The acoustics of yer and non-yer vowels [e] and [o] in Slovak

S. Benus[a] and M. Rusko[b]

[a]Constantine the Philosopher University, Stefanikova 67, 94901 Nitra, Slovakia
[b]Institute of Informatics of the Slovak Academy of Sciences, Dubravska cesta 9, 845 07 Bratislava, Slovakia
milan.rusko@savba.sk

Yers of Slavic languages are vowels that alternate with zero and historically developed from high short vowels [ĭ] and [ŭ]. In Slovak, both front and back yers were preserved and surface as [e] and [o] respectively. For example, the second [o] in *kotol* 'cauldron-Nom-Sg.' is a yer because it disappears with adding a suffix vowel: *kotla* 'cauldron-Gen-Sg.'. Compare with a non-yer [o] in *kostol* 'church-Nom-Sg.', *kostola* 'church-Gen-Sg.'. Traditional phonological accounts of this difference assume that yer vowels are underlyingly different from non-yer vowels and merge with [e] and [o] through a phonological process. Therefore, these accounts predict that yer and non-yer vowels should be phonetically identical since they enter the phonetic component already merged as [e] or [o].

The results of our acoustic experiments show that yer vowels are phonetically different from non-yer vowels. The most salient differences were observed in the first formant and duration: yers have lower F1 than non-yers, and for some subjects they are also shorter. This finding supports the view that the phonetic component has access to deep phonological alternations and that phonetics-phonology is a single cognitive system in which the components have different granularities and interact bidirectionally.

1 Introduction

Slovak, together with other Slavic languages, developed a phonological system in which the presence of mid vowels /e/ and /o/ sometimes alternates with their absence. For example, the second [o] in *kotol* 'cauldron-Nom-Sg.' disappears with adding a suffix vowel: *kotla* 'cauldron-Gen-Sg.' and not **kotola*. Compare this with [o] in *kostol* 'church-Nom-Sg.', that remains even if the suffix vowel is added, *kostola* 'church-Gen-Sg.', and not **kostla*. Note that word-medial cluster [stl] is well-formed in Slovak: *rástli* 'they grew' *nestlacil* 'he did not press'. Vowels that alternate with zero historically developed from high short vowels [ĭ] and [ŭ], and are traditionally called yers. In Slovak, both front and back yers were preserved and surface as [e] and [o] respectively. Hence [o] in *kotol* is analyzed as a yer vowel because it alternates with zero while [o] in *kostol* is a non-yer vowel.

Traditional phonological accounts of this difference [10] assume that yer vowels are underlyingly different from non-yer vowels. Hence, Slovak vowel system contains both 'regular' /e/ and /o/ as well as /e/ and /o/ that derived from yer vowels through so called 'yer-vocalization and subsequent lowering [7, 10]. Crucially, traditional accounts predict that yer and non-yer vowels should be phonetically identical since they enter the phonetic component already merged as [e] or [o].

The underlying assumption of these accounts is that the relationship between the discrete phonological system and the continuous phonetic system is unidirectional: phonology → phonetics. In other words, the cognitive system representing speech first performs discrete phonological computations and then there is a deterministic transducer that can transform the results of such computations into the continuous actions of the body effector system such as articulatory activity.

However, this idea of derivational precedence has been challenged by recent findings. These results lead to proposals that view speech as a unified cognitive system where variation is represented at multiple distinct but inter-dependent levels of different granularities [9]. One such proposal presents the hypothesis that continuous phonetic details are relevant for discrete morpho-phonological alternations [1, 4]. More specifically, articulatory data collected with magnetometry and ultrasound show that the horizontal position of the tongue body during stem-final transparent vowels {[i:], [i], [e:]} in Hungarian correlate with the [±back] quality of the suffix vowel(s) as determined by vowel harmony. We show that this result cannot be attributed to coarticulation and argue that transparent vowels, despite assumptions in the phonological literature, participate articulatorily in vowel harmony. If transparency in vowel harmony is construed as interplay of phonetic and phonological systems, we argue that several recalcitrant and seemingly unrelated generalizations about the data in Hungarian vowel harmony may receive a unified explanation

We set out to test the assumptions of the derivational and alternative models of phonetics-phonology interface discussed above on the Slovak mid vowels /e/ and /o/. If the yer and non-yer vowels in Slovak are in fact different phonetically, it is plausible that this difference is being used by the cognitive system of speech. If, on the other hand, the results show that the yer and non-yer vowels are phonetically non-distinguishable, a closer look at the differences between Slovak mid vowels and Hungarian transparent vowels might reveal insights into the limits of the role of phonetic differences in the cognitive systems.

