QUANTITY DISTINCTION IN THE HUNGARIAN VOWEL SYSTEM—JUST THEORY OR ALSO REALITY?

Katalin Mády (1,2) and Uwe D. Reichel (1)

(1) Institute of Phonetics and Speech Processing, University of Munich, Germany
(2) Institute of German Studies, Pázmány Péter Catholic University, Hungary
mady|reichelu@phonetik.uni-muenchen.de

ABSTRACT

According to most current theories, the Hungarian vowel system involves 14 vowels that correspond to seven vowel pairs, each differentiated by quantity. However, there are phenomena both on the phonological and the phonetic level which suggest that for low, mid, and high vowels a separate evaluation of the quantity opposition is necessary. In order to test this, we conducted a perception test, in which embedded and isolated vowels spoken by a native Hungarian speaker were to be identified by native listeners.

The results show that the perception of vowel length and vowel quality (i.e. the formant structure) closely interacts in Hungarian. Low vowels, for which short and long realisations differ in quality, i.e. in vowel height, were seldom identified incorrectly. For embedded high vowels, duration was not obviously regarded as a crucial cue for identification by the subjects, nor were they clearly differentiated by the speaker. Mid vowels showed a mixed behaviour: they were differentiated regarding their duration and formant structure in production, however, this information was only partly used by the listeners. The fact that vowel quantity distinction in Hungarian is only maintained where there is a perceivable quality difference shows that the role of quantity is not as dominant as it has been regarded for long.

Keywords: Hungarian, vowel system, perception, quantity, quality.

1. INTRODUCTION

Hungarian contains 14 vowels which are generally regarded as seven quantity pairs of the classes /i/, /y/, /u/, /e/, /ø/, /o/, /a/. (In this paper, we refer to vowel classes by the quality of the long vowel.) Several arguments support the assumption that quantity has a primary distinctive function:

(1) There are minimal pairs with all vowel classes where distinction is achieved by vowel quantity: *hal* – *hál* (‘fish’ – ‘he sleeps’), *kor* – *kór* (‘age’ – ‘disease’), *birtok* – *bírtok* (‘property’ – ‘you bear’).
(2) Short and long vowels can occure in open or closed syllables. Thus, vowel length (= quantity) is an independent distinctive feature and not part of a complementary distribution.
(3) The orthography is based on the assumption that long and short vowels correspond to each other (long vowels are marked by an accent: <í ű ú é ő ó á>).

On the other hand, some vowel pairs differ largely in their phonetic realisations, while others (almost) do not. Thus, unrounded central low /aː/ has a back mid-low or low counterpart that is usually regarded as round mid-low /ɔ/ by Hungarian phoneticians, and mid-high front /eː/ goes together with mid-low front /ɛ/.

The /a/ and /e/ classes are often classified as **low** [8], while mid-high /o/ and /ø/ are regarded as **mid** as opposed to the **high** vowels /i/, /y/ and /u/.

There is no general agreement as to what extent mid and high vowels differ in quality according to their length. Past and current research has led to considerably different conclusions: some authors claim that all mid and high vowel pairs are realised differently, others have not found any remarkable difference in any of these classes, while a third group of authors underline the interconnection between degree of opennes (i.e. high vs. mid) and quality difference: high vowels differ less in quality than mid vowels. (For an overview, see [3].) In a recently developed Hungarian speech synthesis system, long mid and long high vowels were generated from their short counterparts [6].

On the other hand, the distribution of vowel length is restricted in some positions. Firstly, word final /o/ and /ø/ are always long. Secondly, in Educated Colloquial Hungarian (but not in orthography and Elaborated Hungarian!), vowel quality is predictable for high vowels in most positions, with the excepction of stressed nonfinal syllables [8].

Another difference between the categories high, mid, and low is observed when the vowel is in mor-

pheme final position. Low vowels are always lengthened when they are followed by a suffix, e.g. for plural: *fa – fák* ('tree(s)'), *este – esték* ('evening(s)') etc. (N.B. Hungarian is a so-called agglutinating language, thus, most grammatical relations are expressed by suffixes.) This rule does not apply to word final high vowels: *kapu – kapuk* ('gate(s)'), *bébi – bébik* ('baby' – 'babies') etc. As short mid vowels never occur word finally, the above rule is not relevant for this category.

