Journal of Phonetics 40 (2012) 535–549

Contents lists available at SciVerse ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Phonetic variation in Slovak yer and non-yer vowels

Štefan Beňuš [a,b,*]

[a] Constantine the Philosopher University, Department of English and American Studies, Štefánikova 67, 94974 Nitra, Slovakia
[b] Slovak Academy of Sciences, Institute of Informatics, Bratislava, Slovakia

ARTICLE INFO

Article history:
Received 3 December 2010
Received in revised form
6 March 2012
Accepted 8 March 2012
Available online 6 April 2012

ABSTRACT

We examine the phonetic characteristics of yer and non-yer vowels in Slovak in an effort to improve our understanding of the link between phonological differences and their phonetic realizations. We test the wide-spread assumption of phonological analyses that yer vowels are phonetically identical to their non-yer counterparts with measures of vowel duration, vowel quality and the patterns of coarticulation with surrounding sounds in both acoustic and articulatory data. Moreover, we compare these patterns with the patterns arising from the variation in speech rate. Our results provide tentative support for the hypothesis that yer vowels in Slovak are phonetically weaker than their non-yer counterparts. The relevance of this observation for the models of phonetics–phonology interface is discussed.

1. Introduction

Yers (sometimes also jers) is a term widely used in the Slavic phonological literature for vowels that alternate with zero (Gussmann, 1980; Lightner, 1965; Rubach, 1993; Scheer, 2006; Szpyra, 1992; Yearley, 1995). Slovak, together with other Slavic languages, has developed a phonological system in which the presence of mid-vowels /e/ and /o/ in some words alternates with their absence. For example, the vowel [o] in *párok* 'sausage-Nom-Sg.' disappears when a suffix vowel is added: *párk-u* 'sausage-Gen-Sg.', *párk-om* 'sausage-Inst-Sg.' and not **párok-u* or **párok-om*. Compare this with the vowel [o] in *nárok* 'entitlement-Nom-Sg.', which remains even if the suffix vowel is added, *nárok-u* 'entitlement-Gen-Sg.', *nárok-om* 'entitlement-Inst-Sg.', and not **nárk-u* or **nárk-om*. Vowels that alternate with zero developed historically from high short lax vowels [ĭ] and [ŭ] of Old Church Slavonic, and in Slovak, both front and back yers were preserved and surface as [e] and [o] respectively. Hence [o] in *párok* is a yer vowel because it alternates with zero while [o] in *nárok* is a non-yer vowel. A sample of words with yer vowels and their alternations with zero in suffixed forms are shown in (1).

(1) Alternations with yers in Slovak.

Nom.Sg.	Transcription	Gen.Sg.	Instr.Sg.	Gloss
palec	[palets]	palc-a	palc-om	'thumb'
laket'	[lakec]	lakt'-a	lakt'-om	'elbow'
pes	[pes]	ps-a	psom	'dog'
kotol	[kotol]	kotl-a	kotl-om	'cauldron'
párok	[paːrok]	park-a	párk-om	'sausage'

To our knowledge, all phonological accounts of yer vs. non-yer paradigms assume that yer vowels are underlyingly different from non-yer vowels (e.g. Rubach, 1993). The formalizations differ – yers as abstract vowels have different featural representation (e.g. the [tense] feature, Gussmann, 1980; Lightner, 1965), they differ from full vowels by the lack of a root node or a melodic specification (Rubach, 1993; Yearley, 1995), or a different government status (Scheer, 2006) – but the presence of an underlying difference is a cornerstone of all accounts. This difference is needed because phonological grammar must be able to target vowels for deletion in forms like *párku*, but not in *nároku*, or alternatively, target vowels for preservation in *nároku* and not in *párku*. Since the phonological system is assumed to operate on categorically discrete representations, o_{yer} must be a different category from $o_{non\text{-}yer}$ in such a system.

The presence of stems like *park* 'park-Nom-Sg.' prevents an otherwise appealing account of vowel epenthesis based on syllabification and coda phonotactics of these alternations (e.g. Szpyra, 1992). Moreover, as pointed out by Rubach (1993) and Scheer (2006), in languages like Slovak with both front and back yers, an epenthesis account is problematic since the type of the putatively inserted yer vowel could not be predicted independently. In more recent Optimality Theoretic accounts, the coda phonotactic restrictions formalized as violable OT constraints play an important role in the generation of the surface forms, yet the assumption of an underlying difference between yer and non-yer vowels remains unchallenged (e.g. Jarosz, 2006; Yearley, 1995).

The second characteristic that is shared among the phonological accounts is the assumption that, once this underlying difference between yer and non-yer vowels has been utilized by the phonological system, the original yer-vowels effectively merge with non-yer vowels into a single vowel category. Again,

* Correspondence address: Constantine the Philosopher University, Department of English and American Studies, Štefánikova 67, 94974 Nitra, Slovakia.
Tel.: +421 37 6408455.
E-mail addresses: sbenus@ukf.sk, sb513@nyu.edu

0095-4470/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2012.03.001

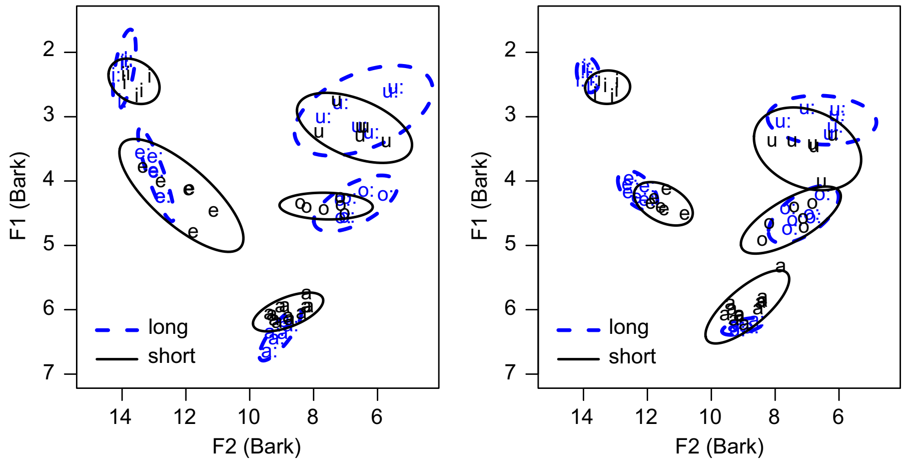

Fig. 1. An example of Slovak vowel inventory for one subject producing pVpa words. Formant space is defined in Bark (x=F2, y=F1), the left panel shows short (black full lines) and long (gray (blue) dotted lines) vowels in stressed syllables and the right panel in unstressed syllables (adapted from Beňuš & Mády, 2010). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

formalizations differ – the rule of lowering (Gussmann, 1980; Lightner, 1965), or linking of unassociated melodies (Rubach, 1993), or high ranking OT markedness constraint (phonetically motivated in Jarosz, 2006) against non-high tense vowels – but the principle of a complete surface phonetic neutralization of yer and non-yer vowels remains. Crucially, all accounts thus predict that yer /o/ in words like *párok* and non-yer /o/ in words like *nárok* should be phonetically identical, because there is only a single phonetic representation of both yer and non-yer vowels. This seems to be a reasonable prediction because the intuition of native speakers is that /o/ vowels in *párok* and *nárok* are the same.

This assumption, however, was to our knowledge never rigorously tested experimentally. The study of alternations involving yers featured prominently in the development of the phonological theory in the past (Non-Linear representations, Lexical Phonology, Government (CVCV) Phonology, Optimality Theory) but virtually no attention was paid to test the crucial assumption of all accounts that yer and non-yer vowels are phonetically identical. A pilot acoustic study compared the production of pairs of yer and non-yer vowels in Slovak (Benus & Rusko, 2008) and suggested that yer vowels might be phonetically different from non-yer vowels. The most salient differences were observed in the first formant and duration: yer vowels had a slightly lower first formant than non-yer vowels, and, for some subjects they were also shorter. An intriguing, yet speculative, analysis of this finding was proposed: yer-vowels /e/ and /o/ might preserve some of the phonetic qualities of the original short high lax vowels [ĭ] and [ŭ] of Old Church Slavonic.

However, in addition to the limited scope, the study of Benus and Rusko had other limitations. First, the results were small in the size, not robust, and the statistical tests neither averaged the data across repetitions or subjects, nor applied a repeated measures design. While problems with the size and the robustness of the effect are to be expected – after all, phonetically trained phonologists never suggested a potential phonetic difference – the limitations of scope and statistical analyses could be addressed. Second, segmenting vowels from liquids [l] and [r] based on the acoustic signal is very challenging. Since many yer vs. non-yer alternations in Slovak involve [l] or [r], precise measurements of the vowel productions become very difficult. Moreover, an acoustic analysis cannot assess kinematic properties, such as the relationship between the velocity and displacement of the gestural movements, that have been shown to be affected by prosodic structure (e.g. Cho, 2006), and also function in differentiating tense and lax vowels, for example in German (e.g. Hoole & Mooshammer, 2002).

It is the goal of the current paper to present the first systematic acoustic and articulatory investigation of yer vowels by comparing them to their non-yer counterparts. Despite great advances in modeling the relationship between phonetics and phonology, our understanding of the extent to which phonetic variability is attributable to the phonological system is still limited. In other words, the nature of the boundary between more granular phonology and less granular continuous phonetics is still an open issue. Recently, several models argued that a thorough understanding of phonetic characteristics leads to better and more coherent phonological explanations (e.g. papers in Hayes, Steriade, & Kirchner, 2004; Gafos & Benus, 2006) while some other proposals argue for the role of phonetics in the diachronic developments but a more modular approach to the phonetics–phonology interface in the synchronic models (e.g. Blevins, 2004; Barnes, 2006). Our contribution to this debate is a thorough examination of the phonetics of a very deep abstract morpho-phonological alternation of yer and non-yer vowels. If the assumption of the phonological accounts is verified, and yer vowels are phonetically identical to their non-yer counterparts, we would have experimental evidence for an area of "self-contained phonology" or modularity in phonetics–phonology interface. If this assumption is not supported, and yer vowels differ from non-yer ones phonetically, the identifications of sources for such difference (in phonology or elsewhere) should lead to better understanding of phonetic variability. In either case, our findings should be useful in seeking cognitive models that formalize the division of labor between discrete-like phonology and continuous phonetics.

1.1. Slovak vowels

Slovak is a West-Slavic language with a five-vowel system of monophthongs /i/, /e/, /a/, /o/, /u/ and a full phonemic quantity contrast for all vowels in all positions in standard colloquial Slovak. Fig. 1 shows Slovak vowel qualities in stressed and unstressed vowels by a single subject (adapted from Beňuš & Mády, 2010).[1]

[1] Low front vowels are not considered in this study since in the speech of most speakers they have merged with mid-front vowels and the contrast survives only in the speech of older speakers of a few dialects.

The primary word stress falls on the leftmost syllable of the prosodic word and rather weak secondary stress is said to fall on every other odd number syllable following the first one counting from the left (Král' & Sabol, 1989). Although traditional literature claims that phonemic quantity and stress placement do not affect the quality of vowels, a recent quantitative investigation of the relationship between the quantity, quality, and lexical stress in Slovak vowels on a limited corpus of 2 subjects (Beňuš & Mády, 2010) showed that shorter vowels (either phonemically short or due to the absence of lexical stress and fast speech rate) are phonetically slightly more centralized than longer vowels.

1.2. *Weakness of yer vowels*

The difference in the behavior of yer and non-yer vowels has been formalized in several ways. Phonologically, yer vowels can be construed as deficient compared to non-yer vowels. This deficiency is demonstrated on several levels. For example, as Jarozs (2006) suggested, Polish underlying yer vowels [ɪ,ɘ] can be considered more marked than their non-yer counterparts [ɛ,ɔ] because the former require simultaneously [+high] and [−tense] or [−high] and [+tense] articulation associated with an antagonistic effect. This is because tongue body raising (formally [+high]) is facilitated by the advancement of the tongue root ([+tense]) and both of these actions also result in F1 lowering. Hence, high vowels paired with advanced tongue positions result in sympathetic effects both articulatorily and acoustically, and high vowels with a more retracted position result in an antagonistic effect (Archangeli & Pulleyblank, 1994).

