Control of phonemic length contrast and speech rate in vocalic and consonantal syllable nuclei

Štefan Beňuš[a)]
Department of English and American Studies, Constantine the Philosopher University,
Štefánikova 67, 94974 Nitra, Slovakia

(Received 25 September 2010; revised 19 July 2011; accepted 25 July 2011)

This paper investigates the mechanisms controlling the phonemic quantity contrast and speech rate in nonsense p_1Np_2a words read by five Slovak speakers in normal and fast speech rate. N represents a syllable nucleus, which in Slovak corresponds to long and short vowels and liquid consonants. The movements of the lips and the tongue were recorded with an electromagnetometry system. Together with the acoustic durations of p_1, N, and p_2, gestural characteristics of three core movements were extracted: p_1 lip opening, tongue movement for (N)ucleus, and p_2 lip closing. The results show that, although consonantal and vocalic nuclei are predictably different on many kinematic measures, their common phonological behavior as syllabic nuclei may be linked to a stable temporal coordination of the consonantal gestures flanking the nucleus. The functional contrast between phonemic duration and speech rate was reflected in the bias in the control mechanisms they employed: the strategies robustly used for signaling phonemic duration, such as the degree of coproduction of the two lip movements, showed a minimal effect of speech rate, while measures greatly affected by speech rate, such as p_2 acoustic duration, or the degree of p_1-N gestural coproduction, tended to be minimally influenced by phonemic quantity.

© *2011 Acoustical Society of America*. [DOI: 10.1121/1.3624824]

PACS number(s): 43.70.Bk, 43.70.Fq, 43.70.Kv [AL] Pages: 2116–2127

I. INTRODUCTION

The understanding of the temporal patterning in speech and formal mechanisms relating the functional requirements to the observable acoustic and articulatory events represents one of the major challenges in the study of human speech (e.g., Port *et al.*, 1995). This paper investigates the mechanisms controlling the realization of syllable nuclei as a function of phonemic quantity contrast and variations in speech rate. Data come from Slovak, which is a language that allows long and short vowels {[i], [e], [a], [o], [u], [i:], [e:], [a:], [o:], [u:]} as well as long and short liquid consonants {[r], [r:], [l], [l:]} as syllable nuclei, and in which vowels and syllabic liquids display similar linguistic behavior (Rubach, 1993; Pouplier and Beňuš, 2011; Beňuš and Mády, 2010). This property of Slovak facilitates the investigation of three inter-related issues discussed below.

A. Functional similarity and articulatory control of vowels and liquids

The first functional similarity of liquids and vowels in Slovak is that both can stand as syllable nucleus. Hence, we can examine the quantity of both vowels and consonants without the confounding factor present in many other languages in which vowels form syllable nuclei and consonants form syllable onsets or codas. Studies of the articulatory control of duration in syllable nuclei typically investigate the consonantal gestures that surround these nuclei. For example Hertrich and Ackermann (1997) and Hoole and Mooshammer (2002) studied the short-long contrast in German vowels by examining the kinematics and coordination patterns of the consonantal gestures immediately surrounding the target vowels. However, our knowledge of the articulatory control of duration in the syllable nuclei themselves is minimal.

In consonants, studies of the articulatory control of the phonemic quantity contrast compare the realization of singletons and geminates (Löfqvist, 2005, 2007; Smith, 1995). However, these two categories do not differ only in quantity but in most phonological accounts also in their syllable affiliations: singletons are onsets while geminates are typically analyzed as coda-onset clusters. In Slovak, both long and short liquids form syllable nuclei and thus allow investigations of the linguistic contrast of consonantal length independent of syllable structure.

The second functional requirement, that Slovak vowels and liquid consonants share, is signaling phonemic quantity contrast. Articulatory gestures producing vowels and consonants have considerably different kinematics—mainly in terms of movement stiffness (Browman and Goldstein, 1990). Stiffness varies negatively with the duration of the movement (Ostry and Munhall, 1985): stiffer consonantal movements are shorter than less stiff vocalic ones. In the notional dynamic model most commonly used to describe task-oriented articulatory movements—a damped mass-spring (Saltzman and Kelso, 1987)—stiffness represents a major determinant of duration in single articulator movements with comparable initial and target locations [Browman and Goldstein (1990); see Simko and Cummins (2010) for a recent expanded model that includes articulator masses

---

[a)] Also at: Institute of Informatics, Slovak Academy of Sciences, Dúbravská 9, 84701 Bratislava, Slovakia. Author to whom correspondence should be addressed. Electronic mail: sbenus@ukf.sk and links more straightforwardly to muscle forces]. One goal of this paper is to improve our understanding of how different kinematic requirements for vowels and consonants are reconciled in shared functional requirements for phonemic length contrast.

B. Intra- and inter-gestural control of phonemic contrast and speech rate variation

In addition to stiffness, the second primary articulatory mechanisms for achieving durational variations are changes in the coordination between multiple gestures. In an abstract sequence of two adjacent gestures, varying just the overlap between them, all else being equal, the duration of the sequence will necessarily change. Although both phonemic contrast and speech rate affect intra-gestural and inter-gestural characteristics, the relative contribution of the control strategies seems to depend on the type of the functional requirement. Paralinguistic contrasts tend to be implemented preferably with intra-gestural changes related to stiffness, while some linguistically meaningful contrasts tend to involve preferably inter-gestural ones. For example, Shaiman (2001, 2002) compared the effects of speech rate variation and the number of consonants in the coda on the kinematics and temporal coordination of lips and jaw movements surrounding the target vowels in English. She found that speech rate affected the labial movements globally by significantly increased peak velocity, stiffness, and decreased time to peak velocity for faster rates. However, little evidence for increased coproduction of the labial movements as a function of increased speech rate was found. Hoole and Mooshammer (2002) reported that the durations of the consonantal opening movement into a vowel and consonantal closing after the vowel did not significantly differ for tense and lax vowels under the variation in speech rate. Finally, speech rate also consistently affects intra-gestural characteristics such as velocity profiles so that speech gestures at fast rates have fewer and more symmetrical velocity peaks than in slower rates (Adams *et al*., 1993; Hoole and Mooshammer, 2002; Munhall and Löfqvist, 1992). Despite these observations, one should keep in mind pervasive interspeaker variability in their strategies for achieving speech rate variations (e.g., Adams *et al*., 1993; Ostry and Munhall, 1985; Shaiman, 2001).

In linguistically meaningful contrasts, on the other hand, Shaiman found that the overlap between opening and closing jaw movements was the sole control strategy for vowel shortening due to increased coda complexity. Hoole and Mooshammer (2002) found that the truncation of the opening consonantal movement preceding a vowel by the closing consonantal movement following the vowel was the most salient feature of the phonemically lax (short) German vowels compared to their tense (long) counterparts. Moreover, differences in the syllable affiliation of consonants, i.e., if a consonant belongs to an onset or a coda of a syllable, have been linked to the differences in the temporal coordinations of the gestures involving the production of these consonants (see Krakow, 1999, for a review).

In consonants, on the other hand, the control of the phonemic quantity contrast seems to primarily involve intra-gestural characteristics. Löfqvist (2005) reported that short singleton consonants were produced with greater stiffness than long geminate ones. Smith (1995) found that the closing movements of the lips were slower for the geminate than for the single consonant in Japanese and Italian. However, several studies pointed out that stiffness scaling, while prominent, cannot be the sole control of phonemic quantity. For example, longer duration of oral closure or constriction for geminate consonants compared to singletons may also be due to the adjustments of the deceleration phase of the closing movement (Löfqvist, 2005). The lengthening of the relative acceleration phase of the lip closing gesture following a vowel was suggested as one of the strategies controlling the quantity contrast in German vowels (Hertrich and Ackermann, 1997). The stiffness of the lingual movement also did not differ as a function of geminate/singleton contrast in Moroccan (Zeroual *et al*., 2008).

By examining the realization of Slovak long/short and normal/fast vowels and liquids in syllable nuclei, this paper adds to our understanding of the interplay and relative contribution of stiffness scaling and the degree of coproduction as the primary controls of durational variation.