Despite the counterarguments above, most authors (both phonologists and phoneticians) agree that quantity is a primary distinctive feature in the Hungarian vowel system. However, Kassai [2] and Kovács [3] point out the necessity of a distinction according to vowel height. Furthermore, Kassai refers to Sauvageot (1964) where the latter claims that quantity distinction is disappearing from Hungarian and is only maintained for vowel pairs that are distinct in their quality. In all other cases, vowel length is neutralised.

The distinctive role of vowel length in Hungarian has only partly been confirmed by recent perception experiments. While Gósy [1] found an obvious influence of duration on vowel length decisions in her synthesised material, Kovács [3], who used embedded vowels uttered by a human speaker, came to the conclusion that the perceptive boundaries of vowel length are clearly connected to duration under and above a given cut-off value, while durations between these boundaries are mainly influenced by the formant structure.

In a previous paper, we have presented a comparison between German and Hungarian vowels [5]. Among others, we investigated vowel duration of short and long vowels in both languages. We found that Hungarian short and long vowels differ far less from each other in their duration than German ones. This result was unexpected, as in German vowel length differences are mostly accompanied by a difference in vowel quality (tenseness), thus, one would expect that vowel duration does not play a central role in this language. On the other hand, in a language like Hungarian that is said to have a quantity-based vowel system, one would expect clear duration differences.

In the present paper, we will investigate the perceptive role of vowel duration and vowel quality (i.e. the formant structure). We will seek an answer to the question (1) whether there is an interaction between vowel length and vowel quality, (2) whether identification by listeners is related to the acoustic features of these categories, and (3) whether high, mid, and low vowels trigger different perceptional behaviours. Another paper in this volume describes a different approach to the same data [7], in which human perception is being compared with machine learning based on decision trees.

2. MATERIAL AND METHODS

The experiment was based on a corpus in which all the vowels of Hungarian were uttered between voiced or voiceless velar consonants in a carrier sentence. All tokens were repeated ten times by a 21-year-old male speaker of Hungarian. The data we used were originally recorded as part of a former articulatory study with electromagnetic articulography, the outcomes of which will be related to our present findings in a following paper.

37 native speakers of Hungarian (14–44 years, 10 males, 27 females) participated in the study. In the first part, they were asked to identify the vowel in nonsense words with the structure /gVgɔ/ or /kVkɔ/. They were given the possibility to listen twice to each stimulus, then they had to choose one of the 14 vowels displayed on the screen. In the second part, they had to listen to isolated vowels which all had the same length (40 msec, corresponding to the shortest duration in the data set) and were weighted by a Tukey window (taper sections each set to 5 msec). The subjects were told that the second part of the experiment was based exactly on the same vowels as they had listened to in the first part, i.e. they knew that half of the items originated from long vowels.

The analysis of perception of embedded and isolated vowels was based on several acoustic measurements performed in Praat: duration of the first and second vowel and of both consonants (as absolute and relative values), pitch, and the first three formants in Hertz and in Bark (the median was calculated with a window size of 20 msec and a time step of 5 msec). In the present paper, duration of the first (= target) vowel and its F1 and F2 will be discussed in detail. The impact of temporal structure, pitch, and consonant voicing are discussed in [7].

3. RESULTS

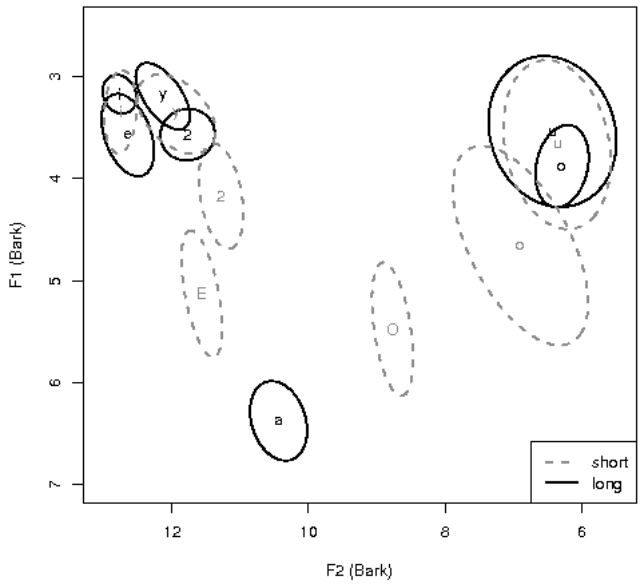
3.1. Vowel duration and formant structure in production