Furthermore, in some accounts, the underlying specification of yer vowels is deficient compared to their non-yer counterparts in the sense that their specification is not supported on the surface. For example, yer vowels are underlyingly unassociated to the melodic tier while all surface vowels must be associated (Rubach, 1993; Yearley, 1995). Alternatively, they are specified with a [tense] feature that is artificially invoked only for the yer vs. non-yer contrast and plays no other role in the phonology of the language (Gussmann, 1980; Lightner, 1965). In the government model of phonology (Scheer, 2006), yers are formalized as dependent (i.e. incapable of government) and thus contrast with non-yers that are treated as heads (capable of government) by virtue of being always phonetically expressed and thus contentful.

Finally, yers could differ from non-yers also based on their frequency. While the type frequency of words with yer and non-yer vowels does not seem to have any systematic pattern,[2] the paradigm frequency is clearly biased in favor of non-yer vowels. Yers only appear in the forms with phonologically zero suffixes. For example, Slovak noun declensions have six cases for singular and plural for each of the three genders, and out of these 36 word forms, only three have a phonetically zero suffix: the Nominative singular of masculine nouns and Genitive plural of feminine and neuter nouns. Since yer vowels surface only in this limited number of word forms, they are less frequent than the non-yer vowels that appear in all declined forms.

The aim of the present paper is to test the assumption of all phonological accounts that the phonological deficiency of yers does not translate into the production level. In other words, we ask if yer vowels are indeed phonetically identical to non-yer vowels, and more specifically, we test if the phonological deficiency is linked to the phonetic weakness of yer vowels. Measuring the weakness of vowels phonetically is not a straightforward issue and we use weakness in this paper as an umbrella term for several phonetic dimensions that assess the patterns in the production of vowels. In the remainder of this section we describe the dimensions of weakness examined in this paper and Section 2.6 describes the actual measures of weakness in more detail.

Most commonly, syllables receiving word stress and word-initial syllables are considered strong while unstressed and non-initial syllables are considered weak. Compared to the vowels of strong syllables, those in the weak syllables tend to be shorter, and produced with greater undershoot of the targets measurable as smaller displacements and/or velocities (e.g. Lindblom, 1963). For Slovak vowels, Beňuš and Mády (2010) showed that phonetic weakness due to fast speech rate and the absence of lexical stress made Slovak vowels quantitatively shorter and qualitatively more centralized. Based on these observations we hypothesize that phonetically weak Slovak vowels should have shorter duration and should be more centralized.

The centralization tendency for weak vowels is also linked to their coarticulatory properties. Recasens and colleagues (e.g. Recasens, 1985, 1999; Recasens et al., 1997) showed that the degree of articulatory constraint correlates positively with the resistance to coarticulation from surrounding sounds as well as with the aggressiveness in influencing these sounds. Hence, more peripheral vowels, which are more articulatorily constrained, resist coarticulation from adjacent consonants and vowels and exert their influence on them more than less peripheral vowels. Therefore, if yer-vowels are phonetically weak, we expect them to resist coarticulation from adjacent vowels and consonants less than non-yer vowels do.

With respect to the coarticulatory characteristics of vowels, we test three levels, as illustrated in Fig. 2. First, we analyze the coarticulation properties of the target vowel V_T with the preceding vowel V_1 and assume that the more similar the production of V_T to the production of V_1, the lesser the coarticulation resistance of V_T, and hence, the weaker the V_T. Given that the first vowel is prosodically stronger than the second since it receives word stress, we assume that the direction of this V-to-V coarticulation will be progressive (i.e. carrying over from V_1 to V_T). However, it could be the case that the coarticulation between the two vowels is primarily regressive, in which case a smaller distance between V_1 and V_T would signal greater coarticulatory resistance, and thus greater phonetic strength, of the target V_T vowel. Therefore, we will test the effect of yer vs. non-yer origin on the coarticulation between V_1 and V_T and determine the direction of V-to-V coarticulation in this sequence.

Second, we test the degree of coproduction between the tongue body vocalic movements of V_T and the lingual consonantal movements of the surrounding consonants. Our hypothesis is that the weaker the vowel, the more coproduction between the vowel and the adjacent consonant should be present. This is because a weak achievement of the vocalic target allows more leeway (i.e. greater scope) for the tasks of producing consonantal constrictions, which results in greater encroachment of the consonants into the vowel production and effectively to greater overlap of the vocalic and consonantal movements.

Third, we examine the temporal overlap of C_1 and C_2 consonants. We assume that greater coproduction of these consonants correlates with V_T weakness. This is because a weaker vowel would allow the adjacent consonants to overlap more with

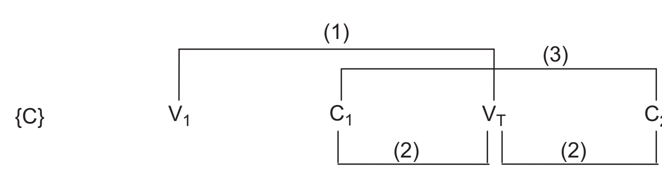

Fig. 2. Schematic illustration of the coarticulatory patterns affecting the production of the target vowel V_T that will be tested in this paper.

[2] See for example the type frequency for the stimuli words listed in (2).

it, and hence, would allow the two consonants flanking the vowel to overlap more with each other.

Finally, we compare phonetic weakness related to speech rate variation to the potential weakness stemming from the yer origin of the vowels. We hypothesize that the qualitative patterns of weakness due to faster speech rate should be comparable to the weakness due to the yer origin of the vowels. In sum, our hypothesis that yer vowels are phonetically weaker than non-yer vowels includes several sub-hypotheses to be tested experimentally: yer vowels are shorter, more centralized, more coproduced with surrounding sounds (both vowels and consonants), and more similar to vowels in fast rate than non-yer vowels.

the first syllable was always identical within the lexical pair. Ideally, both preceding and following consonants were identical, as in *ábel* vs. *kábel*, or *nárok* vs. *párok* but at the minimum they agreed in the place of articulation, as in *námet* vs. *rámec[ts]* or *kapor* vs. *mramor*. We tried to include as many labial consonants as possible because they have a minimal effect on the tongue body movements. Although we could not create a stimulus list in which each pair satisfied all the above considerations, our list of ten pairs listed in (2), five with vowel /e/ and five with /o/, is sufficiently controlled phonetically, natural, and representative. The numbers in brackets correspond to lemma frequencies extracted from the electronic corpus of the Slovak language (http://korpus.juls.savba.sk/).

(2) Stimuli list

YER			NON-YER		
kábel (5564)	[kaːbel]	'cable'	Ábel (381)	[aːbel]	'Name'
Čapek (940)	[tʃapek]	'Name'	papek (181)	[papek]	'twig'
cumel (5)	[tsumel]	'pacifier'	čumel (143)	[tʃumel]	'he stared'
obec (128,656)	[obets]	'village'	obed (20240)	[obet]	'lunch'
rámec (116,986)	[raːmets]	'frame'	námet (8892)	[naːmet]	'idea'
párok (1523)	[paːrok]	'sausage'	nárok (34915)	[naːrok]	'requirement'
nebol (> 100,000)	[ɲebol]	'he wasn't'	jebol (14)	[jebol]	'he fell (curse)'
kufor (7796)	[kufor]	'suitcase'	humor (13630)	[humor]	'humor'
kapor (2510)	[kapor]	'carp'	mramor (2323)	[mramor]	'marble'
smútok (15,699)	[smuːtok]	'sadness'	sútok (600)	[suːtok]	'confluence'

The rest of the paper is structured as follows. Section 2 presents the methodological issues related to data collection, quantification, and analysis. Section 3 presents the results comparing vowels in fast and normal rate in terms of duration, vowel quality, and patterns of coarticulation with neighboring sounds and follows with similar comparisons for yer and non-yer vowels. Section 4 discusses the main findings and their relation to current understanding of the interface between phonetics and phonology.

2. Methodology

2.1. Subjects

Five native speakers of Slovak between the age 20 and 40 years (two females and three males) participated in this study. Their speech did not deviate from customary patterns of standard colloquial Slovak. The subjects were naïve as to the purposes of the study, and none reported any speech, hearing, or language problems.

2.2. Material

Material consisted of pairs of real words so that a yer vowel appeared in one member of the pair and a non-yer vowel occurred in the corresponding pair. The pairs were designed in such a way that the two members of each pair occurred in an environment as similar as possible. In this way, the effect of the phonological origin of the vowels on their production was minimally obscured by spurious phonetic and prosodic differences in the environment or positions of the target vowels. The target vowel always appeared in the second unstressed syllable of the word, flanked by a single consonant on each side.[3] The vowel of

[3] Slovak, like other Slavic languages, has several words with yers in the first syllable such as *pes* 'dog' or *deň* 'day' but we could not find suitable pairs of words with non-yer vowels with which we could compare them.

Both yer and non-yer target vowels were non-initial, unstressed, non-peripheral, and phonologically short. Hence, compared to other vowels in Slovak, they were all already significantly weaker in terms of these structural characteristics and any further differences in their weakness will be attributed to the yer vs. non-yer origin of these vowels.

2.3. Procedure

Subjects read the target words embedded in a prompt sentence at normal and fast speech rates. For speech rate variation we used an ecological approach adapted from Adams, Weismer, and Kent (1993) and Hoole and Mooshammer (2002) and described also in Beňuš (2011). First, we elicited a subset of prompt sentences from each subject in five self-selected speech rates during a pre-test session. Then, taking into consideration the salience of perceptual contrast between the rates and the consistency of prosodic patterns, we selected two sentences with the target word [kaːbel] for each subject so that they represented normal and fast speech rates of that subject respectively. Finally, during the actual data collection in alternating blocks of normal and fast speech rates, the appropriate sentence was presented as a speech rate cue randomly before each 3–8 prompt sentences and the subjects were instructed to match the rate of their test sentences as closely as possible to the rate of the cue sentences.

Prompt sentences consisted of a coordinated structure in which the first clause included a conjugated form of the lexical item, and the second clause contained the target form listed in (2) that was analyzed. Sample prompt sentences for one yer and one non-yer word are listed in (3). All prompt sentences were presented visually in a randomized order in ten blocks (five for normal rate and five for fast rate) on a computer screen in standard Slovak spelling. Since the stimuli list included 20 lexical items, we collected 200 tokens for each subject (5 repetitions, 2 rates) for a total of 1000 tokens.

(3) *Čítame s humorom a humor parádne.*
We read with humor-Instr. and humor-Nom. beautifully.
Čítame s kufrom a kufor parádne.
We read with suitcase-Instr. and suitcase-Nom. beautifully.

2.4. Data collection and processing

Articulatory data were collected with electro-magnetic articulography (EMA, AG500, Carstens Medizinelektronik, IPS Munich), which tracks the movements of receivers attached to active articulators at a sampling rate of 200 Hz. After applying standard calibration and cleaning procedures (Hoole & Zierdt, 2010), seven such receivers were placed in a mid-sagittal plane: on the upper lip, lower lip, the lower incisors to record jaw movement, and four sensors on the tongue that were glued in roughly equidistant intervals between the tongue tip area and the velar/dorsal region of the tongue. We will refer to these four sensors TT, TB1, TB2, and TD respectively. In addition to tracking the movements of the active articulators, four reference sensors were attached: two sensors behind each ear, one on the nose, one above the upper incisors. The information from the movement of these sensors was used in post-processing to correct for head movement during data collection. Movement data were filtered with 60 Hz for the tongue tip sensor, 20 Hz for all other sensors attached to the active articulators, and 5 Hz for the four reference sensors. The data were corrected for head movement, and rotated to each subject's occlusal plane (Hoole & Zierdt, 2010). Acoustic signal was captured with a directional Sennheiser MKH 40 microphone with a sampling rate of 32,768 Hz and downsampled during post-processing to 16,384 Hz.