2 Experiment

We designed a production experiment to test if the yer and non-yer vowels in Slovak are produced differently. Four college students, three females and one male, took part in the experiment in exchange for a course credit.

2.1 Stimuli

We created six pairs of words, three with /e/ and three with /o/ target vowel. The stimuli are shown in Table 1. The crucial phonological alternation of vowels can be seen from the comparison between the singular and plural forms. While the words with yer-vowels do not contain this vowel in the plural, the non-yer words preserve the vowel in both singular and plural forms. In general, a yer vowel is preserved only in those forms that have phonologically zero suffix, such as the Nominative singular case here. Yer vowels are absent in all cases where a suffix vowel is present.

The stimuli pairs are designed in such a way that keeps the consonantal environment as similar as possible within the pairs. Moreover, all target vowels appear in the prosodically identical position: the second syllable of the word. This is an unstressed syllable because Slovak is a language with fixed lexical stress that falls on the leftmost syllable of a prosodic word.

Yer-vowels			**Non-yer-vowels**		
Sing.	**Pl.**	**Gloss**	**Sing.**	**Pl.**	**Gloss**
palec	palc-e	thumb	balet	balet-y	balet
kopec	kopc-e	hill	obed	obed-y	lunch
báseň	básn-e	poem	jaseň	jasen-e	ash tree
kotol	kotl-e	kettle	kostol	kostol-y	church
kapor	kapr-e	carp	nábor	nábor-y	recruiting
štítok	štítk-y	label	prítok	prítok-y	tributary

Table 1 Stimuli pairs used in the experiment. The acute accent on vowels denotes length, the 'hacek' denotes palatal consonants, c = [ts], y = [i]

The stimuli were presented in two blocks. In the first block, each of the 24 words (12 words in singular and plural) were presented as a picture together with the text of the frame sentence: *{Jeden, dva}*______ *{hore, dole}* {one, two} ______ {up, down}. Fig. 1 shows the stimuli prompts for the *kotol* -- *kostol* pair. Hence, there were 48 stimuli sentences. In the second block, the subjects only saw the pictures without the text. In this block, the singular words, that were the target stimuli, were four times more common than the plural words: there were 96 sentences with the singular words and only 24 with the plural. Hence, there were 120 stimuli sentences in the second block. In total, there were 168 stimuli sentences with 120 target singular words. Due to a minor error in the design, three more stimuli sentences were added, making the total number of target sentences 123. Stimuli were randomized within each block.

Fig.1 Example of four stimuli prompts for the pair *kotol-kostol*.

2.2 Procedure

Subjects were seated in front of a laptop screen in a quiet room with the experimenter present. DMDX software was used for presenting the stimuli on the screen [3]. The examples of stimuli prompts are shown in Fig. 1. Within each block, the prompts were shown automatically with 2500ms time window for the subjects to say the sentence. Their speech was recorded and digitized at 44 kHz using Praat [2] and a Solton dynamic microphone fixed in a stand and positioned in front of the subject's mouth. Subjects were asked to say what they see on the screen in a natural conversational style. The design worked smoothly, the errors in the productions were minimal.

2.3 Data labeling and analysis

Praat [2] was used for the labelling and extraction of the data. First, each word of each stimuli sentence was identified, and then standard identification of individual sounds of the target word were applied focusing on the consistency in the labelling. In addition to standard oral markers identifying the onset and offset of the target vowel, we also labelled the beginning of glottal activity for that vowel if the preceding consonant was a voiceless plosive.

After labelling, a Praat automatic routine was designed to extract the acoustic features. We collected the duration of the vowel, and formant values at the mid-point of the vowel and at the ¼ of the total duration time after the onset and before the offset. We will refer to these temporal points as MID-POINT, POINT 1 and POINT 2 respectively. The raw values were then normalized using z-score normalization to allow the statistical analysis when the values of all subjects are pooled together.

SPSS software package was used to performed statistical tests. Formant and duration values served as dependent variables.

3 Results

The main hypothesis tested in this experiment was that yer and non-yer vowels are phonetically identical. We will discuss the results of vowel quality and vowel duration separately.

3.1 Vowel quality

The Anova tests with the formant data from all subjects pooled showed a significant difference between yer and non-yer vowels in the first formant measured at MID-POINT and POINT 1, $F(1, 480) = 4.17$, $p = 0.042$ and $F(1, 480) = 4.03$, $p = 0.045$ respectively. Yer vowels have lower values of the first formant than non-yer vowels. Hence, yer-vowels can be described as phonetically higher than non-yer vowels because the first formant is inversely related to the articulatory height of the tongue body.