Duration measurements showed that in production, vowel length was mostly associated with durational differences, thus, long vowels had an overall longer duration than short vowels. Long and short /a/ were kept apart by a gap between 100 and 130 msec. For /o/, /ø/, and /e/, a duration of 100 msec could belong to both vowel quantities, all other short vowels being shorter and vice versa. On the other hand, high vowels showed a relatively large overlap between 70 and

120 msec where both short and long vowels could appear.

The formant structure revealed an interesting distribution (Fig. 1): the low vowels /aː/, /ɔ/, /ɛ/ and the short mid vowels /o/ and /ø/ clearly differed from each other and from all other vowels. Long mid front and long high front vowels were located closer to each other, but there was little or no overlap between /iː/ and /eː/, /yː/ and /øː/, respectively. However, high short vowels were distributed over the area of both neighbouring vowels, that is to say, /i/ could not be differentiated from /iː/ and /eː/, etc. To put it more generally, while both long mid and long high front vowels were produced by the speaker with relatively little variation and were kept apart regarding their formant structures, high short vowels were characterised by large variation and a less specific formant structure. This is also true for short /u/, while the formant values for /oː/ overlap those for /uː/.

Figure 1: Short and long vowels in the articulatory space. Vowel quality is given in SAMPA.



3.2. Perception of isolated vowels

Vowel identifications were summarised for all subjects to each vowel class (see image file 1, which contains a confusion matrix with all data). There was a strong tendency to identify long vowels as their short counterpart. This tendency was not found for low vowels: /aː/ was identified correctly in 93.9%, as opposed to mid-high /eː/ with only 39.7% correct identifications. Thus, while the short stimulus duration was probably misleading for the subjects, the decisions were not exclusively based on duration.

The tendency to perceive originally long vowels as short ones can often be explained by the distribution of vowels in the articulatory space: as the formant structure of /eː/ partially overlaps the formant structure of /i/, it is not surprising that subjects tended to identify /eː/ as /i/. The same was true for /iː/, /øː/, /yː/, and /uː/.

It is somewhat surprising, however, that the long mid vowels /oː/ and /øː/, the short realisations of which obviously differ from the long ones in their formant structure, were identified correctly in less than 20% of the cases. In order to test whether other factors than F1 and F2 were responsible for these misidentifications, additional t-tests were performed based on delta formants (F3–F2, F2–F1, F1–F0) for long mid vowels that were identified as a long, or as a short mid vowel. However, none of the differences was significant (given $\alpha = 0.05$).

3.3. Perception of embedded vowels

In the embedded context, subjects were provided with the entire CVCA sequence. As shown in the image file 2, the low vowels /ɛ/, /aː/ and /ɔ/ were correctly identified in more than 98% of the cases. On the other hand, /eː/ was almost never misheard as /ɛ/ but quite frequently as /iː/ (17.3%).

Mid vowels showed a specific behaviour: while erroneous identifications of the short segments were mostly substitutions by their long counterpart (/oː/ for /o/: 4.8%, /øː/ for /ø/: 11.1%), long vowels were seldom misheard as short vowels, they were rather perceptually shifted upwards, especially in the case of /eː that was most often misheard as /iː/.

The same tendency was observed for high vowels, where the proportion of incorrectly identified vowels in embedded context was larger. The correct identifications of short high vowels ranged between 70% and 75%, and in case of an erroneous identification, they were most often heard as their long counterparts (around 18%). Long /iː/ and /yː/ were less error-prone, while the proportion of correct identifications of /uː/ was again relatively low. In the case of false identifications, long mid vowels were preferred for /iː/, /yː/, and /uː/. This tendency was independent of the formant structure and of duration (for other factors, like voicing of the flanking consonants and F0, see [7].