2.5. Data labeling and extraction

Using the Praat labeling environment (Boersma & Weenink, 2010) and following standard procedures, a trained annotator determined the temporal intervals of the three segments (CVC) in the second syllable of each target word. The beginning of the onset consonant was marked at the cessation of the formant structure of the preceding vowel, the beginning of the target vowel at the zero crossing of the first cycle of the modal voice with a formant structure, and the end of the vowel at the cessation of the formants. Because the boundary between the vowel and coda /l/ consonants could not be, in many cases, reliably determined, the interval for the analyzed target vowel in these cases included a complete syllable rime.

Using Matlab-based procedures for the visualization and labeling of the articulatory movements developed by M. Tiede, a single annotator (different from the annotator of the acoustic signal) identified the kinematic landmarks of the consonantal gestures in the final CVC syllable of the target word in the following way. The annotator manually selected the temporal window comprising of the movement to be labeled, and the algorithm first identified the peak velocities of the movement into and out of a constriction (PVEL1 and PVEL2 respectively), and then identified the landmarks for gesture onset (GONS), achievement of target (NONS), release of the target (NOFFS) and gesture offset (GOFFS) on the basis of a percentage threshold (default 20%) of the peak velocity ranges. Finally, the point of minimal velocity between the NONS and NOFFS landmarks was identified as the maximal constriction (MAXC). The gestures of bilabial consonants were identified on the velocity profiles of the Lip Aperture measure (LA) which represents the Euclidean distance between the upper and lower lip sensors. The labio-dental /f/ was labeled on the velocity profile of the sensor attached to the lower lip. The gestures for the alveo-dental consonants /t/, /l/, /r/, and /ts/ were labeled on the vertical velocity profile of the tongue tip sensor, and the gesture for /k/ on the vertical velocity profile of the tongue dorsum sensor. Fig. 3 illustrates one result of such labeling. In several cases, the automatic algorithm gave clearly wrong results during the labeling, which were rectified by adjusting the default thresholds.

Such landmarks, however, could not be reliably identified for the vocalic movements. This is partly because our target vowels were mid and unstressed, which decreased the spatio-temporal expansion of these gestures. Furthermore, it was impossible to find stimuli in which both surrounding consonants were labial. Hence, at least one of the consonants immediately adjacent to the target vowel, and sometimes both consonants, required the tongue to produce a lingual constriction. Due to natural overlap of the vocalic and consonantal gestures (e.g. Öhman, 1966) formalized articulatorily as blending of two gestures that control a single effector articulator (Saltzman & Kelso, 1987), the vocalic gestures were greatly obscured by the adjacent consonantal gestures. For these reasons, and despite the effort in designing the stimuli, the unique movement for the target vowel could not be determined.

2.6. Dependent variables

We assessed the phonetic weakness of the target vowels with dependent variables that fall into three categories following the discussion in Section 1.2: duration, quality, and coarticulatory characteristics. Given very small differences discussed above, we will take a less conservative approach and consider a hypothesis supported if at least one of the dependent variables yields statistically significant difference in the hypothesized direction of the effect and the other measure(s) do not yield a statistically significant effect in the opposite direction.

In duration, we hypothesized that yer vowels are shorter than non-yer vowels. One acoustic and one articulatory measure of duration in C1VC2 sequences were used: interval between C1 acoustic release and C2 closure (*DurAc*),[4] and between the articulatory release of the closure for C1 and the achievement of target for C2 (C2-Nons–C1-Noff, *DurArt*, shown in Fig. 4). *DurAc* measures the duration of the modal voice activity without aspiration and is thus more perceptually biased than *DurArt*.

In vowel quality, we expected yer vowels to be less peripheral than non-yer vowels. Hence, horizontally, yer /e/ should be more retracted and yer /o/ more fronted than their non-yer counterparts. Given known non-linearities between the acoustic and articulatory measures of vowel frontness (e.g. Stevens, 1989), both acoustic and articulatory measures were used. Acoustically we tested vowel quality with the values of the second formant, extracted at the temporal midpoint of the vowel taken as a midpoint between C1-Noff and C2-Nons labels (*F2*). Articulatorily we examined the horizontal positions of the two tongue body receivers at the same temporal point (*TB1-x*, *TB2-x*).

The last group of measures examined the coarticulatory properties of target vowels. The overall hypothesis was that yer vowels would show more coarticulation with the surrounding sounds than non-yer vowels. First, we investigated the degree of coarticulation of the target vowel (V_T in Fig. 4) with the preceding vowel in the first stressed syllable (V_1 in Fig. 2) by calculating the Euclidean distance in the F1–F2 space between the values

[4] In case of C2 liquid, C1-release to C2 release was used.

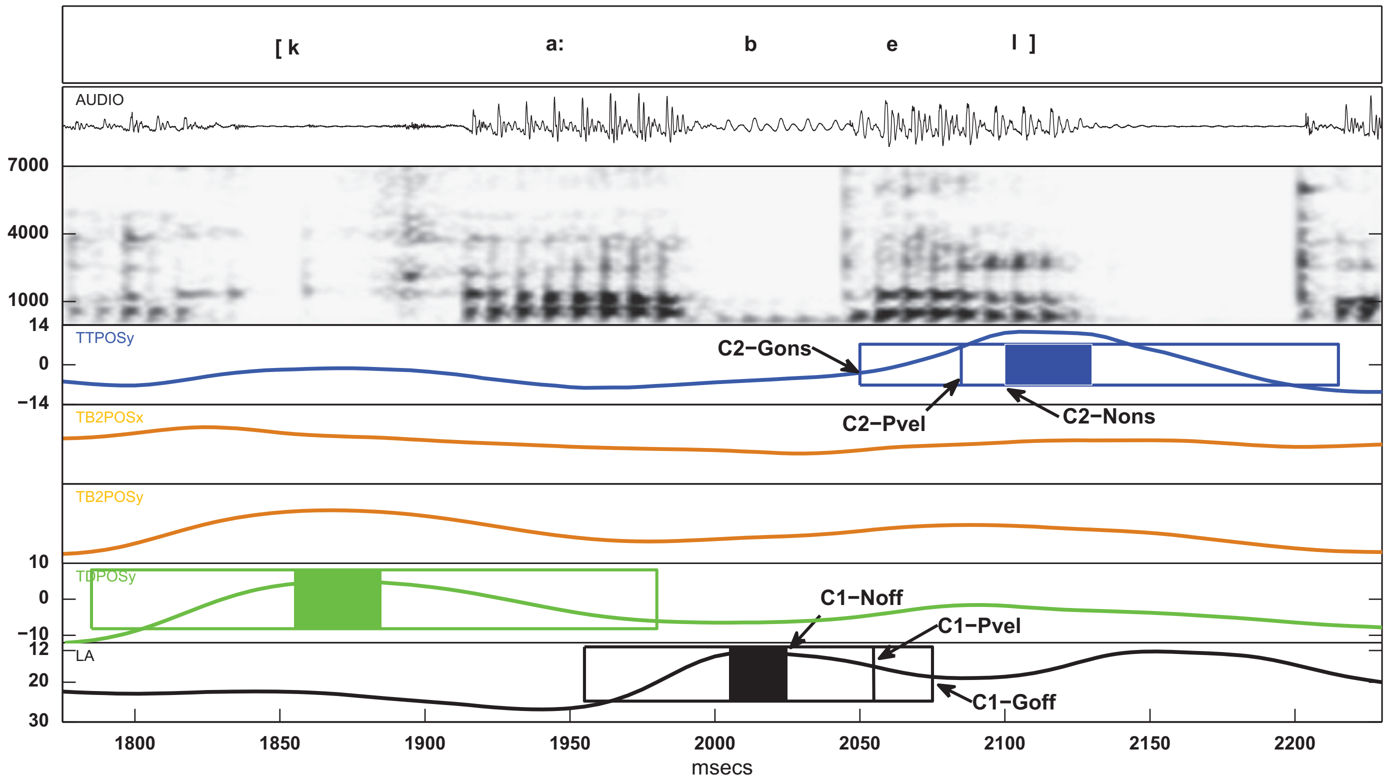

Fig. 3. Example of kinematic landmarks in the word [kaːbel]. Time in ms is on the *x*-axis. Panels along the *y*-axis from top to bottom: audio signal, spectrogram, vertical movement of the tongue tip (TT) sensor, horizontal movement of the tongue body (TB2) sensor, vertical movement of the tongue body (TB2) sensor, vertical movement of the tongue dorsum (TD) sensor, Lip aperture (LA=euc. distance between the upper and lower lip sensors). The filled rectangles correspond to the interval between the achievement and the release of the target, and empty rectangles to the onset and offset of the gesture.

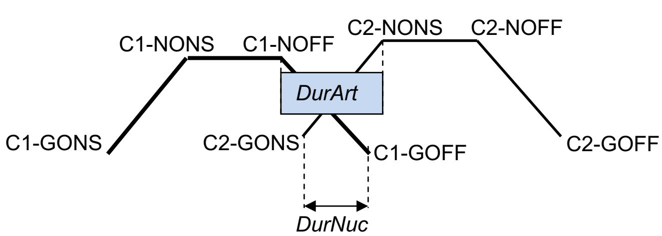

Fig. 4. Temporal intervals for the examination of the vocalic movements based on the gestural landmarks of the flanking consonants; see text for details.

extracted at the temporal midpoint of V_T and the values extracted 10 ms before the acoustic offset of V_1 ($V_1-V_T EucDist$), i.e. 10 ms before the formation of the constriction of the following consonant labeled in the acoustic signal. The smaller the value of this Euclidean distance, the more coarticulated were the two vowels. As discussed in Section 1.2, the Euclidean distance measure, however, gauges the weakness of the target vowels only under the assumption that V-to-V coarticulation is primarily progressive. Therefore, to test, if this was the case in our data, we examined the effect of phonological category (yer vs. non-yer) and speech rate not only on the Euclidean distance measure but also on the formant values extracted from both temporal points separately.

Secondly, we examined the overlap of the target vowel with the surrounding consonants using the time functions of the horizontal and vertical movements of the two tongue body receivers within the intervals defined by the release of C1 and the achievement of target for C2 (C1-Noff−C2-Nons), which corresponds to *DurArt* measure mentioned above and is illustrated in Fig. 4. If the tongue moves differently for yer vs. nonyer vowels, these time functions should be different. One technique for assessing the global properties of trajectories is the discrete cosine transformation (DCT, Harrington, 2010). This mathematical operation decomposes the signal into a set of cosine waves at frequencies $k=0$, 0.5, 1, etc. and the amplitudes of these waves are called DCT coefficients.[5] The first three coefficients correspond to the signal amplitude, slope, and curvature respectively. Due to the presence of adjacent lingual consonants, the slope and the curvature of the time functions for the vertical and horizontal tongue movements, corresponding to the second and the third DCT coefficients, assess the coarticulation of the vowels with these consonants. Following the discussion in Section 1.2 we assume that the degree of overlap between adjacent vowel and consonant indicates the weakness of the vowel and that greater overlap between vowels and lingual consonants results in flatter movement of the tongue both in terms of its slope and curvature. If yer vowels are weaker than non-yer ones, they should be produced with flatter slope and curvature of the tongue body movement. Because we were interested in the size of the effect and not its direction (i.e. whether slopes and curvatures were positive or negative) absolute values of the second and third DCT coefficients were used in the analysis.

In addition to assessing spatial characteristics of the tongue movements, another way of gauging the consonant-vowel overlap is to extract the kinematic and dynamic characteristics of the

[5] To normalize for trajectory duration, we used a Matlab procedure designed by A. Gafos for the repetitions of each target pair of yer and non-yer words. Hence, for each 10 trajectories (2 tokens, 5 repetitions), the script determined the shortest trajectory and equalized all other trajectories to have the same number of points as this shortest one (aligning from the left). This was done separately for the two receivers, two rates and the horizontal and vertical dimension. A built-in Matlab procedure for calculating DCT coefficients was then used on these data.

consonantal movement preceding the target vowels (constriction opening) and following it (constriction closing). These measures thus allow including also non-lingual consonantal movements in assessing phonetic weakness. We expected that if a consonantal movement is slower and less stiff, it encroaches into the vowel more. Hence, if yer vowels were weaker, they were expected to be surrounded by slower and less stiff consonantal movements. In the notional dynamic model most commonly used to describe articulatory movements – a damped mass-spring (Saltzman & Kelso, 1987) – stiffness represents a major determinant of movement duration and varies positively with velocity and negatively with displacement. We extracted peak velocity (C1-Pvel, C2-Pvel), time-to-peak-velocity (C1-Goff − C1-Pvel, C2-Pvel − C2-Gons, and assessed movement stiffness as peak velocity over displacement (the Euclidean distance between the position of the receiver at C1-MAXC and C1-Goff or C2-Gons and C2-MAXC).