C. Locus of phonemic length control

Phonologically, length contrast is assumed to affect the realization of individual segments. However, the articulatory control of phonemic vowel length is not necessarily localized on the gesture representing the vowel but may be distributed globally to affect the organization of multiple gestures. For example, phonemic length in German vowels differs from the mechanism underlying intrinsic vowel durations, which suggests that phonemic vowel length affects more than just the vocalic movement (Hertrich and Ackermann, 1997). Moreover, the coordination patterns between the consonantal movements flanking the vowel participate in controlling phonemic quantity contrast in German (e.g., Hoole and Mooshammer, 2002). These findings support the proposal that vowel duration is a property of syllables rather than individual segments (e.g., Venneman, 1991). In German, short nuclei are typically followed by consonants in the same syllable, i.e., coda consonants, while long nuclei are not. This functional relationship between a short vowel and the following coda consonant might provide rationale for the interdependency between the quantity of a vowel and the kinematics of the consonantal gestures around it. In Slovak, the phonemic length contrast in nuclei is independent of the presence of a following coda consonant. We can thus test if the difference in the phonotactic patterns between languages like German and Slovak affects the control strategies for cuing the length contrast.

Moreover, in addition to the coordinations between the consonants flanking the nucleus, the effects of phonemic quantity on the coordination of the nucleus gesture with respect to the surrounding consonantal gestures is also of interest. For example, Löfqvist and Gracco (1999) reported that in V_1CV_2 sequences with bilabial consonants larger tongue body movements for V_2 start earlier with respect to the C lip movement than smaller movements. Since long

Slovak nuclei (both vowels and consonants) are expected to involve larger movements, a significant effect of phonemic quantity on the degree of overlap of the onset and nucleus gestures might be expected. Alternatively, as discussed in Sec. I B, durational variation might be achieved by the degree of coproduction in the opposite direction: longer acoustic duration (due to phonemic length or rate) might also arise from smaller degree of overlap.

Hence, data from Slovak can provide more information about the prosodic nature of phonemic quantity, i.e., to what extent quantity is cued locally on the nucleus and how it affects the organization of the surrounding segments.

II. METHOD

A. Subjects

Two female and three male native speakers of Slovak served as subjects for this study. They were naïve as to the purposes of the study, all were between 20 and 40 yr old, none reported any speech, hearing, or language problems, and their speech did not deviate from customary patterns of standard colloquial Slovak.

B. Material

Material for this study consisted of nonsense words in the form of pNpa, in which N represents all 14 possible nuclei in Slovak: [i], [e], [a], [o], [u], [r], [l], [i:], [e:], [a:], [o:], [u:], [r:], [l:]. Subject produced pNpa words in the carrier phrase *Čítame _ pyšne* "We read _ proudly." Word stress in Slovak is always on the first syllable, and the target word received sentence stress.

Speech from normal and fast rates was collected. During a pre-test session, we elicited a subset of prompt sentences from each subject in a variety of speech rates following Adams *et al*. (1993). Based on the clarity of perceptual contrast between the rates and the consistency of prosodic patterns, we then selected two renditions of the sentences with the target word [pa:pa] as an example of normal and fast speech rates, respectively. For each subject, his/her example sentences were then presented auditorily and randomly before each 3–8 prompt sentences during the actual recording session and subjects were instructed to match the rate of their test sentences as closely as possible to the rates of the example sentences (Hoole and Mooshammer, 2002). The advantage is that this setup avoids the somewhat unnatural method of metronome prompting, thus providing a more ecological approach for eliciting laboratory speech rate variation.

Prompt sentences were presented visually in a randomized order in alternating five blocks of normal and fast speech rates on a computer screen in standard Slovak spelling. This procedure resulted in 140 tokens for each subject (5 repetitions, 2 rates, 14 nuclei) for a total of 700 tokens.

C. Data recording and processing

Acoustic and articulatory data were recorded at the IPS in Munich. Acoustic data were collected with a Sennheiser MKH 40 microphone with a sampling rate of 32 768 Hz and downsampled during post-processing to 16 384 Hz. Articulatory data were recorded with electro-magnetic articulography at a sampling rate of 200 Hz (AG 500, Carstens Medizinelektronik). Seven sensors were attached to the active articulators approximately along the mid-sagittal line: the upper and lower lips, the lower incisors to record jaw movement, and four roughly equally spaced sensors on the tongue: the most anterior sensor TT about 1 cm behind the actual tongue tip, the most posterior sensor TD in the velar/dorsal region of the tongue, TB1 closer to TT, and TB2 closer to TD. Additionally, four reference sensors were attached—behind each ear, on the nose, and on the maxilla—and were used in post-processing to correct for head movement during speech. Standard calibration and post-processing procedures were applied for each data recording session (Hoole and Zierdt, 2010). Movement data were filtered at 5 Hz for the reference sensors, 20 Hz for all active sensors, and 60 Hz for the tongue tip sensor due to greater flexibility of this articulator.

D. Data labeling and extraction

A trained annotator manually aligned boundaries between the consonants and vowels to the acoustic signal using primarily the visual information in the oscillogram and spectrogram. Consonant closures were labeled at the cessation of the formant structure and their releases at the discontinuous increase of energy associated with the burst. Articulatory labeling identified individual gestures using a semi-automatic procedure developed by M. Tiede that determines kinematic landmarks in the velocity profile of articulatory movements. The tongue movements of the liquids were labeled on the vertical velocity profile of the TT sensor, and those of the vowels on the tangential velocity of the TB1 sensor defined as a square root of the sum of squared horizontal and vertical velocities. Labial gesture used the velocity of the Lip Aperture measure (LA) which represents the Euclidean distance between the upper and lower lip sensors. Figure 1 illustrates one result of such labeling.

Given a manually selected temporal window comprising of the movement to be labeled, the algorithm first identified the peak velocities of the movement into and out of a constriction (PVEL1 and PVEL2, respectively, always corresponding to adjacent peak and valley in the velocity profile, shown with dashed lines in Fig. 1), and then identified the kinematic events movement onset (GONS), achievement of the constriction (NONS), release of the constriction (NOFF), and movement offset (GOFF) on the basis of a percentage threshold (default 20%) of the peak velocities. Finally, the point of minimal velocity between the NONS and NOFF landmarks was identified as the maximal constriction (MAXC, shown with white vertical lines inside the filled boxes in Fig. 1). If the automatic algorithm gave clearly wrong results during the labeling, thresholds were adjusted. In some cases, the relevant landmarks of the movement could not be reliably determined and were discarded from the analysis. This was mostly the case for the onset movement of the target vowel in *pe(:)pa* (54 of 100 tokens) because the preceding vowel in the carrier phrase was also [e], and the offset movement of the target vowel in *pa(:)pa*

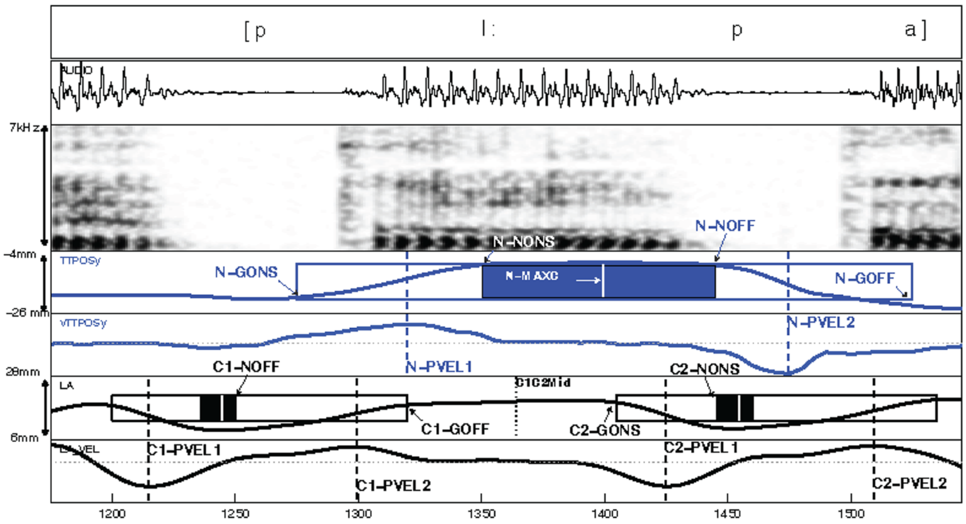

FIG. 1. (Color online) Example of kinematic landmarks in a token of [pl:pa]. Time in ms is on the x axis. Panels along the y axis from top to bottom: audio signal, spectrogram, vertical movement of the tongue tip (TT) sensor, vertical velocity of TT, Lip aperture (= euc. distance between the upper and lower lip sensors), and LA velocity.

