Word-internal versus word-peripheral consonantal duration patterns in three languages

Melissa A. Redford[a)]
Department of Linguistics, University of Oregon, Eugene, Oregon 97403

(Received 12 April 2006; revised 8 December 2006; accepted 11 December 2006)

Segmental duration patterns have long been used to support the proposal that syllables are basic speech planning units, but production experiments almost always confound syllable and word boundaries. The current study tried to remedy this problem by comparing word-internal and word-peripheral consonantal duration patterns. Stress and sequencing were used to vary the nominal location of word-internal boundaries in American English productions of disyllabic nonsense words with medial consonant sequences. The word-internal patterns were compared to those that occurred at the edges of words, where boundary location was held constant and only stress and sequence order were varied. The English patterns were then compared to patterns from Russian and Finnish. All three languages showed similar effects of stress and sequencing on consonantal duration, but an independent effect of syllable position was observed only in English and only at a word boundary. English also showed stronger effects of stress and sequencing across a word boundary than within a word. Finnish showed the opposite pattern, whereas Russian showed little difference between word-internal and word-peripheral patterns. Overall, the results suggest that the suprasegmental units of motor planning are language-specific and that the word may be more a relevant planning unit in English. © 2007 *Acoustical Society of America.* [DOI: 10.1121/1.2431339]

PACS number(s): 43.70.Bk, 43.70.Fq, 43.70.Kv [AL] Pages: 1665–1678

I. INTRODUCTION

Since Kozhevnikov and Chistovich (1965) and Lehiste (1970) researchers have used segmental duration patterns to gain insight into the representation of sequential speech action. Decades of work have provided a wealth of information on the intrinsic and sequential factors influencing articulation as well as on the relationship between linguistic units and speech planning. A widely accepted conclusion that has emerged from this work is that articulatory timing is determined in part by the dynamics of sequential segment articulation and in part by suprasegmental units. Research continues in order to ascertain the nature and representation of the suprasegmental units. The current study fits within this tradition.

Specifically, the current study focuses on consonantal duration patterns that are attributed to the syllable, which has been repeatedly proposed as the most basic suprasegmental unit of articulatory organization in the speech plan. A principal problem with this proposal is that the syllable is ill-defined in that its boundaries are often ambiguous. Such ambiguity makes it difficult to test whether or not the syllable is a relevant unit of speech motor control. The current study attempts such a test, but also explores an alternative hypothesis; namely, that syllable-related segmental duration patterns emerge from the articulatory dynamics of producing meaningful bits of speech—lexical items, for example—rather than from a phonological plan for executing these items in prespecified chunks.[1] This possibility suggests that some putative syllable effects on segmental duration may be better explained by sequential and communicative factors rather than by the syllable per se. In this study, the effect of these other factors on consonantal duration patterns are explored and the results are discussed with reference to a model of speech production that is inspired by MacNeilage's (1998) frame/content theory.

A. Syllable-related segmental duration patterns

Production studies on the acoustic correlates of the syllable have found that vowel duration varies with syllable structure and consonant duration varies with syllable position. For instance, phrase-medial syllable-initial consonants are longer than syllable-final consonants (Boucher, 1988; Tuller and Kelso, 1991; Keating, Wright, and Zhang, 1999; Redford and Diehl, 1999) and the internal members of consonant clusters are shorter than the external member (Haggard, 1973; Klatt, 1976; Umeda, 1977). Perception studies indicate that these duration patterns are also used by listeners to infer syllable boundaries (Christie, 1977; Tuller and Kelso, 1991; Redford and Randall, 2005). For example, an intervocalic consonant-consonant sequence that is a possible onset cluster to the following vowel will be syllabified as an offset-onset sequence when produced with a short-long pattern, and as an onset cluster when produced with a long-short pattern (Redford and Randall, 2005). Although most of the work identifying consonantal duration patterns as correlates of syllable structure has been conducted in English, some work suggests that the relationship may be universal (e.g., Quené, 1992; Prieto, 2002; Maddieson, 1985, 2004).