Finally, we assessed the overlap of the target vowels with surrounding consonants by testing the degree of coproduction of the two surrounding consonants. We hypothesized that greater coproduction of consonants around the target vowel correlates with greater coproduction of the target vowel with these consonants, and thus, that yer vowels should show a greater coproduction of the consonants around them than non-yer vowels. We used two measures. The first was the interval between the gesture offset for C1 and the gesture onset for C2 *DurNuc* (C2-Gons − C1G-off, Fig. 4). The lower the value, the greater the overlap of the two consonantal movements. Note that this calculation may also yield negative values. Whereas *DurNuc* is a rather naïve measure based on the onset and offset of consonantal gestures, the second measure, *Peak-to-Peak ratio*, is a more global and dynamic measure of truncation between the movement away from the constriction of C1 and the movement toward the constriction of C2. Harrington et al., (1995) and Hoole and Mooshammer (2002) both found that the ratio of the interval between the velocity peaks of the C1-opening and C2-closing gestures over the interval between the release of C1 to the achievement of target for C2 provides a good measure of the temporal coproduction of the two movements. In our case, this measure was calculated following the landmarks illustrated in Fig. 3 as (C2-Pvel − C1-Pvel)/(C2-Nons − C1-Noff).

2.7. Statistical analysis

We employed a mixed-models approach implemented with the R software package lmer for determining the effects of fixed factors Category (yer vs. non-yer) and Tempo (normal vs. fast) and their interaction on dependent variables (Baayen, 2008). Subject, LexicalPair, and Repetition were random factors in the model. The primary reason for using this test is that it allows filtering out the variation between the subjects as well as the variation between the 10 pairs of target words within a single test (Harrington, 2010). The mixed-model approach thus also obviates the need for the normalization of the data that depend on the physiology of articulators and the placement of the sensors and a need for repeated measures design both for subjects and tokens. A disadvantage of this test is a somewhat problematic assessment of the degrees of freedom needed for determining *p*-values. We subjected the results of the model fitting into *R*'s anova function that tests whether the model terms are significant and returns *F*-values for each fixed factor and interactions between these factors. We followed the conservative approach of Reubold, Harrington, and Kleber (2010) and set the degrees of freedom to 60, and the alpha level to 0.01. Under these conditions, all *F*-values greater than 8.49 will be considered as significant at $p<0.01$.[6] Finally, since factors Vowel (/e/ vs. /o/) and LexicalPair were not independent (five pairs contained /e/ and five contained /o/), we could not use Vowel as an independent factor and LexicalPair as a random factor in one test. Therefore, we tested the interaction of Vowel with Category and Tempo and ran separate tests for /e/ and /o/ if this interaction was significant.

3. Results

We start by describing phonetic differences between normal and fast rate and then differences between yer and non-yer vowels. The first set of results is more robust and clearer than the second and this order of presentation also facilitates the comparison of the type of phonetic weakness predicted for fast speech rate with the type of weakness predicted in yers.

3.1. Weakness due to faster rate

Measured in the acoustic signal (*DurAc*), vowels in fast rate were robustly shorter than in normal rate. However, the difference between the means was rather small at only 13.2 ms. A similar result was obtained in the articulatory data (*DurArt*): fast rate vowels were shorter than normal rate ones with the difference between the means of 18.7 ms. Mixed-models tests with Subject, LexicalPair, and Repetition as random factors confirm the significance of the rate effect on both measures; $F=316.2$, $p<0.01$ and $F=443.8$, $p<0.01$ respectively. There was no significant main effect of Vowel and no significant interaction between Tempo and Vowel.

Assessing vowel quality acoustically, F2 was not significantly affected by Tempo, but the interaction between Tempo and Vowel was significant; $F=19.2$, $p<0.01$. Separate tests showed that fast-rate /e/ had lower F2 than normal rate /e/; $F=8.0$, $p<0.05$, the difference between the means was 33 Hz, and fast-rate /o/ had higher F2 than normal-rate one; $F=13.2$, $p<0.01$, the difference of means 34.3 Hz. Hence, fast-rate vowels were more centralized horizontally (/e/ was more retracted and /o/ more fronted) than normal-rate vowels.

Assessing the hypothesized horizontal centralization articulatorily, no significant effect of Tempo was found on the temporal midpoint in the horizontal dimension (*TB1-x*, *TB2-x*). Hence, the hypothesized horizontal centralization of target vowels in fast rate was observed only in the acoustic measures.

The final set of measures evaluated the coarticulation of target vowels with the preceding vowel and with the flanking consonants. First, assessing the distance between the target vowel and the vowel that precedes it ($V_1-V_T EucDist$), we found no significant effect of speech rate. This result is unexpected and might be related to already centralized productions of unstressed mid vowels even in normal rate and minimal shortening of duration due to increased rate (less than 20 ms on average) reported above.

Second, we tested the coproduction of the target vowel with the adjacent lingual consonant(s) with the slope (*DCT2*) and the curvature (*DCT3*) of the time functions extracted from the horizontal and vertical movements of two sensors attached to the tongue body (TB1 and TB2). Hence, there were eight dependent variables (data from 2 sensors, 2 dimensions, and 2 DCT coefficients) and we thus ran eight mixed-models tests. The effect of speech rate was robust and had the predicted direction: for both sensors and dimensions, the slopes and curvatures in the fast rate were significantly flatter than in the normal rate; *F* values ranged between 30.8 and 103.7. Hence, the target vowel was coproduced with the adjacent lingual consonant to a greater extent in fast rate than in normal rate.

Analyzing the coproduction of the consonants flanking the target vowels we observed that, as expected, the consonants

[6] *F*-values greater than 7.2 will be considered significant at $p<0.05$.

Table 1
Summary of the results testing phonetic weakness due to fast rate; ☑ corresponds to a significant effect supporting the hypothesis.

<table>
<tr><th>V_F=fast, V_N=normal</th><th>Hypothesis</th><th>Measure</th><th></th><th>/e/</th><th>/o/</th></tr>
<tr><td>V-duration</td><td>V_F shorter than V_N</td><td>DurAc</td><td></td><td colspan="2">☑</td></tr>
<tr><td></td><td></td><td>DurArt</td><td></td><td colspan="2">☑</td></tr>
<tr><td>V-quality</td><td>V_F more centralized than V_N</td><td>Ac (F2)</td><td></td><td>☑</td><td>☑</td></tr>
<tr><td></td><td></td><td>Art (TB{1,2}-x)</td><td></td><td>–</td><td>–</td></tr>
<tr><td>V_1-V_T coarticulation</td><td>$|V_1-V_F| < |V_1-V_N|$</td><td>V_1-V_TEucDist</td><td></td><td>–</td><td>–</td></tr>
<tr><td>(C)V_TC coproduction</td><td>slope V_F < slope V_N</td><td>DCT2 {TB1,TB2}</td><td>hor.</td><td colspan="2">☑</td></tr>
<tr><td></td><td></td><td></td><td>vert.</td><td colspan="2">☑</td></tr>
<tr><td></td><td>curvature V_F < curvature V_N</td><td>DCT3 {TB1,TB2}</td><td>hor.</td><td colspan="2">☑</td></tr>
<tr><td></td><td></td><td></td><td>vert.</td><td colspan="2">☑</td></tr>
<tr><td>C_1C_2 coproduction</td><td>$|C_1V_FC_2| < |C_1V_NC_2|$</td><td>Peak-to-Peak ratio</td><td></td><td colspan="2">☑</td></tr>
<tr><td></td><td></td><td>DurNuc</td><td></td><td colspan="2">☑</td></tr>
</table>

flanking target vowels in fast rate showed significantly greater coproduction than the consonants flanking the vowels in normal rate as measured with *Peak-to-Peak ratio* (F=42.8, $p<0.01$), and also in *DurNuc* (F=13.8, $p<0.01$).

Phonetic weakness of unstressed Slovak mid vowels due fast rate compared to normal rate is summarized in Table 1. The data support our hypotheses that fast rate vowels are shorter, more centralized based on F2, have flatter slopes and curvatures of both horizontal and vertical movement of the tongue body and also display greater coproduction of the flanking consonants.

3.2. Weakness due to yer origin

3.2.1. Vowel duration

In duration as the measure of phonetic weakness we hypothesized that if yer vowels are weaker than non-yer ones, the former should also be shorter than the latter. A mixed model test with Category (yer/non-yer) as the dependent variable and Subject, LexicalPair, and Repetition as random factors showed that acoustically (*DurAc*), yer vowels were very slightly, but significantly, shorter than non-yer vowels; F=10.1, $p<0.01$. The mean difference, however, was miniscule at 2.5 ms. A further examination revealed that this difference was largely attributable to the rhymes of the target syllables rather than the vowels. Thus, when excluding the three pairs where the boundary between the vowel and the syllable coda could not be reliably determined (/el/, /ol/), the effect of Category in the remaining 7 pairs was no longer significant while the effect of Category in the 3 liquid-coda pairs was more robust (F=16.1, $p<0.01$, mean difference 5 ms). Hence, syllable rhymes in yer words were shorter than the same rhymes in non-yer words.

The articulatory measure of vowel duration *DurArt* was not significantly affected by Category in the pooled data (F=5,0, ns.), but the test reported a significant interaction with Vowel (F=20.7, $p<0.01$). Separate tests for the two vowels showed that yer /o/ vowels were significantly longer than non-yer /o/ vowels; F=25.4, $p<0.01$, difference of means 7.8 ms.

In sum, the effect of Category on vowel duration was minimal and inconsistent: in some measures yer vowels were longer than non-yer ones (/o/ on *DurArt*) and in others they were shorter (/el/ and /ol/ rhymes with *DurAc*). Hence, yer vowels could not be considered shorter than non-yer vowels.

3.2.2. Vowel quality

Vowels articulated with more centralized articulation positions are also considered phonetically weaker than vowels with more peripheral articulations; hence, we expected yer vowels to be more centralized than non-yer ones. Acoustically, there was no main effect of Category on F2 but it interacted significantly with Vowel. Yer /e/ had a significantly lower F2 than non-yer /e/ (F=9.5, $p<0.01$, diff. of means 38 Hz), and no significant effect of Category was observed with /o/.

To assess possible vowel quality differences with articulatory measures we examined the horizontal positions of the two receivers at the temporal midpoint placed on the tongue body (*TB1-x*, *TB2-x*) because this articulator is the main determinant of vowel quality. The mixed models test showed a significant effect of Category on vowel /o/ such that yer vowels were more centralized than non-yer vowels. The TB2 sensor for yer /o/ vowels was horizontally more fronted than for non-yer vowels (F=18.3, $p<0.01$, difference of means 0.8 mm). Category did not have a significant effect on vowel /e/ in this measure.

In sum, yer vowels can be characterized as more centralized horizontally since front /e/ yer vowels were acoustically more retracted than non-yer vowels and back /o/ yer vowels were articulatorily more fronted than non-yer vowels.

3.2.3. Coarticulation with the preceding vowel

Another measure of phonetic weakness is the degree of resistance to coarticulation from surrounding sounds. We start with exploring the coarticulation patterns between the target vowel and the vowel that precedes it. In other words, we assess the effect of the vowel in the initial stressed syllable on the vowel in the second unstressed syllable using measure V_1-V_T*EucDist* described in Section 2.6. If yers are weaker than non-yers, the initial vowel (V_1) should have a greater effect on yer vowels than non-yer ones. Category did not affect this measure significantly in the pooled data (F=6.7, $p>0.05$), but the interaction with Vowel was significant (F=10.5, $p<0.01$). In separate tests, yer /e/ vowels had significantly smaller distance to, and thus were more coarticulated with, the preceding vowels than non-yer /e/ vowels (F=21.7, $p<0.01$).

