categorisation. The same tendencies were observed for unstressed vowels, although categorisation was less clear throughout all vowel classes.

The results support the assumption that quantity neutralisation for unstressed high vowels is also reflected in an ongoing perceptual loss for the same vowel class. At the same time, categorisation of unstressed mid vowels is also becoming weaker at present. It is still to be tested whether this loss is connected to

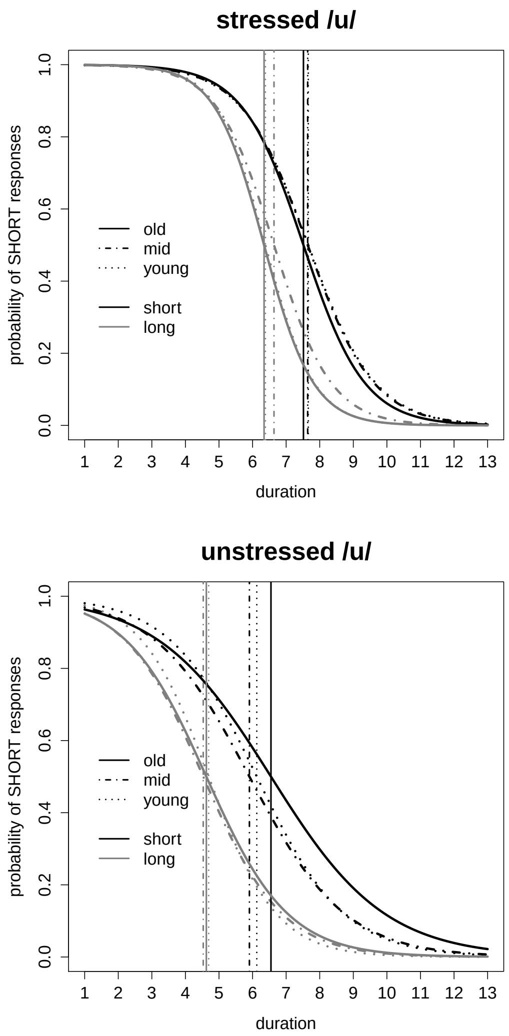

Figure 3: *Logistic regression curves for stressed and unstressed* /u/*, number of short responses over duration increasing in 13 steps. Vertical lines mark the inflection points for each group.*

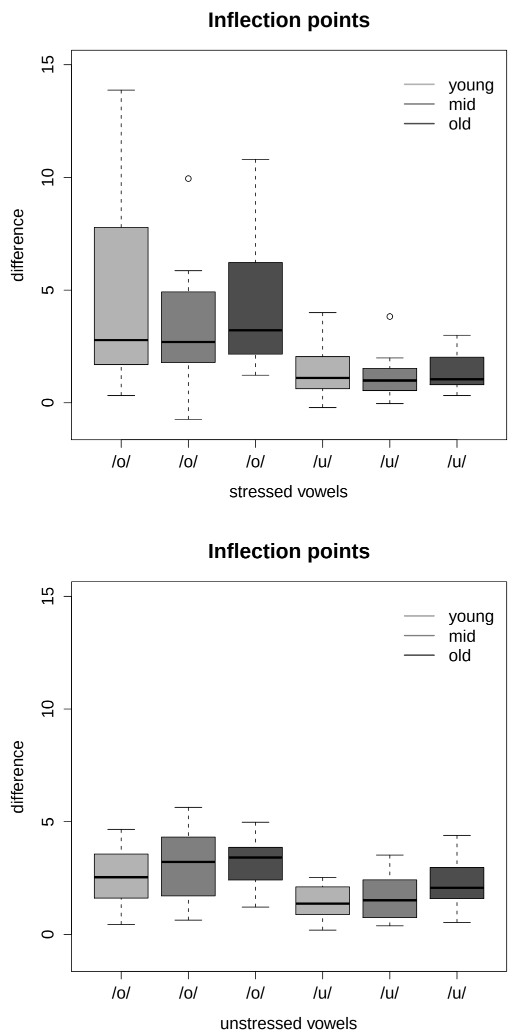

Figure 4: *Difference between inflection points for equivalent short and long vowels.*

a current neutralisation process in the production of unstressed mid vowels.

5. Acknowledgements

This research was part of the excellence project *Speaker characteristics and speech community* supported by the German Research Council. I would like to thank to Christoph Draxler, IPS Munich, for setting up the experimental tool.

6. References

[1] Siptár P. and M. Törkenczy. 2000. The phonology of Hungarian. Oxford: University Press.

[2] Magdics K. 1965. *A magyar beszédhangok akusztikai szerkezete [Acoustic structure of Hungarian speech sounds].* Nyelvtudományi Értekezések. Budapest: Akadémiai Kiadó.

[3] Mády K. and U.D. Reichel. 2007. Quantity distinction in the Hungarian vowel system–just theory or also reality? *Proc. 16th International Congress on Phonetic Sciences*, Saarbrücken, Germany. 1053–1056.

[4] Magdics K. 1960. A szóvégi magánhangzók rövidülése a köznyelvben [Shortening of word-final vowels in colloquial speech]. *Nyelvtudományi Közlemények* 62. 301–324.

[5] Ohala J.J. 1981. The listener as a source of sound change. In Masek, C.S., R.A. Hendrick and M.F. Miller (eds.). *Papers from the parasession on Language and behavior.* Chicago: Chicago Linguistic Society. 178–203.

[6] Fónagy I. 1958. *A hangsúlyról [On stress].* Nyelvtudományi Értekezések. Budapest: Akadémiai Kiadó.

[7] Kassai I. 1979. *Időtartam és kvantitás a magyar nyelvben [Duration and quantity in Hungarian].* Budapest: Akadémiai Kiadó.

[8] Mády K. and F. Kleber. 2010. Variation of pitch accent patterns in Hungarian. *Proc. 5th Speech Prosody Conference*, Chicago, USA. 100924:1-4.