The fact that segmental duration patterns covary with syllable structure and help cue syllable boundary perception has often been taken as evidence that syllables are supraseg-

[a)]Electronic mail: redford@uoregon.edu

mental units of articulatory organization. Specifically, segments are often thought to be temporally adjusted (or timed relative to one another) within a syllable frame (Kozhevnikov and Chistovich, 1965; Lehiste, 1970; Haggard, 1973; Klatt, 1976; Fowler, 1979; Boucher, 1988; Browman and Goldstein, 1988; Campbell and Isard, 1991; Campbell, 1992; Munhall *et al.*, 1992; Turk, 1994; Krakow, 1999; de Jong, 2003). Campbell (1992), for instance, suggested that segments are elastic in their durations and are stretched or compressed uniformly across the syllable, which is independently lengthened or shortened according to prosodic factors or rate of speech. van Santen and Shih (2000) questioned this version of the syllable frame hypothesis, noting that syllable duration increases with the number of segments. In addition, van Santen and Shih showed that intrinsic segmental duration creates systematic and large duration differences between identically structured syllables in the same prosodic context. Their finding suggests that segmental factors may sometimes provide a better explanation than syllables for local consonantal duration patterns.

Another version of the syllable frame hypothesis emphasizes syllabic structure over the temporal constancy of the syllable, but assumes a temporal frame nonetheless. In this more abstract version of the hypothesis, syllable structure complexity results in compensatory lengthening or shortening of segments in particular positions. For instance, Munhall *et al.* (1992) explained that vowel duration decreases within a syllable as the number of postvocalic consonants increases because the additional consonants drive increasing overlap between vocalic and consonantal articulation. The increased articulatory overlap between adjacent vowels and consonants only makes sense if an organizational domain is assumed: some boundary forces earlier and earlier articulation of postvocalic consonants as consonant number increases. Munhall *et al.* (1992) suggested that the relevant domain/boundary was syllabic. However, their experiment examined articulatory organization within monosyllabic words, so it remains unclear whether word-internal syllable boundaries could also force the same pattern of increasing overlap. In the absence of evidence for syllabic effects on word-internal articulatory timing, it unclear whether the relevant organizational domain is the syllable or the word.

In sum, segmental factors may often explain durational patterns attributed to the syllable, but higher-level units must be invoked to explain certain boundary-adjacent patterns. The identity of these higher-level units is still under dispute.

B. An alternative explanation for consonantal duration patterns

Munhall *et al.*'s (1992) study is not unique in confounding syllable and word boundaries. By far the most popular method of manipulating syllable structure in production experiments is to insert a word boundary where a syllable boundary is desired. This practice is justified by the fact that syllable boundaries are most easily defined with reference to a word boundary, but the problem is that the factors governing production at a word boundary are likely to be different than those governing production word-internally. For instance, we already know that articulatory strengthening and lengthening characterize segmental production at word and phrase boundaries (Fougeron and Keating, 1997; Byrd and Saltzman, 2003; Keating *et al.*, 2003; Tabain, 2003). Similar effects have not been observed word-internally (Keating *et al.*, 1999). Although this may be because word-internal patterns have not been as thoroughly explored as word-peripheral patterns, the absence of word-internal strengthening and lengthening could also indicate that the relevant linguistic units for speech motor planning are those that are also important to message encoding. After all, the motor goal of speech is to convey meaning and the location of a word-internal syllable boundary is not itself meaningful. For example, a word such as *city* can be produced by different speakers with different syllable structures (or, at least, it can be analyzed differently by different linguists) without affecting meaning [see, e.g., Kahn (1976) versus Selkirk (1982)]. Although the different pronunciations of the same word can be described as yielding different syllable structure percepts (cf. [cɪtʰi] versus [ciri]) this need not imply that the planning domain is different in the two cases. Instead, it may imply that different global production strategies have been employed to achieve different communicative goals (e.g., slow and careful versus fast and casual speech)—and these need not refer to syllable structure to produce effects that are attributed to syllable structure (e.g., aspirated /t/ versus the flapped allophone).

If sequential action is organized within word units rather than within syllable units, then the observed correlation between segmental duration patterns and word-edge structure must still be explained. One possibility is that the correlation evolved as a listener-oriented strategy to facilitate speech segmentation (Kohler, 1991; Quené, 1992; Fougeron and Keating, 1997; Keating *et al.*, 1999). If this is the case, then the research emphasis on English may have over-represented the strength of the correlation between segmental duration patterns and structure because English users may have a stronger need to cue boundary location than speakers of other languages: English has a large repertoire of sounds that can occur both word-finally and word-initially, and very little inflectional morphology to make boundaries more predictable. Further, the durational patterns at the edges of English words have been shown to help resolve word boundary ambiguity. For instance, the lengthening of /n/ relative to /s/ in the sentence *help us nail* distinguishes the *us nail* parse from an *a snail* parse (Christie, 1977). Intriguingly, these boundary-distinguishing patterns of relative consonantal duration emerge slowly in children's speech—at least a year after children are able to produce the relevant monosyllabic word shapes (i.e., syllable structures) in connected speech (Redford and Gildersleeve-Neumann, unpublished). This suggests that consonantal duration patterns around word boundaries may not be inherent to the production of different syllable structures, but learned as the child slowly gains control over timing one word relative to another. Again, the motivation for gaining such control may be to highlight meaningful boundaries for the listener. This possibility will be referred to here as the juncture hypothesis.