To determine the direction of the observed V-to-V coarticulation, i.e. if the first vowel affects the second or vice versa, we checked if Category affected the V1-offset for the target /e/ vowel and found no significant effect on neither the first (F=6.6, ns.) nor the second formant (F=0.6, ns.). Taken together with the results from F2, in which yer /e/ was more centralized than non-yer /e/, and the generally greater resistance to coarticulation of the stressed vowels than unstressed ones (reported for Slovak in a different dataset, Beňuš & Mády, 2010), we analyze the observed effect of Category on the V_1-V_T*EucDist* measure for vowel /e/ as evidence for smaller resistance to coarticulation of yer /e/ vowels compared to non-yer ones. The absence of the effect of Category on /o/ might be linked to its lower degree of coarticulatory resistance in general when compared to /e/.

3.2.4. Coproduction with surrounding consonants

The second and third coefficients of the discrete cosine transform reflect respectively the slope and the curvature of the time functions extracted from the horizontal and vertical movements of the sensors attached to the tongue body. As discussed in Sections 1.2 and 2.6, given that each token has at least one lingual consonant surrounding the target vowel, these two coefficients can be used for assessing the coproduction of the vowel with the lingual consonant. As there are three pairs of variables (DCT coefficients for slope and curvature, for horizontal and vertical movement, and for TB1 and TB2 sensors), we ran eight separate mixed-models tests. The dependent variable in each test combined one member from each of the three pairs; for example, DCT2 in the horizontal movement of TB1, DCT3 in the vertical movement of TB2, etc. We found a significant main effect of Category in three tests: yer vowels had flatter slopes in the vertical movement of TB1 ($F=12.1$, $p<0.01$), in the horizontal movement of TB2 ($F=7.8$, $p<0.05$), and flatter curvatures in the horizontal movement of TB1 ($F=7.7$, $p<0.05$). When testing the steepness of slopes in the horizontal dimension of TB1, the main effect was absent but Category interacted significantly with Vowel ($F=12.5$, $p<0.01$). In the separate tests for each vowel we observed significantly steeper slopes of yer /e/ vowels than non-yer ones ($F=11.7$, $p<0.01$).

The visualization of the movement of the tongue body sensors complements the analyses of the horizontal and vertical time functions presented above. For example, Fig. 5 shows this movement for the TB1 sensor between the release of /b/ and the achievement of the target for /l/ in the pair [kaːbel–aːbel] for all five subjects separately. Despite great variability in the productions, we can observe that non-yer vowels are produced in general with greater curvatures (apart from S2) that signal greater frontward (S1, S5) or upward (S3) movement than yer vowels. For subject S2, most of the movement toward the target vowel occurred during the b-closures, which explains a somewhat different pattern of movement from the ones observed for the other subjects.

We conclude the section by reporting on the kinematic characteristics of the consonantal movements surrounding the target vowels and the patterns in their temporal coproduction. We start with an analysis of *Peak-to-Peak ratio*, which is a measure of phonetic weakness related to the truncation of adjacent articulatory movements (Section 2.6). According to this measure, the consonantal opening movement preceding the vowel and the closing movement following it were more coproduced for yer /e/ than non-yer /e/ ($F=15.9$, $p<0.01$), no significant effect was observed for /o/ ($F=0.2$, ns.). Similarly, the consonants flanking yer vowels overlapped more than consonants flanking non-yer vowels as measured with *DurNuc*, which is the interval between the offset of the C1 gesture and onset of the C2 gesture ($F=15.6$, $p<0.01$).

The significantly greater overlap of consonantal movements around yer /o/ measured with *DurNuc* might seem puzzling in the view of the result reported for *DurArt* where yer /o/ vowels showed less overlap of the flanking consonants, and thus longer duration, than non-yer /o/ vowels (Section 3.1). One would expect that longer vowel duration should co-occur with less coproduction of consonants surrounding it. Hence the two measures of consonantal overlap provide contrastive findings for yer vs. non-yer /o/ vowels. But there are two primary mechanisms for duration changes—adjustment of movement stiffness and temporal alignment (coordination) of adjacent movements (e.g. Beňuš, 2011). We examined additional kinematic characteristics of the opening movement before the vowel and the closing movement after the vowel separately to investigate the relationship between stiffness and coordination in durational changes for

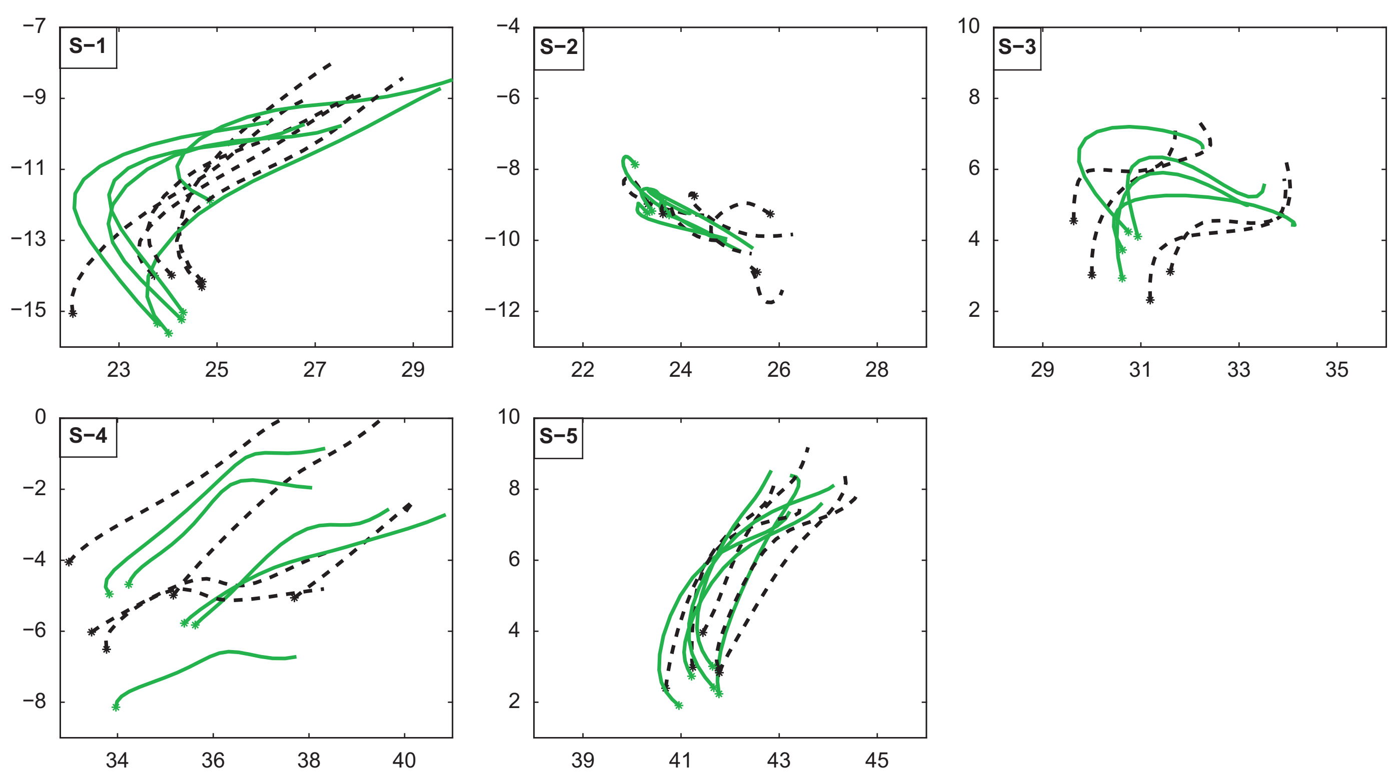

Fig. 5. Movement for the TB1 sensor between the articulatory landmarks of the release of /b/ and the achievement of the target for /l/ in the pair [kaːbel–aːbel] for all five subjects separately in normal speech rate. The movement of the yer vowels [kaːbel] is shown in dark dashed black lines and the movement for non-yer vowels [aːbel] in light solid gray (green) lines. The axes vary since the coordinate system differed for each subject but all five plots show 8 mm on the *x*- and 9 mm on the *y*-axis, each tick equals 2 mm, *x*-axes refer to the horizontal and *y*-axes to the vertical movements of the sensors (the top left corner of each box corresponds to high front position of the sensor). Stars show the temporal onset of the movement. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Summary of the results testing phonetic weakness of vowels due to their yer origin; ☑ corresponds to a significant effect supporting the hypothesis, ☒ marks a significant effect in the opposite direction.

<table>
<tr><th>V_Y=yer, V_{NY}=non-yer</th><th>Hypothesis</th><th>Measure</th><th></th><th>/e/</th><th>/o/</th></tr>
<tr><td>V-duration</td><td>V_Y shorter than V_{NY}</td><td>DurAc</td><td></td><td colspan="2">☑??</td></tr>
<tr><td></td><td></td><td>DurArt</td><td></td><td>–</td><td>☒</td></tr>
<tr><td>V-quality</td><td>V_Y more centralized than V_{NY}</td><td>Ac (F2)</td><td></td><td>☑</td><td>–</td></tr>
<tr><td></td><td></td><td>Art (TB{1,2}-x)</td><td></td><td>–</td><td>☑</td></tr>
<tr><td>V_1-V_T coarticulation</td><td>$|V_1-V_Y|<|V_1-V_{NY}|$</td><td>V_1-V_TEucDist</td><td></td><td>☑</td><td>–</td></tr>
<tr><td>(C)$V_T C$ coproduction</td><td>$slope_Y<slope_{NY}$</td><td>DCT2{TB1,TB2}</td><td>hor.</td><td>☑☒</td><td>☑</td></tr>
<tr><td></td><td></td><td></td><td>vert.</td><td colspan="2">☑</td></tr>
<tr><td></td><td>$curv_Y<curv_{NY}$</td><td>DCT3{TB1,TB2}</td><td>hor.</td><td colspan="2">☑</td></tr>
<tr><td></td><td></td><td></td><td>vert.</td><td>–</td><td>–</td></tr>
<tr><td>C_1C_2 coproduction</td><td>$|C_1V_YC_2|<|C_1V_{NY}C_2|$</td><td>Peak-to-Peak ratio</td><td></td><td>☑</td><td>–</td></tr>
<tr><td></td><td></td><td>DurNuc</td><td></td><td colspan="2">☑</td></tr>
</table>

/o/ vowels. The constriction-opening movement preceding /o/ was longer for yer /o/ vowels than non-yer vowels (F=24.9, $p<0.01$, difference of means 8.7 ms) and the same was observed for the constriction-closing movement following it (F=13.2, $p<0.01$, difference of means 5.1 ms). Additionally, both movements surrounding yer /o/ vowels had longer time to peak velocity (F=11.6, $p<0.01$ for the preceding movement and F=25.1, $p<0.01$ for the following, difference of means 4.6 and 5.6 respectively), and lower stiffness (F=36.1, $p<0.01$ and F=12.8, $p<0.01$). Similar findings were observed also for yer vs. non-yer /e/ vowels.

Hence, the consonantal movements that precede and follow yer /o/ vowels were in general slower, longer and less stiff, which co-occurred with longer *DurArt* intervals. One way of construing this finding is that movement stiffness is decreased around yer /o/, but the temporal coordination is not adjusted. This observation from the stiffness corroborates the results from CV coproduction with slopes and curvatures and shows that yer /o/ vowels were co-produced with the surrounding consonants over a longer period of time than non-yer /o/ vowels. Yer /e/ vowels seem to show both decreased stiffness as well as tighter coordination of the two consonants as shown by the significantly lower *Peak-to-Peak ratio* and *DurNuc* of yer /e/ vowels than non-yer ones together with no significant effect on *DurArt*.

Table 2 summarizes main findings related to the weakness of yer vowels. Following the evaluation proposed in Section 2.6, the sub-hypotheses that yer vowels are shorter and that they have flatter slopes in the horizontal movement of tongue body sensors were not supported. On the other hand, data for at least one yer vowel support the sub-hypotheses that yer vowels are more centralized, coarticulate more with the preceding vowel, and that the consonants surrounding yers are more coproduced than consonants around non-yers.