The contrast between yer and non-yer vowels was not produced uniformly by the subjects. Two-way Anova showed the interaction between the variables YER and SUBJECT only for POINT 1; $F(1, 482) = 2.65$, $p = 0.048$ for the mid-point, and only tendency was observed for MID-POINT; F(1, 482) = 2.08, p = 0.1. Separate one-way Anova tests for each subject showed that the difference between yer and non-yer vowels was significant only for two of the four subjects, Subject 1: F(1, 119) = 7.86, p = 0.006 for MID-POINT, and Subject 4: F(1, 119) = 7.14, p = 0.009 for POINT 1.

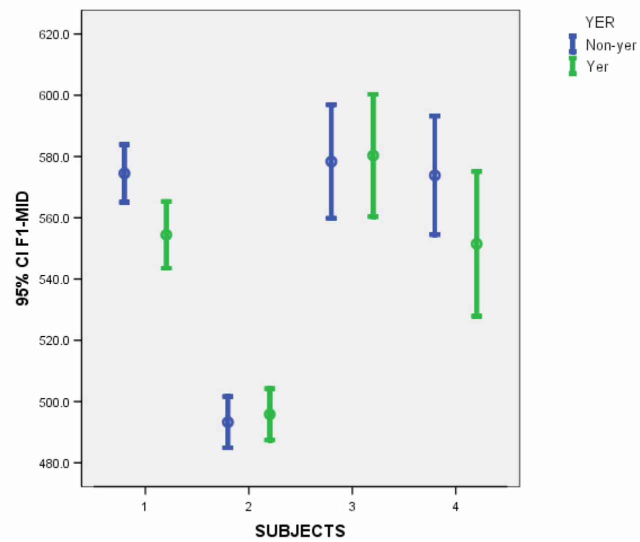

Fig.2 Means and error bars of raw first formant values extracted from the temporal midpoint of the target vowels for all four subjects.

Two-way Anova did not show significant interaction between the variables YER (yer vs. non-yer) and VOWEL (/e/ vs. /o/); F (1, 482) = 1.16, p = 0.282 for the mid-point. Despite this, the two vowels behaved differently. Separate one-way Anova tests for each vowel showed that the difference between yer and non-yer vowels at MID-POINT, was significant for /o/ (F(1, 239) = 8.3, p = 0.004) but not significant for /e/ (F(1, 239) = 0.387, p = 0.535). Similar results were obtained also for the data extracted at POINT 1. Consider the F-value from the test where both vowels were pooled (F = 4.17) and the F-value from the separate test for vowel /o/ (F = 8.3). The comparison of these two values shows that the difference between yer ad non-yer vowels in the general test can be attributed mostly to the difference in vowel /o/.

Given the difference between vowels /e/ and /o/ reported above, and the difference among the subjects, we tested the difference between yer and non-yer vowels also for the four subjects and each vowel separately. Table 2 summarizes the results. It can be seen that only /o/ yer and non-yer vowels were produced differently, and only by two subjects.

	Subject			
Vowel	**1**	**2**	**3**	**4**
/o/	√	X	X	√
/e/	X	X	X	X

Table 2 Significant differences between yer and non-yer vowels separately for each vowel and subject.

Finally, we also looked at the production of target vowels in individual lexical pairs. We already reported that /o/ produced significant differences between yer and non-yer vowels while /e/ did not. Of the three pairs with /o/, two showed significant differences. Of the three pairs with /e/, only one showed a significant difference.

In terms of raw formant values, the significant differences between yer and non-yer vowels were relatively small. They were around 20-30 Hz.

To summarize the results in vowel quality, /o/ vowels that are phonologically considered as yer-vowels were produced with higher position of the tongue by two of four subjects. The production of yer and non-yer /e/ did not vary significantly.

3.2 Vowel duration

Two measures of duration were considered. In the first measure, vowel onset was defined at the release of the oral constriction of the preceding consonant. We will refer to this measure DUR-ORAL. In the second measure, vowel onset was defined glottally at the beginning of voicing. We will refer to this measure DUR-GLOTTAL. These two measures were different for 4 out of 6 stimuli pairs since in these pairs the target vowels were preceded by plosives with measurable positive voice onset time.

The Anova tests with the duration data from all subjects pooled showed a significant difference between yer and non-yer vowels with, DUR-GLOTTAL measure, F(1, 480) = 30.95, p < 0.001. Yer vowels were shorter than non-yer vowels. However, of the four pairs that differ on the two duration measures, voicing of the plosives that precede the target vowel differs in two pairs: *kopec-obed* and *kapor-nábor*. In both pairs, the yer vowel follows a voiceless consonant while the non-yer vowel follows a voiced one. Hence, the DUR-GLOTTAL measure might be affected by this stimuli design. The Anova test with the two mentioned pairs excluded showed that yer vowels were still different from non-yer vowels, but the difference is much smaller, F(1, 319) = 4,19, p = 0.041.