(51 of 100 tokens). In analyses using these problematic landmarks, all data from the relevant vowel were excluded.

E. Acoustic and articulatory measures

Section I discussed three broad issues of this study. Table I lists acoustic and articulatory measures that were used to gain information relevant to these issues in our data. The description of the measures (shown in italics) in this section complements their definitions listed in Table I. Following the idea that relational measures better reflect phonemic contrasts than absolute measures (e.g., Port, 1981), we used *long/short-ratio* to assess if phonemic length is realized similarly for vowels and liquids. The acoustic duration of all three target segments (*durAcc*) was used for testing if phonemic length contrast was realized only on nuclei or affected also adjacent labial consonants.

The relevance of articulatory control in cuing the acoustic duration of nucleus was checked by testing the relationship between *durAcc* of the nucleus on the one hand and inter-gestural *durArt* (Hertrich and Ackermann, 1997) and intra-gestural *durPlateau, durGesture* on the other hand.

In order to test the effect of phonemic quantity and speech rate on the intra-gestural characteristics, we examined the kinematic characteristics of three movements: lip opening preceding the nucleus, tongue movement toward nucleus target, and lip closing following the nucleus. Specifically, we focused on the raw values of movement *duration*, its *peak velocity*, the *time-to-peak-velocity*, and *movement displacement*. Additionally, two derived measures of stiffness were used: *stiff-ratio* and *parameter C*. The first estimates stiffness of each movement directly based on the peak velocity and displacement of the movement. *Stiff-ratio* was

TABLE I. Summary of acoustic and articulatory measures; see Sec. II D and Figure 1 for the definition of articulatory landmarks

		Measure	Description/Definition
Acoustic		long/short-ratio	mean of long nuclei/mean of corresponding short ones
		durAcc	duration of p_1, N, and p_2 in the target p_1Np_2 sequence
Intra-gestural	Raw	duration	temporal difference: C1-GOFF – C1-NOFF, C2-NONS – C2-GONS
		peak velocity	velocity values at PVEL1, PVEL2
		time-to-peak-velocity	temporal difference: C1-PVEL2 – C1-NOFF, C2-PVEL1 – C2-GONS, N-PVEL1 – N-GONS
		displacement	spatial (Euclidean) difference: |C1-GOFF – C1-NOFF|, |C2-NONS – C2-GONS|, |N-GONS – N-NONS|
		durPlateau	nucleus only; temporal difference: N-NOFF – N-NONS
		durGesture	nucleus only; temporal difference: N-GOFF – N-GONS
	Derived	stiff-ratio	peak velocity/displacement
		parameter C	(peak velocity/displacement)/duration
Inter-gestural		durArt	C2-NONS – C1-NOFF
		symmetry	time-to-peak-velocity/duration
		peak-to-peak ratio	(C2-PVEL1 – C1-PVEL2)/(C2-NONS – C1-NOFF)
		pN-lag	temporal difference: N-PVEL1 – C1-PVEL1
		N-ONS-lag, N-MAXC-lag, N-NOFF-lag	temporal differences: *C1C2Mid* – N-NONS, *C1C2Mid* – N-MAXC, *C1C2Mid* – N-NOFF, *C1C2Mid* = C1-GOFF + (C2-GONS – C1-GOFF)/2

used rather than a more traditional measure of the slope of the regression line between the movement amplitude and its peak velocity in order to unify the statistical analyses of the factor effects. However, all reported qualitative patterns of *stiff-ratio* correspond to the patterns obtained with the slopes of the regression lines. *Parameter C* normalizes *stiff-ratio* for movement duration and describes the peakedness of the movement velocity profile (Adams *et al.*, 1993; Hertrich and Ackermann, 1997).

To test the effect of phonemic quantity and speech rate on the inter-gestural coordination of the opening and closing lip movements, discussed in Sec. I B, we employed two derived measures. First, the *symmetry* of velocity profiles provided an estimate of truncation between the two lip movements (Hoole and Mooshammer, 2002; Hertrich and Ackermann, 1997). Values of 0.5 correspond to perfectly symmetrical profiles. Late peaks (i.e., values greater than 0.5) in the lip opening movements suggest "regressive" truncation of the lip-opening movement by the following lip-closing one. Early peaks in the lip-closing movement suggest "progressive" truncation of the lip-closing movement by the preceding lip-opening one. The presence of both patterns suggests bi-directional truncation of the two movements. Second, *peak-to-peak ratio* assessed the temporal coproduction of the two movements by relating the interval between the velocity peaks of the lip movements to the interval between the closure release of lip opening and closure onset of lip closing. Both Harrington *et al.* (1995) and Hoole and Mooshammer (2002) found that greater values of this ratio correspond to greater coproduction.

Finally, we employed two measures of inter-gestural coordination between the lip gestures and the gestures for the syllable nuclei to assess the issues discussed in Sec. I C. First, *pN-lag* captured the degree of overlap between the initial consonant and the following nucleus using the velocity peaks of the movements toward the target as a relatively most stable gestural landmark. Second, to assess the coordination of the nucleus gesture with respect to both flanking lip movements, and the effect of quantity and speech rate on this coordination, we first calculated the temporal midpoint between the lip-opening and lip-closing gestures (*C1C2Mid*), which represented a crude measure of coordination between the two lip gestures. We then examined the temporal lag of nucleus closure achievement, maximal constriction, and closure release with respect to this midpoint: *N-NONS-lag*, *N-MAXC-lag*, and *N-OFF-lag*. A stable relationship between *C1C2Mid* and the landmarks of the nucleus gesture would suggest a dependency between inter- and intra-gestural controls of duration while systematic differences between the three lag measures would point to a bias between the two control mechanisms.

F. Statistical analysis

We tested the effect of four independent variables: Quantity (long/short), SpeechRate (normal/fast), NucleusType (consonantal/vocalic), and GestureType (cv-lip-opening/cv-lip-closing). A mixed-model approach to determining the effects of fixed factors and their interactions on dependent variables implemented in R software package was used as a primary tool for statistical analysis (Baayen, 2008). Output F-values were taken as the main indicators of the robustness of effects that the factors had on the dependent variables and for the comparisons among these effects. Following Reubold *et al.* (2010) a conservative value for the degrees of freedom 60, and an α level of 0.01 was used for the calculation of significance. Under these conditions, all $F > 8.49$ were considered as significant at $p < 0.01$. All reported tests had Subject as a random factor. In this way, the variation among the subjects was excluded from the F-value calculation. This is important for some non-normalized values such as movement displacement that depend on the physiology of articulators and the placement of the sensors.

III. RESULTS

A. Acoustic duration

We first explore the characteristics of the quantity contrast between long and short Slovak nuclei. The boxplot in Fig. 2(left) shows the nuclei *long/short ratio* separately for vowels and liquids. The quantity contrast was robustly present in both rates, and the *long/short ratio* was similar for the consonantal and vocalic nuclei. In normal rate, the ratios were on average greater than two; hence, long nuclei were more than twice as long as the short ones. In fast rate, the ratio was smaller, but still above 1.5 on average. A Repeated Measures Anova test with the *long/short ratio* as the dependent variable and NucleusType (consonantal vs vocalic) and SpeechRate (normal vs fast) as within factors showed no significant effect of NucleusType ($F^{1,1} = 0.8$, $p = 0.42$) and a significant effect of SpeechRate so that the ratio was greater in the normal rate than in the fast rate ($F^{1,1} = 10.2$, $p = 0.033$). There was no significant interaction between the two factors. An RM Anova was employed in this case rather than the mixed model design since the calculation of ratios required averaging of repetitions, which yielded a relatively small number of tokens (7 nuclei × 2 rates × 5 subjects = 70). For small populations

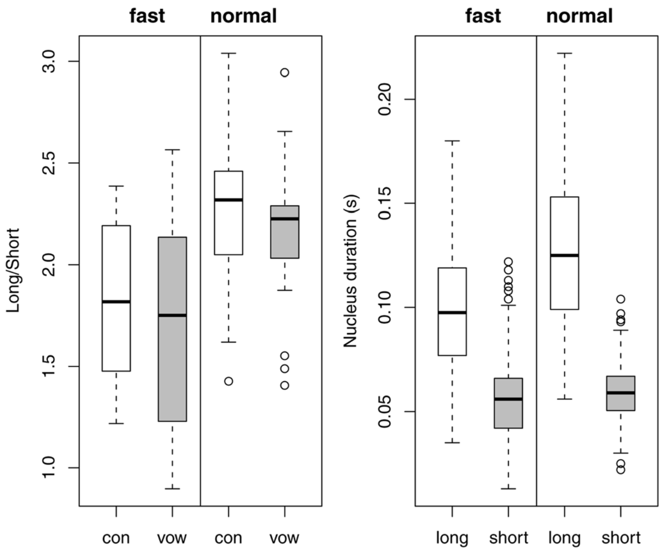

FIG. 2. (left) Ratio of long/short nuclei split for nucleus type (consonantal (con) vs vocalic (vow)) and speech rate (fast vs normal). (right) Raw duration of nuclei split for phonemic quantity (long vs short) and speech rate.