3.2.5. Comparison of yer-origin and rate

This section compares the findings from the previous two subsections in order to evaluate the prediction that the weakness due to faster rate might manifest similar phonetic effects as weakness due to yer origin. There was no significant interaction between Category and Tempo on any of the dependent variables, which suggests that the two types of weakness are in fact phonetically similar. Additionally, several patterns in the individual results support the predicted similarity between the two types of weakness. These correspond to the pairs of cells in Tables 1 and 2 that share a tick. First, yer /e/ was more centralized than non-yer /e/, which was similar to fast-rate /e/ that was also more centralized than normal-rate /e/. Yer /o/ was also more centralized than non-yer /o/, which was similar to fast-rate /o/ that was also more centralized than normal-rate /o/.[7]

Second, although speech rate had no significant effect on the coarticulation of the target vowel with the preceding vowel (V_1-V_T*EucDist*), as seen in Fig. 6, the direction of the significant Category effect (smaller values for yer /e/ than non-yer /e/) corresponded to the direction of the Tempo effect (smaller values in fast rate /e/ vowels than normal rate /e/ vowels).

Third, yer /e/ vowels patterned together with vowels in fast rate and showed more coproduction of the surrounding consonantal movements as measured by *Peak-to-Peak ratio*.[8] These findings are summarized in Fig. 7. It should be noted that Figs. 6 and 7 are based on pooling over speakers and lexical pairs and thus show a lot of overlap between categories despite the significance reported by the statistical tests that filter the effects of speaker and lexical pair.

Hence, on these three measures, yer /e/ vowels (and yer /o/ less robustly) pattern together with vowels in fast rate and they thus support the idea that yer vowels behave similarly to vowels in fast speech rate and can be therefore characterized as phonetically weaker than the same non-yer vowels.

Two findings, however, do not show a similar type of weakness in fast rate and yer-vowels. In duration, faster rate did not have a similar effect as yer origin since vowels in fast rate were significantly shorter than in the normal rate but yer vowels were not shorter than non-yer ones, and yer /o/ were even longer on *DurArt* than non-yer /o/. It was suggested in Section 3.2.4 that the consonantal movements around yer /o/ vowels were slower and less stiff, while this decreased stiffness was accompanied by tighter coordination of the two consonantal movements for yer /e/ but not yer /o/. As concerns speech rate, consonantal movements around vowels in normal rate were also slower (F=90.1, $p<0.01$ for the opening movement preceding the vowel and F=13.1, $p<0.01$ for the closing movement following the vowel) and less stiff (F=46.0, $p<0.01$ for the opening movement and F=18.2, $p<0.01$ for the closing movement). However, this decreased stiffness in normal rate compared to fast rate was

[7] We are merging here the results from the acoustic and articulatory data and although not all acoustic and articulatory measures showed statistical significance, the pattern of more centralization both for yers and fast-rate vowels is clearly present in the data.

[8] Note that Tempo shows a greater effect on /e/ than on /o/ on this measure, which is also observed in the effect of Category.

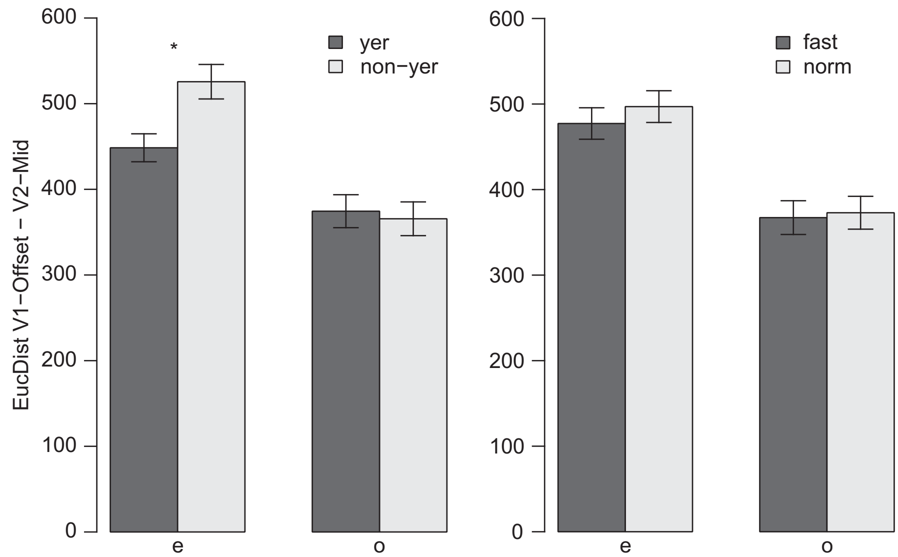

Fig. 6. The effect of category (yer/non-yer) in the left panel and speech rate (fast/norm) in the right panel on the Euclidean distance calculated in two dimensional $F1-F2$ space between the values at the midpoint of the target vowel and 10 ms before the offset of the preceding vowel (V_1-V_TEucDist).

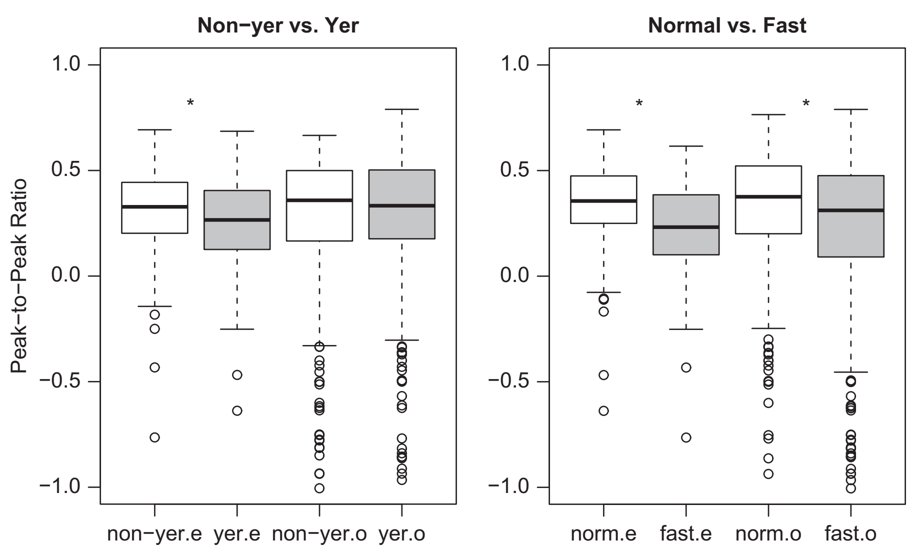

Fig. 7. The effect of category (left) and speech rate (right) on the Peak-to-Peak ratio measure separately for the two vowels.

accompanied by greater *DurNuc* interval (as reported in Table 2). Hence, there seems to be a continuum of weakness. Fast rate vowels are most weak as they are realized with increased stiffness and increased temporal coproduction of the flanking consonantal movements. This strategy is employed for yer /e/ vowels as well. Yer /o/ vowels are medially weak showing decreased stiffness but no change in the temporal coordination of the flanking consonantal movements. Vowels in normal rate are the least weak ones showing both lower stiffness and decreased coproduction of the consonantal movements.

The second result that seems to go against the prediction that the phonetic weakness due to fast rate is similar to the weakness due to yer origin concerns the slopes of the horizontal trajectories in which TB2 showed flatter slopes both in fast rate vowels and yer-vowels. However, recall, that while TB2 showed steeper slopes for yer /e/, significantly flatter slopes were reported for yer vowels on TB1. Since TB1 and TB2 sensors are not independent as they track the movement of tissue forming a single articulator, the significant effects of Category on the slope of horizontal movement in the opposite directions on these two sensors yields these findings suspicious. Moreover, the effect of yer origin that disagrees with the one of tempo in the horizontal movement of TB1 sensor might also be attributable to segmental differences between the yer-nonyer pairs.[9] Given the disagreement between the two tongue body sensors, potential confounding effect of segmental environment, and the disagreement between the patterns for fast and yer vowels (on TB1), the analysis of yer weakness in the slopes of the horizontal movement of the

[9] Further analyses showed that steeper slopes of the TB1 sensor in yer /e/ tokens than non-yer tokens were produced in three lexical pairs out of five: [tsumel–tʃumel], [ra:mets–na:met], and [obets–obet]. In the first pair, the initial [tʃ] of the non-yer tokens compared to [ts] on the yer tokens could cause a more retracted position of the tongue body and consequently a steeper slope for the frontward horizontal movement towards the target /e/ vowel. In the remaining two pairs, it is plausible that the frontward movement towards a longer final consonant [ts] has to start slightly sooner than the same movement toward a shorter [t], which again causes the horizontal movement for the yer vowels to have slightly steeper slopes.

tongue body is inconclusive and the sub-hypothesis concerning weakness of yers on this measure is regarded as not supported.

4. Discussion

This paper set out to test the assumption of phonological analyses that vowels /e/ and /o/ in Slovak with the phonological status of yer vowels are phonetically identical to the same vowels that do not have this status. We hypothesized that yer vowels might be phonetically weaker than their non-yer counterparts and tested this weakness with measures of vowel duration, vowel quality and the patterns of coarticulation with surrounding sounds in both the acoustic and articulatory domain, and compared these patterns with the patterns arising from the variation in speech rate. Our results show that the assumption of the phonological analyses about the phonetic identity between yer and non-yer vowels cannot be unequivocally supported in our data. The alternate hypothesis, that yer vowels in Slovak are phonetically weaker than their non-yer counterparts, received moderate support. Although no single result provides conclusive evidence and some inconsistencies were found, when taken together, the results favor supporting this alternate hypothesis.

The majority of the statistically significant effects of phonological category (yer vs. non-yer) are consistent with the analysis that yer vowels are phonetically weaker than non-yer ones. Yer /e/ vowels resisted the coarticulation with the preceding vowels less than their non-yer counterparts, and they were also articulated in more centralized positions horizontally. Moreover, the consonantal movements surrounding yer /e/ vowels were more coproduced (i.e. mutually truncated) than non-yer /e/ vowels. Finally, the vocalic yer /e/ horizontal movements were more coproduced with the movements of adjacent lingual consonants than the same movements for non-yer vowels as shown in the slopes on the vertical trajectories of the tongue body sensors. Given both greater coproduction of consonants surrounding yer /e/ than non-yer /e/ as well as lower stiffness of these movements, yer /e/ is accompanied by localized decrease of stiffness in the consonantal movements before and after the vowel together with tighter temporal coordination of these two movements.

Yer /o/ vowels were more fronted than their non-yer counterparts, they had flatter curvatures of the time functions extracted from the horizontal movement of the TB2 sensor, and more overlap of the surrounding consonantal movements on the *DurNuc* measure. Furthermore, the consonantal movements that preceded and followed yer /o/ vowels were in general slower, longer and less stiff, and thus, they were co-produced with the vocalic movements over a longer period of time than the movements surrounding non-yer /o/ vowels. Despite some similarities, the differences between yer /e/ and /o/ might suggest that the class of yers has lost (or is losing) its coherence as a unitary class. The individual members may have drifted in their own evolution paths, by forces inherent to the phonetics of the specific vowel and/or lexical forces, and that is why an obvious single phonetic feature (or a set of features) characterizing the yer class has not been found.[10]

Finally, no significant interaction between phonological category and speech rate was reported, suggesting qualitative similarity between two types of weakness. Moreover, yer vowels behaved similarly to vowels in fast rate on several measures, especially those testing the coarticulation patterns. For example, both yer vowels and vowels in fast rate showed a greater overlap of surrounding consonants (*DurNuc* and for /e/ on *Peak-to-Peak ratio* measures), flatter slopes and curvatures of time functions extracted from the movements of tongue body sensors, or shorter acoustic durations. Hence, the hypothesis that qualitative patterns of weakness due to faster speech rate are comparable to the weakness due to the yer origin of the vowels found support in our data.