Of course, the juncture hypothesis begs the question of why particular durational patterns should be associated with particular word-onset and -offset structures. Whereas initial lengthening and strengthening may increase the perceptual prominence of a consonantal onset, thereby signaling the beginning of a word, there is no equivalent explanation for why the second consonant in a word-onset or -offset sequence would be shorter than the first. However, if we assume that the word is the most basic unit in the speech plan, then these patterns could simply represent within-unit coarticulation or articulatory overlap. That is, an account like the one offered by Munhall *et al.* (1992) for vowel shortening in monosyllabic words might be invoked to explain why internal members of a word-onset or -offset cluster are truncated.

This example suggests an important corollary to the juncture hypothesis; namely, that word-internal durational patterns may not refer to syllable structure, but may instead reflect the articulatory dynamics of sequential segment articulation. These dynamics are influenced by segment identity and sequencing as well as by the word's rhythmic structure. Specifically, the different articulatory configurations associated with different segments are achieved at different rates, even when other factors are held constant, resulting in different intrinsic segmental durations (Klatt, 1976; Umeda, 1975, 1977). When segmental context is varied, the same configurations will be achieved in more or less time; that is, coarticulation affects segmental duration (Umeda, 1975, 1977). Rhythm also affects articulatory dynamics. This assertion is especially well documented for the rhythmic structure derived from lexical stress (Edwards and Beckman, 1988; de Jong, Beckman, and Edwards, 1993; de Jong, 1995; Harrington, Fletcher, and Roberts, 1995). For example, de Jong characterized lexical stress as localized hyperarticulation: "more extreme targeting of articulator movement" (de Jong, 1995, p. 497). More extreme articulator movement takes longer to execute, resulting in longer segmental durations (e.g., Moon and Lindblom, 1994).

In sum, the word is proposed as an alternative unit to the syllable to explain boundary-related consonantal duration patterns in English. By extension of the hypothesis, word-internal patterns are assumed to emerge from within-unit articulatory dynamics.

C. The current study

The syllable frame and juncture hypotheses make different predictions regarding word-internal segmental duration patterns. The syllable frame hypothesis predicts that word-internal patterns will reflect syllable structure. The juncture hypothesis predicts that such patterns emerge from within-unit articulatory dynamics, which are affected by lexical factors such as stress and the identity and sequencing of the sounds within a word. The goal of the current study was to test between these predictions to determine whether the syllable frame or juncture hypothesis provides a better explanation for segmental duration patterns. To do this, we used disyllabic nonsense word stimuli with intervocalic consonant sequences and manipulated syllabic affiliation of the consonants using stress and sequence order. The resulting consonantal duration patterns were compared to those that occurred at the edges of words, where word-boundary location was held constant and stress and sequence order were varied.

The full manipulation was carried out in English. The results were then compared to those from Finnish and Russian, where the stress manipulation was across languages. The experiment was extended from English to Finnish and Russian to obtain more general conclusions. The extension was also motivated by the juncture hypothesis, which posits that boundary-adjacent consonantal patterns in English may be due to listener-oriented boundary highlighting. If it occurs, such highlighting could interfere with effects from stress and sequencing at the edges of words, making the word-internal results less interpretable as well (for lack of comparison). The assumption is that boundary highlighting is less crucial in Finnish and Russian. Both languages have extensive inflectional morphology, providing strong top-down cues to segmentation that do not exist in English. Also, word-edge phonotactics in Finnish are restrictive, providing an additional aid to segmentation.