(less than 100), a repeated measures design is preferred over the ordinary mixed model design (Harrington, p.c.).

The boxplot in Fig. 2(right) shows the interaction between Quantity and SpeechRate in the absolute durations of nuclei. While long vowels were robustly shorter in the fast rate than in the normal rate, the difference between the short vowels in the two rates was minimal. A linear mixed models test with the dependent factor of nucleus duration showed significant main effects of Quantity ($F = 1265.6$), SpeechRate ($F = 160.0$), and their interaction ($F = 89.8$). While the effect of SpeechRate was robust for long vowels ($F = 237.5$), it was much weaker in short ones ($F = 10.3$). Hence, the increased speech rate decreased the duration of short vowels only marginally. These observations can be linked to the notions of greater stretchability of long vowels compared to the short ones and incompressibility of short vowels (e.g., Klatt, 1973) in that the effect of various shortening factors—in our case phonemic contrast and speech rate—tend to decrease when these factors are combined. Finally, the factor NucleusType had no significant effect on the duration of the nucleus.

To examine the prosodic nature of phonemic quantity contrast, we next tested if the phonemic quantity distinction extends beyond the nuclei and affects the acoustic durations of the surrounding sounds. Figure 3(left) shows that the consonant preceding the nucleus was longer in normal rate than in fast rate ($F = 416.6$), but also longer when long nuclei followed than when short ones followed ($F = 45.4$). In addition, the initial onset consonants were also slightly longer when followed by consonantal nuclei than by vocalic ones ($F = 27.7$). A linear mixed-model test showed no significant interactions among these three main factors.

A similar test with the duration of the consonant following the nucleus as the dependent factor [Fig. 3(right)] showed even greater main effects for SpeechRate ($F = 685.5$), and NucleusType ($F = 62.7$), and a weakly significant interactions between Quantity and SpeechRate ($F = 10.1$). Subsequent separate tests for normal and fast rates showed that consonants following long nuclei were longer than the ones following short nuclei only in the fast rate ($F = 16.9$). Finally, the duration of the final vowels in the target words was affected only by SpeechRate ($F = 120.2$).

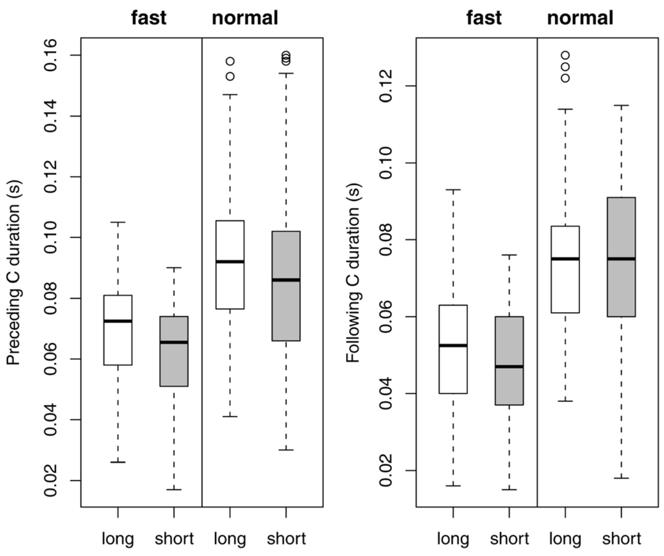

FIG. 3. Effects of phonemic quantity and speech rate on the acoustic closure duration of (left) the preceding and (right) the following consonants.

The cumulative effect of phonemic quantity and speech rate on the nucleus and the C_1 onset resulted in a better separation of the four categories (long-normal, long-fast, short-normal, short-fast) in the duration of the first syllable than in the duration of the nucleus or onset alone. Recall the weak effect of SpeechRate in the short nuclei reported above ($F = 10.3$). The same test with the duration of the whole first syllable as the dependent measure showed increased robustness of this effect ($F = 180.1$). This observation supports the suprasegmental nature of nucleus quantity and is in line with models expressing quantity as weighted ratios of durations of multiple participating units (e.g., Port, 1981).

B. Relationship between the acoustic and kinematic characteristics of nucleus duration

We now test if the lip opening and closing movements surrounding the nuclei represent a relevant control mechanism for the variability in the acoustic duration of the nuclei. Hertrich and Ackermann (1997) argued that significant correlations between the acoustic duration of nuclei and the opening-closing cycle of the lip movement surrounding them (Pearson's coefficients ≥ 0.8 for /a/ and ≥ 0.5 for /i/ and /u/) show that the phonemic quantity contrasts in German nuclei is controlled by the articulatory features of these lip movements. Using similar methodology for the Slovak data, we found highly significant correlation between *durArt*, and *durAcc* for all seven nuclei (Pearson's coefficients ≥ 0.9 in normal rate, and between 0.57 for /u/ and 0.95 for /a/ in fast rate). Both consonantal nuclei behaved comparably to the vocalic nuclei. To avoid making part-whole correlations (e.g., Barry, 1983; Benoit 1986), which in themselves would result in a correlation of about 0.7, we have also calculated *durAccPerc* as the percentage of the shorter interval (the part, in our case *durAcc*) of the longer interval (the whole, *durArt*) and tested the correlations between *durArt*, and *durAccPerc*. Pearson's coefficients for all seven nuclei remained significantly positive (0.3 for /u/, 0.63 for /o/ and $\geq$0.7 for the other five nuclei). Hence, the articulation of the opening and closing consonantal movements surrounding syllable nuclei provides a relevant control mechanism for durational variation in Slovak nuclei.

Investigating the relationship between the acoustic duration of nuclei (*durAcc*) and the durations determined from the nucleus gestures of the tongue (*durPlateau* and *durGesture*), and using the same strategy for avoiding the part-whole issue as described above, two general patterns emerged. First, the duration of the entire nucleus gesture also provides a relatively robust relationship to the acoustic duration for all nuclei; highly significant positive Pearson's coefficients between 0.54 and 0.82 were found for all seven nuclei. Second, the acoustic duration of the nuclei positively correlates only with the plateaus of the tongue tip gestures of the two liquids ($r(100) = 0.23$, $p = 0.023$ for /l/, $r(98) = 0.46$,

$p < 0.001$ for /r/), while the plateaus of the vocalic gestures show either negative or insignificant correlations.

C. Intra-gestural characteristics of the lip movements preceding and following the nucleus

Following Hertrich and Ackerman (1997) and Shaiman (2001), we expected that the lip movements associated with long nuclei and in normal speech rate should be longer, slower, and with greater displacement in comparison to the movements associated with short nuclei and those in fast speech rate. The predictions for the difference of such scaling between the opening and closing movements were less clear. The results from investigating acoustic durations of the preceding and following consonants, which are reported in Sec. III A, indicate that the duration of the preceding consonant was affected by phonemic quantity of the nucleus more than the following consonant, and that the opposite applied to the effect of speech rate. Given that longer contact duration for consonants is typically associated with less stiff and slower movements (e.g., Löfqvist, 2005), we expected a greater effect of the quantity contrast on the kinematics of the opening gesture of the onset consonant (henceforth cv) than on the closing gesture of the consonant following the nucleus (henceforth vc). However, data from the short/long contrast of German vowels did not show this general asymmetry in the effect of vowel quantity on the durations of opening and closing movements (Hertrich and Ackermann, 1997). Finally, the effect of nucleus type (consonantal vs vocalic) on the kinematics of the opening and closing movements has not been previously examined to our best knowledge. To address these issues, we tested the effects of QUANTITY (long/short), SPEECHRATE (normal/fast), NUCLEUSTYPE (consonant/vowel), and GESTURETYPE (cv-opening/vc-closing) on raw kinematic parameters of movement *duration, peak velocity, time to peak velocity*, and *displacement*, followed by testing their effects on derived measures of *stiff-ratio*, and *parameter C*.