Several partial results, however, did not show a significant difference in the weakness of yer and non-yer vowels, and two significant effects went against the analysis advocated above: yer /e/ had a steeper slope of the time function extracted from the horizontal trajectory of the TB1 sensor than the slopes from non-yer /e/ vowels, and the interval between the release of the consonantal gesture preceding yer /o/ and the achievement of the target for the consonantal gesture following yer /o/ were longer than the same intervals extracted from non-yer /o/ tokens. However, as discussed in Section 3.2.4, the first might be linked to a greater involvement of the TB1 section of the tongue body in the production of alveolar consonants or to differences in the consonantal environments between the yer and non-yer lexical pairs. The second of these effects should be seen in the context of inconsistent results relating to the overlap of consonants surrounding /o/: this overlap was smaller for yer /o/ than non-yer /o/ on *DurArt*, not different on *AccDur*, and greater for yer /o/ than non-yer /o/ on *DurNuc*. In relation to that, we also observed that the consonantal movements around yer /o/ vowels were longer, slower, and less stiff than the same movements around non-yer /o/ vowels. Hence, for /e/ the greater overlap of consonants with the target yer vowels was observed in greater temporal coproduction of the flanking consonants (seen most clearly in *Peak-to-Peak ratio* results). For /o/, greater overlap of yer than non-yer vowels with the surrounding consonants seems to be caused only by lower stiffness of the consonantal movements surrounding yer vowels with no changes to the temporal coproduction of the two movements.

The significant differences between yer and non-yer vowels suggesting greater phonetic weakness of the former over the latter category were not particularly robust and were rather small in size. This, however, was to be expected given that not only naïve speakers but phonetically trained phonologists assumed that the two categories were phonetically identical. Moreover, the observed effects were spread over several dependent variables that may be correlated with each other in non-trivial ways; recall for example the relationship between various measures of duration and consonant–vowel coproduction. Conceivably, a version of a multivariate analysis with this data might provide a better understanding of these relationships, but the complexities stemming from the nature of this multi-speaker dataset put such an analysis beyond the scope of the current paper. Finally, the stimuli pairs are not true minimal pairs in the sense that we could not control for the quality of the initial consonants. It is known that coarticulation has quite a large span, and some of the observed effects might plausibly arise from the differences in the initial consonants. For example, /e/ in *kábel* may differ from /e/ in *Ábel* in subtle ways precisely because of the presence of the initial /k/ in the former and its absence in the later token. Although some partial results may be plausibly attributed to these differences, the random distribution of the differences in the stimuli pairs prevents a coherent account along these lines. Despite the limitations mentioned in the last two paragraphs, the account based on greater phonetic weakness of yer vowels compared to non-yer vowels provides the most coherent explanation of the patterns observed in our data.

Given that yer and non-yer vowels do in fact represent two phonological categories, the natural question is how these subphonemically different categories could arise and be formalized. The most plausible source of the difference is the competition

[10] Thanks to A. Gafos (p.c.) for pointing out this view to me.

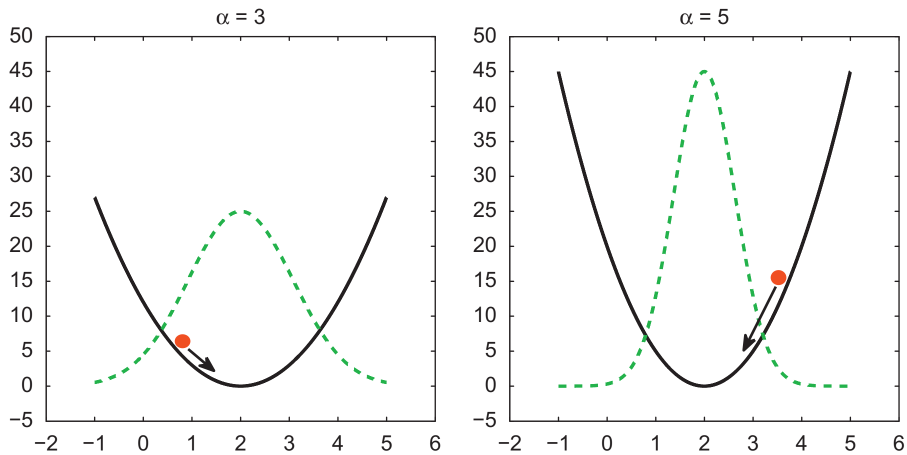

Fig. 8. Potentials (black full lines) and probability distributions (gray (green) dotted lines) of the dynamic system $V(x)=\alpha(x-2)^2$+noise as a variation of the weight α. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

between the full and zero realization of the yer vowels. Yer vowels surface only with morphologically zero suffixes – e.g. *kábel-ø*, *kábl-a*, **kábel-a* – which is for nouns, as discussed in Section 1.2, only in 3 out of 36 possible forms. Hence, the two consonants flanking a yer vowel, /b/ and /l/ in case of *kábel*, are more commonly realized adjacent to each other as a cluster (C_1C_2) than separated by a yer vowel ($C_1V_YC_2$). Adjacent consonants are in general more coproduced than the ones separated by a vowel (requiring the opening of the vocal tract which significantly limits the overlap of these consonants).[11]

In a traditional modular approach of generative phonology, the differences required at the phonological level are assumed to be encoded using abstract symbolic representations and discrete logical operations with them. Crucially, they are assumed to be no longer present at the stage at which the surface discrete representations are transformed to the continuous movements of the articulators via transduction. This view of phonology–phonetics relationship is similar to other aspects of language and cognition in general (e.g. Fodor & Pylyshyn, 1981; Harnad, 1990). Our findings suggest that these deeply and discretely encoded phonological differences persist, albeit in very fine details, in the continuous production patterns.

A potential conceptualization of this observation lies in the nature of the relationship between phonology and phonetics. We assume that the two sides of the cognitive system underlying speech (more granular phonology and less granular phonetics) reinforce each other: traces of some cognitive states, that were assumed to be wiped-out, are nevertheless allowed to be encoded phonetically, and these phonetic differences in turn facilitate the acquisition and retention of this rather intangible phonological contrast.[12] Since a traditional modular approach lacks formal tools for expressing and/or encoding such sub-phonemic differences, we briefly discuss two approaches to our data in which this encoding is possible. The first one is based on Articulatory Phonology (AP, Browman & Goldstein, 1986, 1995, 2000), and the second on Exemplar Theory (Johnson, 1997; Kirchner & Moore, 2009; Pierrehumbert, 2001).

Articulatory Phonology assumes that the basic units of the cognitive system representing speech are articulatory gestures. A gesture is a task-oriented dynamically defined unit of action that has both spatial and temporal dimensions. A task of producing a vowel gesture such as /e/ involves, in part, the achievement of a mid-constriction (between wide and narrow) between the tongue body and the front area of the hard palate. The formation of constrictions such as these involves a change in the position of one or more active articulator(s) over time. Therefore, the task of producing speech sounds can be modeled using the mathematical theory of dynamics.

In the model of Articulatory Phonology, the spatial target of a gesture is characterized by two variables: constriction location (CL) and constriction degree (CD), and the movement of every active articulator (lips, tongue tip, tongue body, etc.) is specified for CL and CD variables. The vocal tract variables CD and CL are coupled to prosodic and speech rate effects, which yields a gestural score. This score then serves as input to the task dynamics module that calculates the time-varying response of the vocal tract articulators to a set of gestural control structures (Saltzman & Kelso, 1987).

For our purposes, and simplifying for expository reasons, each gesture is a unit of action characterized by a discrete target, and these targets correspond to the stable values of the CD and CL parameters. The dynamic system that describes the behavior of the parameters CD and CL may be simplified to the first-order differential equation $dx/dt=-k/b(x-x_o)$, which describes a gesture as a movement toward the target x_o of a spatial parameter $x=$ {CL, CD} over time with the stiffness term k (e.g. Gafos, 2006; Benus, 2005). This articulatory movement can be imagined as a ball moving in a potential landscape $V(x)$, which can be derived with the general equation of motion: $dx/dt=f(x)+\text{noise}=-dV(x)/dx+\text{noise}$, which for our first-order equation gives $V(x)=k/2b(x-x_o)^2$. The potentials $V(x)$ drawn in a solid (black) line in Fig. 8 represent an arbitrary situation where the target value $x_o=2$, and $\alpha=k/2b$.

[11] Certain weakness might also be related to a potentially more flexible association between consonants and prosodic positions in yer paradigm compared to the non-yer one. As pointed out by the editor, Berg and Abd El Jawad (1996) found that syllabic affiliations of consonants within words imposed greater constraints for the frequency of speech word-internal errors involving consonants in German than in the non-concatenative language Arabic. Although Slovak is not a non-concatenative language, the paradigmatic difference in the yer class (e.g. *kábel* (CVCVC) for yer and *kábla* (CVCCV) for other forms) may induce a similarly 'looser' associations between segments and the syllabic structure in this class than in the non-yer class in which this paradigmatic difference is missing.

[12] The fact that the distinction between yer and non-yer vowels is rather abstract and difficult to acquire is supported by author's personal observations from first and second language acquisition and proper names conjugations. Children rather late in their language development make mistakes in omitting yers (**palc* instead of *palec*, **lakt'* instead of *lakeť*). Moreover, even very proficient Hungarian speakers of Slovak commonly make mistakes by failing to omit yers in affixed forms (**Ružomberoku* instead of *Ružomberku*), and many times there is a vacillation between the presence and absence of yers in affixed forms of various proper names even in printed newspapers (e.g. Hašeka/Haška, Mareka/Marka).

These potentials thus represent the dynamic mechanism that underlies a task-defined gesture, and the movement of a ball in that potential represents the actual movement of an articulator toward a target. Due to the presence of inherent noise in the system and variable initial conditions, we calculate the probability with which the system reaches a particular value of x. The probability function could be estimated analytically as well as through a simulation (Gafos & Benus, 2006), and it is shown as a (green) dotted line in Fig. 8. The strength of an attractor thus corresponds to the steepness of the probability function. The comparison of the two panels in Fig. 8 shows that the strength of an attractor can be modeled with parameter $\alpha=k/2b$ of the dynamic system $V(x)=\alpha(x-2)^2+\text{noise}$. If $\alpha=3$ (left panel), the attractor is weaker than if $\alpha=5$ (right panel). This, however, applies only to a constant noise term, and increasing noise decreases the strength of the attractor. In our situation, increased noise for yer vowels, given the mentioned bias between zero (CC) and full (CVC) realizations, compared to non-yer vowels, gives rise to lower strength of the attractor for yer vowels.

Hence, one way of conceptualizing the observed phonetic weakness of yer vowels compared to non-yer ones is through the strength of the dynamic attractor underlying the vocalic gesture. The mechanism relating the strength of the attractor and the competition between the forms with and without yer within the yer-words paradigm is the presence of noise in any dynamic system. Standard measures of movement variability reflect both noise and attractor strength, and these two variables influence each other over time.[13] If yer vowels had this attractor weaker, their articulation would be more prone to blending from adjacent gestures (both consonantal and vocalic), which is what we observed in the data. However, the stiffness of the movements does not depend only on the underlying strength of the attractor,[14] but also on prosody and speech rate factors, formalized for example through the π-gesture model of Byrd and Saltzman (2003). Therefore, the observed variability in the weakness of yer vowels might plausibly be linked to the fact that in the AP model, there are multiple underlying sources for phonetic weakness: differences in underlying stiffness of gestures, differences in the amount of noise due to lexical and frequency effects of individual words, or differences in how the prosody component interacts with and modulates the stiffness of gestures and their blending with adjacent gestures. Investigation of these issues in future may suggest adjustments and modifications of the formal AP model.