Given the observed differences in the two vowels and four subjects in vowel quality, we also tested these differences in vowel duration. The results show that in some measures, yer /e/ is significantly shorter than non-yer /e/ while /o/ did not vary significantly, and that Subject 3 is the only one with significantly shorter yer vowels than non-yer ones.

In terms of absolute duration, the difference between yer and non-yer vowels was small, not exceeding 15ms, but it was consistently the yer vowel that was shorter than non-yer vowel.

To summarize, only small duration differences were observed. Interestingly, however, they were observed on a different vowel and with different subject than the vowel quality differences.

3.3 Orthography

The design of the experiment also enabled us to test if the presence or absence of text in the stimuli prompts affected the production of the target vowels. But, there were much fewer tokens in the picture-text condition than in the picture-only condition. We hypothesized that the production of target vowels would be more similar if the text was present than in the picture-only mode. This hypothesis was not supported since there were no significant differences in the vowel quality or duration between the two conditions. We noticed, however, that in terms of the first formant, the target vowels in the picture-only condition were produced less varied than in the text-picture condition.

Discussion

The results of our pilot acoustic experiments on Slovak vowels are at odds with the traditional derivational view of phonetics-phonology interface. This view predicts that yer and non-yer vowels are phonetically identical since their phonological difference is neutralized before the result of phonological computation is transferred to the phonetic component. Our results, however, showed that at least for some subjects and some measures, yer and non-yer vowels differ. The most salient differences were observed in the first formant and duration: yer vowels have lower first formant than non-yer vowels, and for some subjects they are also shorter.

A potential explanation for the difference in F1 can be sought in the exemplar model of speech production and perception (e.g. [8]). Following this model, phonological representations encode phonetic details beyond the scope of standard segmental and featural representations. Hence, variability in the production of target mid vowels /e/ ad /o/ might be achieved by averaging and/or randomization over a set of memorized exemplars of the category. For example, if the yer vowels often occurred under the coarticulatory influence of high vowels, the category of yer vowels, defined by the exemplars affected by such coarticulation, might have on average higher F1 value. This higher F1 would persist even in the environments without such coarticulatory influence.

In our experiment, the statistical analysis was based only on the singular words where the target vowels have similar environments (see Table 1). But, in the plural the environment was different because the non-yer vowels were followed by the suffix vowel [i] in five of the six words. Hence, the exemplars of non-yer vowels could be different from yer-vowels due to the coarticulatory influence of the suffix vowel. However, the prediction of such a model is that non-yer vowels should be produced higher because the following [i] is a high vowel. As we saw, our data do not support that prediction since it was the yer, and not non-yer, vowels that were produced higher than non-yer vowels. Therefore, despite its ability to produce sub-phonemic differences, the exemplar model can not explain the variation found in our data.

An intriguing speculation is that yer-vowels /e/ and /o/ preserved some of the phonetic qualities of the original short lax vowels [ĭ] and [ŭ] of Old Church Slavonic. Since these original yer vowels were higher and shorter than /e/ and /o/, the phonetic differences we observed support this speculation.

A similar situation was reported for Hungarian front vowels [i] and [e]. In the 10th century, Hungarian had a contrast between front and back unrounded vowels, which was neutralized in the 11th century [5]. Yet, the remnant of this contrast in the form of small but significant differences in the horizontal position of the tongue body was reported in [1]. Hence, both Slovak and Hungarian seem to display phonetic differences that point to incomplete mergers of phonological categories. And, both languages maintain sub-phonemic differences reminiscent of the contrast assumed to be lost sometimes after the 10th century in both languages.

Interestingly, the two strategies that speakers use to make the yer and non-yer vowels different – vowel height and duration – could be used independently. If the two measures are taken together, three of four subjects produced the target vowels differently.

The difference between /e/ and /o/ is surprising. It is also difficult to explain phonetically since the back cavity is assumed to be smaller than the front cavity, and thus the vertical tongue movements are more limited for back vowels than for front vowels [6]. Hence, phonetically, back vowels should be more prone to phonetically motivated contrast neutralization than front vowels. Yet, our findings suggest that the height contrast is more robustly displayed with the back /o/ than with the front /e/.