Both QUANTITY and SPEECHRATE had robust effects on the four raw measures in the expected direction. The lip movements associated with long nuclei were longer (F = 421.6), had lower (F = 186.5) and later (F = 207.7) peak velocity, and greater displacement (F = 444.6) in comparison to the movement associated with short nuclei. Similarly, the movements produced in normal speech rate were longer (F = 217.7), had lower (F = 33.1) and later (F = 77.7) peak velocity, and greater displacement (F = 83.9) in comparison to the movements produced in fast speech rate. Additionally, two interactions between QUANTITY and SPEECHRATE were significant. The phonemic quantity contrast affected movement duration slightly more in normal than in fast rate (F = 11.1). In peak velocity, the significant effect of SPEECHRATE applied only to the short nuclei (F = 90.8) and not to the long ones (F = 0.3). In general, the comparison of the reported F-values shows that the effect of phonemic quantity was more robust on all measures than the effect of speech rate.

A significant difference between the kinematics of cv-opening and vc-closing gestures was observed on three out of four measures. The boxplots in Fig. 4 illustrate these effects. The cv-gesture was longer (F = 164.1), had smaller peak velocity (F = 80.7), and longer time to peak velocity (F = 299.1) than the vc-gesture, while the displacements did not differ significantly (F = 0.0). Additionally, the interaction between QUANTITY and GESTURETYPE was significant in movement duration (F = 16.3), peak velocity (F = 33.7), and time to peak velocity (F = 25.3). The main effect of QUANTITY held for both cv and vc gestures, only the difference between long and short vowels was greater in the duration of

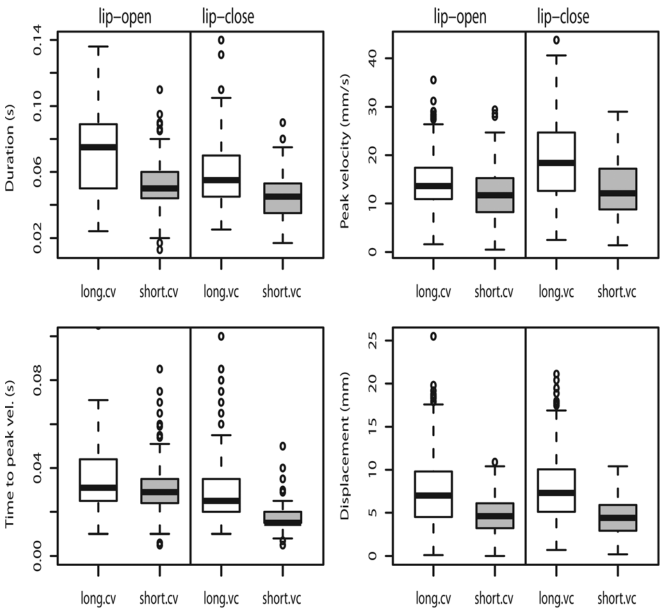

FIG. 4. Effect of phonemic quantity (short vs long) and gesture type (lip-opening vs lip-closing) on (top left) movement duration, (top right) peak velocity, (bottom left) time-to-peak-velocity, and (bottom right) displacement.

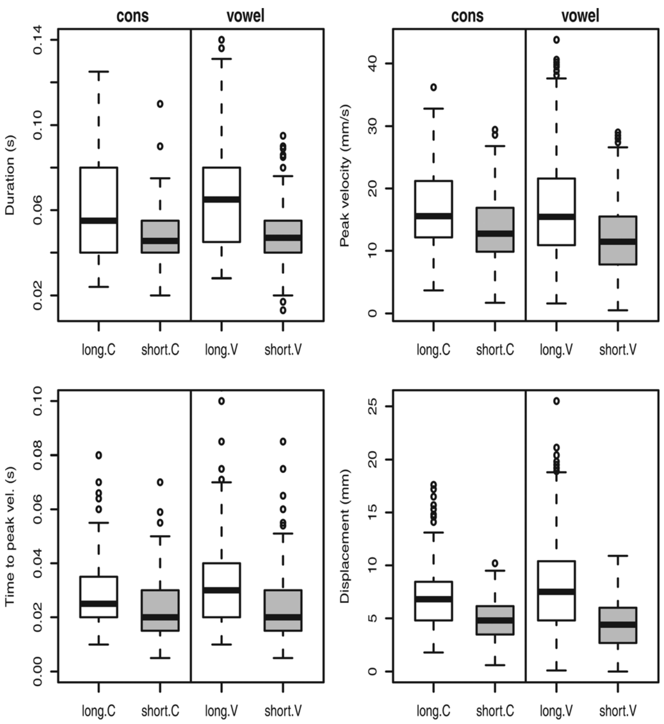

FIG. 5. Effect of phonemic quantity (short/long) and nucleus type (consonantal/vocalic) on (top left) movement duration, (top right) peak velocity, (bottom left) time-to-peak-velocity, and (bottom right) displacement.

the cv-gesture than in the vc-gesture, and greater in the peak velocity and time to peak velocity in the vc-gesture than in the cv-gesture. Hence, the effect of phonemic quantity on the two lip movements was asymmetrical for three of the four measures, and in two of these the vc-gestures were affected more than cv-gestures.

The boxplots in Fig. 5 illustrate the effect of NucleusType on the kinematics of the lip movements. In general, this factor had the weakest effect of the four main factors. Movement duration (F = 9.8) and time to peak velocity (F = 9.6) were slightly greater in the words with vocalic nuclei than in the words with consonantal ones. We also found significant interactions of NucleusType with Quantity in movement duration (F = 12.6) and displacement (F = 15.4). Further inspection of these two interactions revealed that the difference between the consonantal and vocalic nuclei in movement duration was only significant in long nuclei (F = 16.1).

Additionally, long nuclei caused marginally greater displacement of the lip movements in vocalic nuclei than consonantal ones (F = 9.5), while a reverse tendency was observed in short nuclei (F = 8.5), and greater variability, especially in long nuclei, can be observed for vocalic than for consonantal nuclei. These results are in line with the suggestions that jaw movements are greater for vowels than for consonants (e.g., Lindblom, 1983). If we assume that the jaw and the lips belong to a single synergy for the achievement of the lip closure and abstract away from known differences and compensations of the upper lip, lower lip, and jaw in achieving this goal (Gracco, 1994), greater jaw movements for vowels (especially for the long ones) than the consonants may explain the observed effects of NucleusType in the movement duration and displacement.

Finally, we examined the influence of phonemic duration and speech rate on the lip movements with the measures of stiffness (*stiff-ratio, parameter C*). Lip movements for long nuclei had lower stiff-ratio than the movements for short ones (F = 149.6). Lip movements in the normal rate had lower stiff-ratio than in the fast rate (F = 104.9) and the lip opening movements had lower stiff-ratio than the closing ones (F = 46.9). Neither NucleusType, nor any interaction of factors had a significant effect. Hence, even if stiffness was assumed to be a major control parameter for speech rate, the effect of phonemic quantity was slightly more robust than the effect of speech rate. *Parameter C* was expected to be affected by SpeechRate in that normal rate should produce larger values than fast rate (Adams *et al*., 1993). However, it was not affected significantly by any of the four factors and no significant interactions emerged. One of the goals of this paper was to investigate the differences between consonantal and vocalic nuclei. In this sense it is noteworthy that NucleusType had no significant effect on these two derived kinematic measures of the lip movement.

D. Intra-gestural characteristics of the nucleus gesture

One of the systematic patterns reported in the previous section was a more robust effect of phonemic quantity than speech rate on the lip movements surrounding the nucleus, and little or no difference in this effect depending on whether the nucleus was a vowel or a liquid consonant. In this section we investigate how the two main factors (Quantity and SpeechRate) affect the kinematics of the nucleus gesture itself and if there are systematic differences in this effect between the consonantal and vocalic nuclei.