4. DISCUSSION

The results presented above reveal a strong correlation between length and quality in the Hungarian vowel system. Both vowel production and perception seem to rely on different patterns dependent on vowel height. Listeners were able to identify the low

vowels /aː/, /ɔ/, and /ɛ/ with ease, both in embedded and isolated experimental settings. In production, long and short /a/ were distinguished by an obvious difference in duration and in the formant structure. The opposite can be said about the high vowels /i/, /y/ and /u/: short and long vowels had partly overlapping durations, and their identification was not possible on the basis of F1 and F2 (though the distribution was not identical: formants for long high vowels varied much less than for their short equivalents). The patterns observed for mid vowels can be located between those for low and high vowels. While long and short vowels only slightly overlapped regarding their durations and clearly differred in their formant structures, none of this information was necessarily interpreted by the listeners. The mixed class /e/ (short low /ɛ/ and long mid /eː/) revealed a dual behaviour: in production, the two sounds reflected the patterns observed for /o/ and /o/ (little overlap of duration, distinct formant structure), while /ɛ/ was much better identified in both the embedded and isolated context, than was /eː/. As opposed to the mid vowels /o/ and /ø/, /eː/ and /ɛ/ were not interchanged in perception.

These findings support that the role of quantity in the Hungarian vowel system cannot be defined in general, it should be rather regarded according to vowel height. The idea proposed by Sauvageot, that the quantity distinction is loosing its importance in favour of the quality opposition, cannot be fully rejected. It is possible that the Hungarian vowel system is in a changing state. This could explain the transitional behaviour of the mid vowel classes.

The results raise two further questions of a more general nature. Firstly, there seems to be a strong tendency across listeners to hear short vowels as long, but not vice versa. Moreover, the misidentification of long vowels as another long vowel was not enhanced by the formant structure of the data used here (as no overlap for any of the long vowels was found). Clearly, the tendency cannot simply be ascribed to undershoot, at least not based on these data.

Secondly, there is a discrepancy between the production and the perception of the mid vowels /o/ and /ø/. While the speaker clearly produced short and long vowels differently, the listeners did not use this information for their decisions. This finding is contradictory to Lindblom's hyper- and hypoarticulation theory, which states that speakers make as much effort to speak clearly as is required for maintaining the auditorily relevant difference [4]. However, native Hungarian listeners do not seem to be aware of the quality difference between short and long /o/ and /ø/. This contradiction might be based on the articulatory variance of these sounds through different speakers, or on the secondary role of the quality difference for these sounds. As the next step, we will investigate the production of several speakers in order to provide a more detailed analysis of these questions. At the same time, a perception experiment based on different age groups could reveal more information about the question whether the Hungarian vowel system is undergoing a change at present.

5. ACKNOWLEDGEMENTS

We would like to thank to Christian Geng, ZAS, Berlin, for providing us with the acoustic data. The research activity of the first author has been supported by the Alexander von Humboldt Foundation, Germany.

6. REFERENCES

[1] Gósy, M. 2002. Magánhangzók változása az idő függvényében. In: Hunyadi, L., (ed), *Kísérleti fonetika – laboratóriumi fonológia*. Debrecen: Kossuth Egyetemi Kiadó 7–20.

[2] Kassai, I. 1979. *Időtartam és kvantitás a magyar nyelvben*. Budapest: Akadémiai Kiadó.

[3] Kovács, M. 2002. *Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában*. PhD thesis Debreceni Egyetem.

[4] Lindblom, B. 1990. Explaining phonetic variation: a sketch of the H and H theory. In: Hardcastle, W. J., Marchal, A., (eds), *Speech Production and Speech Modelling*. Dordrecht: Kluwer 403–439.

[5] Mády, K., Tronka, K. Z., Reichel, U. D. 2005. Syllable cut and energy contour in vowels: a comparative study in German and Hungarian. *ZAS Papers in Linguistics* 42, 197–217.

[6] Olaszy, G. 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben*. Budapest: Akadémiai Kiadó.

[7] Reichel, U. D. 2007. Comparing human and machine vowel classification. *Proc. 16th ICPhS Saarbrücken* (accepted).

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