However, the phonetic differences we observed are not robust and further research is needed to test the models discussed above. First, there were large individual differences. Two of four subjects showed clear yer vs. non-yer differences while two other subjects produced these vowels in a similar way. Second, the significant difference is largely due to the difference in /o/ and the contrast with yer and non-yer /e/ vowels seem to be phonetically much weaker. Third, duration differences were also significant only for two of four subjects.

One of the limitations of acoustic studies is small reliability of segmenting vowels from liquids. Since many yer vs. non-yer alternations in Slovak involve [l] or [r], precise measurements of the vowel onset and offset becomes very difficult. Currently, we are pursuing articulatory investigation of yer and non-yer vowels. Such a study would allow for better precision in differentiating vocalic and consonantal gestures. Additionally, it would allow the investigation of other aspects of vowel production that acoustics does not allow: kinematic properties of the tongue movement including velocity profiles and the coordination of vocalic gestures with surrounding consonantal gestures. One might imagine that the phonetic difference between yer and non-yer vowels might very well be of a dynamic not static nature, for which an articulatory study is ideally suited.

The significance of this research is in providing more data that could serve for testing the traditional and alternative models of phonetics-phonology interface discussed in the Introduction section. If the yer and non-yer vowels in Slovak are in fact different phonetically, it is plausible that this difference is being used by the cognitive system of speech. For example, sub-categorical differences like these might function as one of the parameters mitigating against a total merger of yer and non-yer vowels, which might de-stabilize the vocalic system in other areas. Potentially, the formal tools of non-linear dynamics developed for the model of Hungarian transparent vowels in [4] can also be employed in a model of this and other incomplete mergers.

If, on the other hand, the results show that the observed acoustic differences are not replicable, a closer look at the differences between Slovak mid vowels and Hungarian transparent vowels might reveal insights into the limits of the role of phonetic differences in the cognitive systems. For example, it might turn out that phonetic differences are used only if they participate in phonological alternations (as we argued is the case in vowel harmony) but not in consideration of the inventory of units that represent the stable characteristics of the phonological systems.

5 Conclusion

In this study we tested the common assumption in the traditional phonological literature that yer and non-yer vowels in Slavic languages are phonetically identical. We designed an experiment in Slovak where subjects produced pairs of words with yer and non-yer vowels /e/ and /o/ in similar consonantal, vocalic, and prosodic environments.

The results of our acoustic experiments suggest that yer vowels are phonetically different from non-yer vowels. The most salient differences were observed in the first formant and duration: yers have lower F1 than non-yers, and for some subjects they are also shorter. We also found that the statistical significance of this overall finding was mostly due to the differences in the production of /o/ and due to the speech of two out of four subjects. Further articulatory exploration of the yer vs. non-yer contrast was proposed.

Our results suggest that the phonetic component has access to deep phonological alternations and that phonetics-phonology should be seen as a single cognitive system in which the components have different granularities and do not necessarily have to interact directionally phonology → phonetics.

Acknowledgments

This work was supported in part by the Ministry of Education of the Slovak Republic, Scientific Grant Agency project number 2/0138/08 Applied Research project number AV 4/0006/07 and by the European Education, Audiovisual and Culture Executive Agency LLP project EURONOUNCE.

euronounce

Education and Culture DG

This project has been funded with support from the European Commission. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

References

[1] Benus, S. & Gafos, A. (2007). Articulatory characteristics of Hungarian 'transparent' vowels. *Journal of Phonetics* 35, 271-300.

[2] Boersma, P. & Weenink, D. (2007) Praat: doing phonetics by computer (Version 5.0.14) [Computer program]. Retrieved March 1, 2008, from http://www.praat.org/

[3] Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. Behavior Research Methods, Instruments & Computers, 35, 116-124.

[4] Gafos, A. & Benus, S. (2006). The dynamics of phonological cognition. *Cognitive Science*, 30 (5), 905-943.

[5] Kálmán, B. (1972). Hungarian historical phonology. In L. Benkő & S. Imre (Eds.) *The Hungarian language* (pp. 49–83). The Hague: Mouton.

[6] Kassai, I. (1979) *Id, tartam és kvantitás a magyar nyelvben.* Budapest: Akadémiai Kiadó.

[7] Lightner, Th. M. (1965). Segmental phonology of contemporary standard Russian. Ph.D. dissertation, MIT.

[8] Pierrehumbert, J. (2001) Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee and P. Hopper (eds.) *Frequency effects and the emergence of lexical structure* (pp. 137-157). Amsterdam: John Benjamins.

[9] Pierrehumbert, J. (2000). The phonetic grounding of phonology. *Bulletin de la Communication Parlee* 5, 7-23.

[10] Rubach, J. (1993). The lexical phonology of Slovak. Oxford: Clarendon Press.