Phonemic quantity affected significantly both plateau duration (*durPlateau,* F = 204.9) and gesture duration (*durGesture,* F = 247.2). Speech rate affected both measures significantly, but less robustly than quantity (F = 28.4 and F = 138.8, respectively). Additionally, consonantal nuclei had a shorter overall duration of gestures than vocalic ones (F = 143.6), this effect was slightly more robust in normal rate than in the fast rate (F = 16.9), and no significant interaction between NucleusType and Quantity was reported. The effect of NucleusType on plateau duration was different. It had no significant main effect, but interacted significantly with Quantity. Short liquids had shorter plateaus than short vowels (F = 19.3), but long liquids had longer plateaus than long vowels (F = 13.4). This also means that Quantity affected plateau duration of consonantal nuclei more robustly than it affected the plateau duration of vowels, which is in line with the results concerning *durPlateau* presented in Sec. III B.

Examining now only the movement toward the target of the nucleus gestures, the consonantal nuclei had robustly greater peak velocities than the vocalic ones (F = 444.8), and less robustly shorter times to peak velocity (F = 26.3). We observed neither the main effect of Quantity nor its interaction with other factors for peak velocities and time to peak velocity. Displacements for long nuclei were greater than for short ones (F = 21.8). SpeechRate had only a marginal effect on time to peak velocity (F = 10.3): shorter times in fast rate.

For the derived measures, the strongest factor was predictably NucleusType. Consonantal nuclei had much greater stiffness (*stiff-ratio*, F = 745.3) and later relative velocity peaks (*symmetry,* F = 232.2) than the vocalic ones. The effects of Quantity and SpeechRate on the derived measures were much less robust than the effect of NucleusType. In comparison to short nuclei, long nuclei had lower *stiff-ratio* (F = 24.7), marginally earlier velocity peaks (F = 8.8), and greater values of *parameter C* in the vocalic nuclei (F = 17.5). Nuclei in normal rate had lower *stiff-ratio* (F = 24.9) and greater values of *parameter C* (F = 13.8) than nuclei in fast rate.

E. Inter-gestural coordination characteristics

We assessed the coordination of the two lip movements with the measures of *symmetry* of velocity profiles and *peak-to-peak* ratio (Sec. II E). The results for the first measure are illustrated in Fig. 6(left) and show the truncation of lip movements surrounding short nuclei compared to the long ones. This is because lip movements surrounding short nuclei had late relative velocity peaks of the opening gesture and early ones in the closing gesture. Moreover, this effect was more robust in the normal rate than in the fast rate. A linear mixed model test showed the main effect of GestureType (F = 164.7), its interaction with Quantity (F = 77.4), as

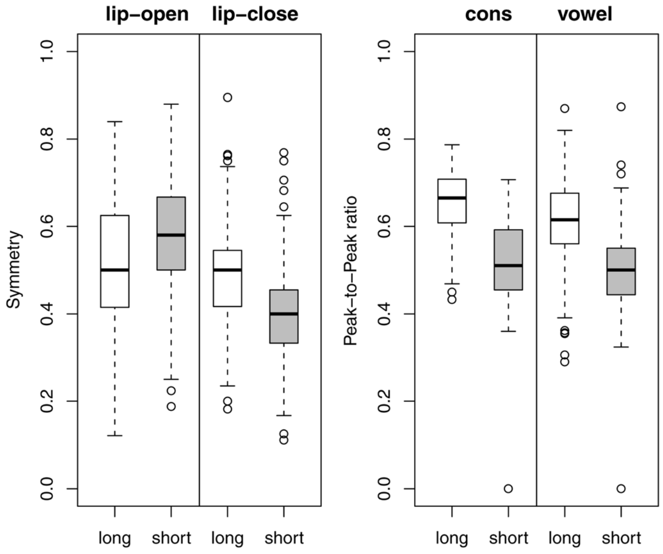

FIG. 6. (Left) Effect of phonemic quantity (short/long) and gesture type (lip-opening/lip-closing) on the *symmetry* of velocity profiles. (Right) Effect quantity and nucleus type (consonantal (cons) vs vocalic (vowel)) on *peak-to-peak ratio*.

well as a weak 3-way interaction GestureType-Quantity-SpeechRate (F = 9.4).

The results for *peak-to-peak ratio* are illustrated in Fig. 6(right) and show that this ratio was robustly smaller for lip movements surrounding short nuclei than the long ones (F = 287.1). SpeechRate showed no significant main effect and only a marginal interaction with Quantity (F = 10.5) so that the effect of phonemic quantity was slightly greater in normal rate than in fast rate. This finding further corroborates the pattern emerging from the data that the temporal coordination of the lip movements is employed for controlling the phonemic quantity distinction to a greater extent than for implementing variations in speech rate. Finally, NucleusType affected *peak-to-peak ratio* significantly (F = 15.3) in that consonantal nuclei showed less truncation of the lip movements, and thus less overlap, than vocalic ones.

Next we tested the coordination between the gestures of the initial consonant and the nucleus using *pN-lag* (Sec. II E). Greater values of this measure represent smaller overlap of the gestures. Consonantal nuclei were robustly less overlapped with the onset gesture than the vocalic nuclei (F = 394.2). There was a weaker effect of SpeechRate (F = 45.1), and a significant interaction between SpeechRate and NucleusType (F = 25.1). Additional tests examining the two nuclei separately showed that SpeechRate had a significant effect only on the consonantal nuclei (F = 64.6) and only a weak tendency on the vocalic nuclei was reported (F = 7.5). Consonantal nuclei overlapped the onset C_1 gesture less in the normal rate than in the fast rate. There was no main effect of Quantity and no significant interaction of Quantity and NucleusType.

Our data also corroborate previous findings from the coordination of vocalic and consonantal gestures in VCV sequences (e.g., Löfqvist and Gracco, 1999) that the tongue body movement for the vowel starts consistently before the closure for the consonant is achieved. In our data this lag was on average 5 ms.

Finally, we tested how phonemic quantity and speech rate control the coordination of the nucleus gesture with respect to both flanking labial gestures. We calculated the lag between *C1C2Mid* and three landmarks of the nucleus gesture (N-NONS, N-MAXC, N-NOFF) defined as *N-NONS-lag, N-MAXC-lag, N-NOFF-lag,* respectively, in Sec. II E. We found that neither Quantity nor SpeechRate significantly affected the interval *N-NONS-lag* (F = 1.7, and F = 1.9, respectively), they both affected significantly *N-MAXC-lag* (F = 61.2, and F = 29.6, respectively), and they affected *N-NOFF-lag* the most robustly (F = 211.2, and F = 47.3, respectively). In all significant effects, the long nuclei and the nuclei in normal rate were shifted "to the right" with respect to *C1C2Mid* point more than the short nuclei or nuclei in fast rate. This effect is sketched in Fig. 7. The achievement of the target for the nucleus gesture (N-NONS) has a stable timing with the *C1C2Mid* point for both long and short nuclei. However, both N-MAXC and N-NOFF landmarks show a greater lag to the *C1C2Mid* for the long vowels than the short ones. Finally, NucleusType had no significant effect on the three measures, only an interaction with Quantity was reported with *N-NOFF-lag* (F = 15.9) so that the contrast due to phonemic quantity was greater in consonantal nuclei than in the vocalic ones.