Another potential conceptualization of the data is based on the Exemplar theory and models that assume that lexical representations encode phonetic details beyond the scope of standard segmental and featural representations (Johnson, 1997; Pierrehumbert, 2001). Recent developments of these models take seriously also the production of speech and provide a "seamless phonetics-phonology interface" (Kirchner & Moore, 2009) by computing the outputs on the continuous signals extracted from exemplar clouds rather than symbolic representations. In this model, a more frequent zero realization of a yer (as a CC cluster) creates a bias among the memorized exemplars of the word so that even when a full realization with the vowel is required (CVC), the selected representation is likely to be similar to the representation without a vowel.[15]

Before this explanation could be considered, however, several issues should be addressed. First, the greatest phonetic difference between a C_iC_j cluster and a C_iVC_j sequence is the presence of a vowel and the perceptual salience of its duration. Hence, of the three dimensions of phonetic weakness, we would expect the effect of yer/non-yer origin in vowel duration to be the most salient. Yet, duration did not show remarkably greater differences than other dimensions of weakness. Second, the differences were observed not only in the temporal coordination of individual gestures but also in intra-gestural characteristics such as stiffness. It is not clear how these effects could be modeled by mere concatenation of multiple substrings extracted from the clouds of exemplars for a given word as proposed by the currently most advanced model of production within the Exemplar Theory (Kirchner & Moore, 2009). Finally, the explanation couched in the exemplar models depends on the assumption that all observed differences could be traced to the surface difference between the production without a vowel (CC, e.g. ká[bl]a) and the production with a vowel (CVC, e.g. ká[bel]). Although many observed differences could be attributed to the difference between CC and CVC structures, it is less clear that the observed variation of V-to-V coarticulation patterns or consonantal stiffness could also be attributed to the CC vs. CVC contrast.

Our data were not designed to tease apart these two approaches to the relationship between sub-phonemic differences and phonological alternations. However, several potentially problematic issues mentioned above for the Exemplar approach are possible to deal with, and even predicted by, the AP approach. For example, the observed differences in V-to-V coarticulation are straightforwardly formalizable through the blending of adjacent vowels (e.g. Benus, 2005; Fowler, 1983). Additionally, patterns in stiffness and coordination of surrounding consonantal movements and similarity to weakness due to speech rate are possible to capture using the π-gesture model. It seems that the AP approach is more flexible and fits the observed patterns better than the Exemplar approach possibly at the expense of lower degree of constrain in the former.[16]

Finally, an important question related to our results and discussion above is how these subtle phonetic differences, which are not likely to be perceptually salient, could have been acquired. Our data suggest that a phonological contrast (yer vs. non-yer), which was assumed to be completely neutralized phonetically, nevertheless displays minute differences in the production. This is reminiscent of incomplete and near mergers (e.g. Charles-Luce, 1997; Ernestus & Baayen, 2006; Labov, Karen, & Miller, 1990; Pierrehumbert, 2003; Port & Crawford, 1989; Warner, Jongman, Sereno, & Kemps, 2004, etc.). Labov et al. (1990) observed that, despite the fact that certain phonetic contrasts have been claimed to be neutralized and subjects do not perceive the contrast, the same subjects consistently maintain the contrast in their productions for sociolinguistic reasons. Pierrehumbert (2003) proposed that, in order for the contrast to persist in production, the maintenance of the contrast must have been motivated in the past while speakers were younger but was subsequently lost. We

[13] Recent methods such as recurrence analysis (e.g. Richardson, Schmidt, & Kay, 2007) began teasing apart the attractor strength and the amount of noise, which promises to lead to better understanding of the mutual relationship between these two.

[14] Browman and Goldstein (1990) suggested that consonantal and vocalic movements differ mainly in terms of stiffness, which suggests that stiffness should be included in the underlying dynamic articulatory representations of speech sounds.

[15] Moreover, the first mention of the target yer word in our frame sentence was also produced in the affixed form without the yer vowel.

[16] The Exemplar model is indeed constrained by the input data, and thus able to encode patterns obtained in that data. Yet, this makes the model also somewhat weak since, given a different kind of input data, nothing in the model prevents encoding different, and possibly unnatural patterns. In other words, the Exemplar model is good at accounting for the generalization in the input data but has difficulty explaining the patterns in phonetics–phonology interface. The AP model is in this sense more constrained since it is firmly based on the physiological and dynamic mechanisms underlying speech.

speculate that this motivation is linked to the idea of mutual phonetic–phonology reinforcement during the acquisition of this contrast, mentioned above, in which phonetic traces of some cognitive states are allowed to be encoded, and these phonetic differences in turn facilitate the acquisition and retention of the phonological contrast. Alternatively, the re-occurrence of this pattern in successive generations might be possible without recourse to a requirement that children actually perceive an acoustic–auditory distinction between yer and non-yer vowels in the speech of adults using the mentioned frequency bias between tokens with full and zero yer realizations. Unfortunately, our data were not designed to tease apart these speculations about learnability and they need to be carefully tested in future experiments. The primary novel result of the current paper is that, phonetically, yer vowels might be subtly weaker than non-yer vowels.

Acknowledgments

This work was supported by an Alexander von Humboldt Fellowship and the preparation of the manuscript was also supported by the VEGA No. 2/0202/11 grant. The author wishes to thank Diamandis Gafos, Jonathan Harrington, Phil Hoole, Stefania Marin, and Marianne Pouplier and anonymous reviewers for valuable comments to earlier drafts of this paper and Susanne Waltl and Yuki Era for assistance with data collection and annotation. All mistakes are mine.

References

Adams, S. G., Weismer, G., & Kent, R. D. (1993). Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research*, *36* 41–54.

Archangeli, D., & Pulleyblank, D. (1994). *Grounded phonology*. Cambridge, MA: MIT Press.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: CUP.

Barnes, J. (2006). Strength and weakness at the interface: positional neutralization in phonetics and phonology. Berlin: Mouton de Gruyter.

Beňuš, Š. (2011). Control of phonemic length contrast and speech rate in vocalic and consonantal syllable nuclei. *Journal of the Acoustical Society of America*, *130*(4), 2116–2127.

Benus, S. (2005). *Dynamics and transparency in vowel harmony.* Unpublished Ph.D. Thesis, New York University.

Benus, S., & Rusko, M. (2008). The acoustics of mid vowels [e] and [o] in Slovak. In *Proceedings of the 155th conference of the Acoustical Society of America*. Paris, France.

Beňuš, Š., & Mády, K. (2010). Effects of lexical stress and speech rate on the quantity and quality of Slovak vowels. In *Proceedings of speech prosody 2010,* Chicago, USA.

Berg, T., & Abd El Jawad, H. (1996). The unfolding of suprasegmental representations: A crosslinguistic perspective. *Journal of Linguistics*, *32*, 291–324.

Blevins, J. (2004). Evolutionary phonology: the emergence of sound patterns. Cambridge: CUP.

Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer, 〈http://www.praat.org〉.

Browman, C. P., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, *3*, 219–252.

Browman, C. P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In: J. Kingston, & M. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech* (pp. 341–397). Cambridge: Cambridge University Press.

Browman, C. P., & Goldstein, L. (1995). Gestural syllable position effects in American English. In: F. Bell-Berti, & L. J. Raphael (Eds.), *Producing speech: Contemporary issues (for Kathering Safford Harris)* (pp. 19–33). Woodbury, NY: AIP Press.

Browman, C. P., & Goldstein, L. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Les Cahiers de l'ICP, Bulletin de la Communication Parlée*, *5*, 25–34.

Byrd, D., & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, *31*, 149–180.

Charles-Luce, J. (1997). Cognitive factors involved in preserving a phonemic contrast. *Language and Speech*, *40*, 229–248.

Cho, T. (2006). Manifestation of prosodic structure in articulation: evidence from lip kinematics in English. In: L. Goldstein, D. Whalen, & C. Best (Eds.), Varieties of phonological competence (pp. 519–540). Berlin, New York: Mouton de Gruyter.

Ernestus, M., & Baayen, H. (2006). The functionality of incomplete neutralization in Dutch: The case of past tense formation. In: L. Goldstein, D. Whalen, & C. Best (Eds.), *Varieties of phonological competence* (pp. 27–49). Berlin, New York: Mouton de Gruyter.

Fodor, J. A., & Pylyshyn, Z. W. (1981). How direct is visual perception? Some reflections on Gibson's 'ecological approach'. *Cognition*, *9*, 139–196.

Fowler, C. A. (1983). Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in sequences of monosyllabic stress feet. *Journal of Experimental Psychology: General*, *112*, 386–412.

Gafos, A. (2006). Dynamics in grammar: comment on ladd and ernestus & baayen. In: L. Goldstein, D. Whalen, & C. Best (Eds.), Varieties of phonological competence (pp. 51–79). Berlin, New York: Mouton deGruyter.

Gafos, A., & Benus, S. (2006). Dynamics of phonological cognition. *Cognitive Science*, *30*, 905–943.

Gussmann, E. (1980). *Studies in abstract phonology*. Cambridge MA: MIT Press.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, *42*, 335–346.

Harrington, J. (2010). *Phonetic analysis of speech corpora*. Oxford: Willey-Blackwell.

Harrington, J., Fletcher, J., & Roberts, C. (1995). Coarticulation and the accented/unaccented distinction: evidence from jaw movement data. *Journal of Phonetics*, *23*, 305–322.

Hayes, B., Steriade, D., & Kirchner, R. (2004). *Phonetically-based phonology*. Cambridge: CUP.

Hoole, P., & Zierdt, A. (2010). Five-dimensional articulography. In: B. Maasen, & P. H.H. M. van Liehout (Eds.), *Speech motor control* (pp. 331–349). Oxford: OUP.

Hoole, P., & Mooshammer, C. (2002). Articulatory analysis of the German vowel system. In: P. Auer, P. Gilles, & H. Spiekermann (Eds.), *Silbenschnitt und Tonakzente* (pp. 129–152). Tübingen: Niemeyer.

Jarosz, G. (2006). Polish yers and the finer structure of output–output correspondence. In *Proceedings of the Berkeley Linguistics Society*.

Johnson, K. (1997). Speech perception without speaker normalization. In: K. Johnson, & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–166). San Diego: Academic Press.

Král', A., & Sabol, J. (1989). *Fonetika a fonológia [Phonetics and phonology]*. Bratislava: Slovenské pedagogické nakladatel'stvo.

Kirchner, R., & Moore R. K. (2009). *Computing phonological generalization over real speech exemplars*. Ms. University of Alberta.

Labov, W., Karen, M., & Miller, C. (1990). Near mergers and the suspension of phonemic contrast. *Language Variation and Change*, *3*, 33–74.

Lightner, Th. M. (1965). *Segmental phonology of contemporary standard Russian*. Ph.D. Dissertation, MIT Press.

Lindblom, B. (1963). A spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, *31*, 773–1781.

Öhman, S. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, *39*, 151–168.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In: J. Bybee, & P. J. Hooper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137–158). Amsterdam: John Benjamins.

Pierrehumbert, J. (2003). Probabilistic phonology: Discrimation and robustness. In: R. Bod, J. Hay, & S. Jannedy (Eds.), *Probability theory in linguistics*. Cambridge, MA: MIT Press.

Port, R., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, *17*, 257–282.

Recasens, D. (1985). Coarticulatory patterns and degrees of coarticulatory resistance in Catalan CV sequences. *Language and Speech*, *28*(2), 97–114.

Recasens, D. (1999). Lingual coarticulation. In: W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques in speech production* (pp. 78–104). Cambridge: Cambridge University Press.

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, *102*, 544–561.

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication*, *52*, 638–651.

Richardson, M. J., Schmidt, R. C., & Kay, B. A. (2007). Distinguishing the noise and attractor strength of coordinated limb movements using recurrence analysis. *Biological Cybernetics*, *96*, 59–78.

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

Saltzman, E., & Kelso, S. (1987). Skilled actions: A task-dynamic approach. *Psychological Review*, *94*(1), 84–106.

Scheer, T. (2006). How yers made Lightner, Gussmann, Rubach, Spencer and others invent CVCV. In: P. Bański, B. Łukaszewicz, & M. Opalińska (Eds.), *Studies in constraint-based phonology* (pp. 133–207). Warsaw: Wydawnictwo Uniwersytetu Warszawskiego.

Szpyra, J. (1992). Ghost segments in nonlinear phonology: Polish yers. *Language*, *68*, 277–312.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, *17*, 3–45.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, *32*, 251–276.

Yearley, J. (1995). Jer vowels in Russian. In: J. N. Beckman, L. Walsh Dickey, & S. Urbanczyk (Eds.), *Papers in optimality theory*. Amherst: GLSA, University of Massachusetts pp. 533–571). Occasional papers in linguistics 18.