IV. DISCUSSION

We investigated the effects of the phonemic quantity of syllabic nuclei and speech rate variations on the realization of C_1NC_2a sequences where C_1 and C_2 were both /p/, and N represented all 14 possible syllabic nuclei in Slovak: ten vowels {[i], [e], [a], [o], [u], [iː], [eː], [aː], [oː], [uː]} and four liquids {[r], [l], [rː], [lː]}. We concentrated on intra- and

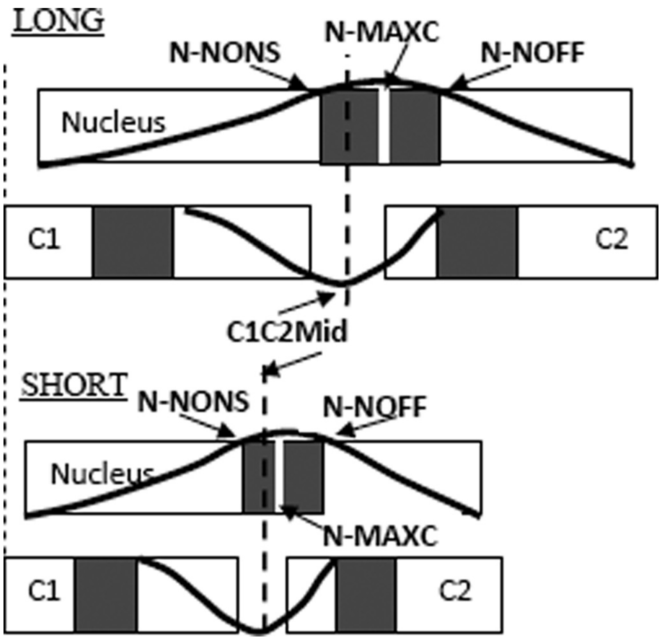

FIG. 7. Schematic illustration of the relationship between *C1C2Mid* point and the landmarks of the nucleus gesture for (top) long nuclei and (bottom) short ones.

inter-gestural characteristics of three core movements: lip opening of C_1, tongue movement for N, and lip closing for C_2. Several systematic patterns emerged from the data.

First, although both phonemic quantity and speech rate variation caused adjustments in both inter- and intra-gestural characteristics, differences in the underlying durational control of phonemic quantity and speech rate were observed in the distribution of these adjustments. Changes in the inter-gestural coordination were a more robust determinant of phonemic quantity differences while changes in the intra-gestural characteristics were more prominent as speech rate control. While both opening and closing lip movements had roughly symmetric profiles for long nuclei, short nuclei displayed significantly later peaks in the opening movement and significantly earlier peaks in the closing movement. Short nuclei also had a robustly smaller peak-to-peak ratio (time between the velocity peaks of the lip opening and closing gestures relative to the interval between lip-closure offset and subsequent lip-closure onset). These findings support the analysis that the lip movements are mutually more overlapped, i.e., temporarily more coproduced, in short nuclei than the long ones. Speech rate affected both measures of the coproduction only very marginally in that the effect of phonemic quantity was slightly more robust in the normal rate than in the fast rate.

Additional support for the crucial role of the temporal coordination of lip movements in cueing phonemic quantity comes from the asymmetrical effect of quantity on the coordination of the nucleus tongue gesture to the two lip movements. The C_1N coordination was not influenced significantly by phonemic quantity. However, the lag of N-MAXC and N-NOFF nucleus landmarks to the *C1C2Mid* point was greater for long nuclei than for short ones (see Fig. 7). At the same time, the adjustments in stiffness of the nucleus closing gesture represented the only robust effect of phonemic quantity on this gesture. These findings suggest that the temporal extension of the nucleus gesture by stiffness adjustment exceeds the temporal window provided by C_1-C_2 coordination control. The effect of speech rate showed the opposite trend on the measures of inter-gestural coproduction. Speech rate significantly influenced C_1N coordination (albeit only in consonantal nuclei), and its effect on the lag between *C1C2Mid* and nucleus landmarks was significant but much less robust than the effect of phonemic quantity.

Finally, the measures of the coproduction of the two labial movements showed absent or minimal influence of the nucleus type. Hence, this coproduction relationship was stable whether the nucleus was a liquid or a vowel. Despite great differences between these two nucleus types in their kinematics (we reported robust differences in peak velocity, stiffness, or the symmetry of velocity profiles) and their coordination with onsets (robustly greater lag of the consonantal nuclei to the C_1 gesture than the vocalic ones), the temporal patterning of the flanking consonantal movements was influenced minimally by nucleus type. Hence, although consonantal and vocalic nuclei are predictably different on many kinematic measures, one way in which their common phonological behavior as syllabic nuclei may be signaled, is through a stable C_1-C_2 coordination. In this sense, our data fit well with a general approach to understanding a syllable as a set of timing requirements for the gestures forming the syllable domain (Krakow, 1999).

The crucial role of the temporal coordination of the flanking consonantal gestures in cuing phonemic duration in Slovak also suggests that this coordination might be independent of the phonotactic patterns. In German, phonemic length of the nucleus is linked to the presence of a coda consonant, but in Slovak long and short nuclei appear without such restrictions. Since both languages realize phonemic quantity contrast on nuclei primarily through the temporal coordination of the flanking consonantal gestures, this relationship between linguistic function and articulatory control is, barring future findings from other languages suggesting otherwise, general and language independent.

The second stable pattern that emerged from the data also concerns the relationship between phonemic quantity and speech rate. In several aspects, these two functional requirements predictably behave cumulatively in the sense that lengthening caused by both phonemic quantity and normal rate is greater than lengthening due to each requirement individually. A very clear example of such pattern in our data was the acoustic duration of the first syllable. However, we also observed instances of complementary rather than cumulative behavior of the two requirements. For example, the effect of phonemic quantity was greatest on the acoustic duration of nuclei, medial on the duration of preceding consonant belonging to the same syllable, and weakest on the onset consonant of the following syllable. Speech rate, on the other hand, affected the acoustic duration of C_2 the most, then C_1 and then the nucleus. Similarly, phonemic quantity affected the kinematics of the lip movements more than speech rate did, yet the onset-nucleus coordination, in which quantity had no significant effect, showed the effect of speech rate for the consonantal nuclei. In general, then, those measures that were affected by phonemic duration robustly tended to show a minimal effect of speech rate, while measures greatly affected by speech rate tended to be influenced by phonemic quantity minimally.

As concerns the kinematics of the nuclei, we found robustly greater stiffness and later relative velocity peaks for the consonants compared to vowels. The closing gesture of the long nuclei had lower stiffness than the short ones despite the absence of the quantity effect on peak velocities. These observations are similar to the results of Löfqvist (2005) for the closing movements of Japanese geminates/singletons consonants in that the long segments had more displacement but not greater velocity than the short segments. Moreover, we observed that the gestures for long nuclei had marginally earlier velocity peaks for the long nuclei than the short ones despite the fact that long nuclei had clearly longer plateaus and overall durations. A shorter relative acceleration phase for long nuclei compared to the short ones means, in turn, a longer deceleration phase for the long nuclei. The adjustments of the deceleration phase observed in our data supports analysis division of a general slow-down of the movement as one of the controls for longer durations (Edward *et al.*, 1991; Hertrich and Ackermann, 1997) since this observation cannot be analyzed as truncation of one gesture by another.

Parameter C of the opening and closing lip movements was not affected significantly by any of the four factors and no significant interactions emerged. This result adds to the finding of Hertrich and Ackermann (1997) that this parameter in general does not systematically control the lip movements for phonemic quantity distinction. However, the absence of the effect of speech rate is surprising given that Adams *et al*. (1993) found systematically greater values of this parameter the slower the speech rate. Also, Hertrich and Ackermann (1997) found greater values of this parameter for longer movements. Our results are more consistent with Munhall *et al*. (1985) that, similarly to the present study, varied speech rate only between the normal and fast rates. However, *parameter C* was significantly, albeit weakly, affected by speech rate in the movements toward the target of the syllabic nucleus in the expected direction: nuclei in the fast rate had lower values than nuclei in the normal rate. This might be due to the fact that the gestures for the nucleus are significantly longer than the labial movements. Overall, it seems that systematic effects of *parameter C* might require a greater range of movement durations.

Finally, our data support the general approach to phonemic length of syllabic nuclei as a suprasegmental phenomenon (e.g., Vennemann, 1991). For example, we observed cumulative effects of the acoustic duration of nuclei and preceding onsets, and the importance of the inter-gestural timing and intra-gestural characteristics of the consonantal gestures surrounding the nucleus. In search for a model capable of generating the observed patterns in this study, the π-gesture model should be considered (Byrd and Saltzmann, 2003; Saltzmann *et al*., 2008). This model formalizes the observation that stiffness scaling (determined from observed kinematic data) and the degree of gestural coproduction function inter-dependently in controlling quantity requirements related to the prosodic structure. The π-gestures are assumed to represent local adjustments to the speed of the central clock controlling the activations of gestures falling within the domain of π-gestures. Both modeling and empirical studies showed that the π-gesture model successfully generates several observed patterns (e.g., Byrd *et al*., 2000, 2006).

In our data, the long tokens were realized with lower stiffness consistently on all three target movements, and with decreased overlap of the C_1 opening and C_2 closing movements. Hence, the observed patterns invite the analysis that the phonemic duration contrast of syllabic nuclei might be controlled similarly to the durational effects arising from the presence of prosodic boundaries. Future modeling work should, however, also examine two observations in our data and test if the parameters of the current model are sufficient in generating the observed patterns. Both of these observations relate to the closing movements for the nucleus gesture. First, the effect of phonemic duration on the stiffness of this gesture was much less robust than its effect on the stiffness of the lip movements associated with syllable margins. If a π-gesture is deployed to signal phonemic quantity contrast and is aligned to the syllable nucleus, we would expect the greatest effect on the nucleus gesture and smaller effects on the lip movements. Second, as we discussed above, while the gestures of the long nuclei were more expanded than those for the short ones, the C_1-C_2 timing, to some extent, acted against this expansion of the nucleus gesture. The question is if the inter-dependence of intra- and inter-gestural adjustments stemming from their source in the activation level and central clock could also accommodate, perhaps as a combination of alignment and shape parameters of a π-gesture, a seemingly independent behavior of these adjustments.

ACKNOWLEDGMENT

This work was supported by an Alexander von Humboldt Fellowship. The author also wishes to acknowledge support from EUSF project IMTS 26240220060 and to thank Phil Hoole, Marianne Pouplier, and two anonymous reviewers for valuable comments to earlier drafts of this paper and Susanne Waltl and Yuki Era for assistance with data collection and annotation. The author retains the responsibility for any errors.

Adams, S. G., Weismer, G., and Kent, R. D. (**1993**). "Speaking rate and speech movement velocity profiles," J. Speech Hear. Res. **36**, 41–54.

Baayen, R. H. (**2008**). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R* (Cambridge University Press, Cambridge), pp. 241–302.

Barry, W. (**1983**). "Some problems of interarticulator phasing as an index of temporal regularity in speech," J. Exp. Psychol.: Hum. Percept. Perform. **16**, 826–828.

Benoit, C. (**1986**). "A note on the use of correlation in speech timing," J. Acoust. Soc. Am. **80**, 1846–1849.

Beňuš, Š., and Mády, K. (**2010**). "Effects of lexical stress and speech rate on the quantity and quality of Slovak vowels," in *Speech Prosody 2010*, Vol. 100185, pp. 1–4. (Available online at http://aune.lpl.univ-aix.fr/~sprosig/sp2010/papers/100185.pdf, last viewed 2/28/11.)

Browman, C. P., and Goldstein, L. (**1990**). "Tiers in articulatory phonology, with some implications for casual speech," in *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*, edited by J. Kingston and M. E. Beckman (Cambridge University Press, Cambridge), pp. 341–338.

Byrd, D., Kaun, A., Narayanan, S., and Saltzman, E. (**2000**). "Phrasal signatures in articulation," in *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, edited by M. B. Broe and J. B. Pierrehumbert (Cambridge University Press, Cambridge), pp. 70–87.

Byrd, D., Krivokapić, J., and Lee, S. (**2006**). "How far, how long: On the temporal scope of phrase boundary effects," J. Acoust. Soc. Am. **120**, 1589–1599.

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

Edwards, J., Beckman, M. E., and Fletcher, J. (**1991**). "The articulatory kinematics of final lengthening," J. Acoust. Soc. Am. **89**, 369–382.

Gracco, V. (**1994**). "Some organizational characteristics of speech movement control," J. Speech Hear. Res. **37**, 4–27.

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

Hertrich, I., and Ackermann, H. (**1997**). "Articulatory control of phonological vowel length contrasts: Kinematic analysis of labial gestures," J. Acoust. Soc. Am. **102**, 523–536.

Hoole, P., and Mooshamer, C. (**2002**). "Articulatory analysis of the German vowel system," in *Silbenschnitt und Tonakzente*, edited by P. Auer, P. Gilles, and H. Spiekermann (Niemeyer, Tübingen), pp. 129–152.

Hoole, P., and Zierdt, A. (**2010**). "Five-dimensional articulography," in *Speech Motor Control*, edited by B. Maasen and P. H. H. M. van Liehout (Oxford University Press, Oxford), pp. 331–349.

Klatt, D. H. (**1973**). "Interaction between two factors that influence vowel duration," J. Acoust. Soc. Am. **54**, 1102–1104.

Krakow, R. A. (**1999**). "Physiological organization of syllables: A review," J. Phonetics **27**, 23–54.

Lindblom, B. (**1983**). "Economy of speech gestures," in *The Production of Speech*, edited by P. F. MacNeilage (Springer, New York), pp. 217–246.

Löfqvist, A. (**2005**). “Lip kinematics in long and short stop and fricative consonants,” J. Acoust. Soc. Am. **117**(2), 858–878.

Löfqvist, A. (**2007**). “Tongue movement kinematics in long and short Japanese consonants,” J. Acoust. Soc. Am. **122**(1), 512–518.

Löfqvist, A., and Gracco, V. L. (**1999**). “Interarticulator programming in VCV sequences: Lip and tongue movements,” J. Acoust. Soc. Am. **105**(3), 1864–1876.

Munhall, K. G., and Lofqvist, A. (**1992**). “Gestural aggregation in speech: Laryngeal gestures,” J. Phonetics **20**, 93–110.

Munhall, K. G., Ostry, D. J., and Parush, A. (**1985**). “Characteristics of velocity profiles of speech movements,” J. Exp. Psychol. Hum. Percept. Perform. **11**, 457–474.

Ostry, D. J., and Munhall, K. G. (**1985**). “Control of rate and duration of speech movements,” J. Acoust. Soc. Am. **77**, 640–648.

Port, R. F. (**1981**). “Linguistic timing factors in combination,” J. Acoust. Soc. Am. **69**(1), 262–274.

Port, R. F., Cummins, F., and McAuley, J. D. (**1995**). “Naive time, temporal patterns and human audition,” in *Mind as Motion*, edited by R. F. Port and T. van Gelder (MIT Press, Cambridge, MA), pp. 339–437.

Pouplier, M., and Beňuš, Š. (**2011**). “On the phonetic status of syllabic consonants: Evidence from Slovak,” Laboratory Phonology **2**(2), doi:10.1515/LABPHON.2011.009.

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

Rubach, J. (**1993**). *The Lexical Phonology of Slovak* (Clarendon Press, Oxford), pp. 1–316.

Saltzman, E., and Kelso, S. (**1987**). “Skilled actions: A task-dynamic approach,” Psychol. Rev. **94**(1), 84–106.

Saltzman, E, Nam, H., Krivokapic, J., and Goldstein, L. (**2008**). “A task-dynamic toolkit for modeling the effects of prosodic structure on articulation,” in *Proceedings of Speech Prosody 2008*, edited by P. A. Barbosa, S. Madureira, and C. Reis, Campinas, Brazil, pp. 175–84. (Available online at http://aune.lpl.univ-aix.fr/~sprosig/sp2008/papers/3inv.pdf, last viewed 2/28/11.)

Simko, J., and Cummins, F. (**2010**). “Embodied task dynamics,” Psychol. Rev. **117**(4), 1229–1246.

Shaiman, S. (**2001**). “Kinematics of compensatory vowel shortening: The effect of speaking rate and coda composition on intra- and inter-articulator timing,” J. Phonetics **29**, 89–107.

Shaiman, S. (**2002**). “Articulatory control of vowel length for contiguous jaw cycles: The effects of speaking rate and phonetic context,” J. Speech Hear. Res. **45**, 663–675.

Smith, C. L. (**1995**). “Prosodic patterns in the coordination of vowel and consonant gestures,” in *Laboratory Phonology IV: Phonology and Phonetic Evidence*, edited by B. Connell and C. Arvaniti (Cambridge University Press, Cambridge), pp. 205–222.

Vennemann, T. (**1991**). “Syllable structure and syllable cut prosodies in Modern Standard German,” in *Certamen Phonologicum II. Papers from the 1990 Cortona Phonology Meeting*, edited by P. M. Bertinetto (Rosenberg and Sellier, Torino), pp. 211–243.

Zeroual, C., Hoole P., and Gafos, A. (**2008**). “Spatio-temporal and kinematic study of moroccan arabic coronal geminate,” in *Proceedings of 8th ISSP*, edited by R. Sock, S. Fuchs and Y. Laprie (Inria, Paris), pp. 133–136.