II. COMPARISON OF WORD-INTERNAL AND WORD-PERIPHERAL CONSONANTAL DURATION PATTERNS IN ENGLISH

Syllable affiliation is most readily determined by position relative to the vowel at the edge of a word: a word-initial consonant is also syllable-initial, a word-final consonant is also syllable-final. Word-internal syllable affiliation is not as easy to determine for all sequence types. In particular, consonant sequences that can occur as word-onset clusters may be syllabified word-internally either as an offset-onset sequence or as an onset cluster to the following vowel. Psycholinguistic studies suggest that the particular syllabification will depend primarily on lexical stress since a stressed vowel attracts consonants (Treiman and Danis, 1988; Treiman and Zukowski, 1990; Derwing and Nearey, 1991; Redford and Randall, 2005). This means that one way to manipulate word-internal syllable boundaries is to manipulate stress, which is what was done in the present experiment. By itself, however, such a manipulation is inadequate for distinguishing between the syllable frame and juncture hypotheses: syllable structure changes with different stress patterns, but so do the local articulatory dynamics. One way to control for this problem is to vary stress while keeping the syllable boundary constant. This can be done at a word boundary. Accordingly, the experiment reported in this section compares consonantal duration patterns within a word (word-internal) and across a word boundary (word-peripheral) as a function of stress, which was manipulated both within the word and across the word boundary.

Another way to examine the independent effects of syllable structure and stress is to examine the effects of stress on word-internal consonant sequences that can only be syllabified as an offset-onset sequences. This can be done by inverting a sequence that forms a legal onset, thereby creating an illegal onset. Manipulating consonant sequences in this way has the added advantage of manipulating consonant position vis-à-vis a vowel and therefore syllable affiliation. For instance, when a stop-liquid sequence is inverted to become a liquid-stop sequence, the liquid becomes syllable-final instead of syllable-initial, and vice-versa for the stop depending on word position and stress.

In addition to changing syllable affiliation, sequencing changes also affect articulatory dynamics. For instance, pre- and postvocalic liquids are articulated differently (Sproat and Fujimura, 1993; Gick *et al.*, 2005). In a prevocalic environment a liquid is formed first with an anterior gesture then with a posterior gesture. The reverse is true in a postvocalic environment. Such changes in articulatory timing could result in position-dependent differences in the acoustic duration of liquids. A change in sequence order will also affect consonant-consonant coarticulation, which could affect acoustic duration. For instance, stop coarticulation with a subsequent liquid will result in greater oral constriction during the stop release than when the stop is coarticulated with a subsequent vowel. Greater oral constriction during the release gesture in a stop-sonorant sequence may increase "the time required to re-establish the transglottal pressure drop necessary for the initiation of voicing" compared with the less constricted gesture associated with the stop release in a stop-vowel sequence (Docherty, 1992; p. 148). This, in turn, means that stop aspiration duration is likely to be longer in a stop-liquid sequence than in a liquid-stop sequence when the stop precedes a vowel in the latter instance (viz. Docherty, 1992).

Sequencing effects arising from vowel-consonant or consonant-vowel coarticulation may be inseparable from effects of syllable position, but it is nonetheless important to examine these effects if only to explain syllable-related patterns in terms of more basic production processes. On the other hand, sequencing effects arising from consonant-consonant coarticulation may be minimized at a boundary that separates consonants, since it is assumed that boundaries reduce coarticulatory pressure. Although an effect of syllable boundary on consonant-consonant coarticulation would provide strong support for the syllable frame hypothesis and would allow us to reject the juncture hypothesis, the lack of an effect will not allow us to reject the syllable frame hypothesis. Some will argue that any absence of an effect of syllable structure on articulation is consistent with the syllable's position in the prosodic hierarchy: research on the articulatory dynamics at word versus higher-level phrase boundaries (accentual, intonational, and utterance) has shown that weaker prosodic boundaries have weaker effects on coarticulation than stronger prosodic boundaries [see, e.g., Tabain (2003)] and syllable boundaries are presumed to be weaker than word boundaries. If one accepts that the weakest prosodic boundaries may have no effect on coarticulation, then it would seem that only the juncture hypothesis can be falsified in the present experiment—a point that is relevant to interpreting the results that are presented after the experimental methods below.

A. Methods

1. Speakers and stimuli

Six native American-English speakers—three male and three female—produced a series of disyllabic nonsense words and nonsense word pairs (see Table I). The words contained stop-liquid or liquid-stop sequences that occurred word-internally or across a word boundary (e.g., *nɛkla* and *nɛlka* versus *mɛnɛk lano* and *mɛnɛl kano*). There were eight words and eight word pairs with each sequence order. Half of the set had intervocalic -kl- or -lk- sequences and half had intervocalic -tr- or -rt- sequences. These two consonant combinations were chosen with a cross-language comparison in mind: the sequences obey word-medial and word-edge phonotactics in English, Russian, and Finnish.

TABLE I. The experimental stimuli were designed so that the intervocalic stop-liquid and liquid-stop sequences occurred either within a word (internal) or across a word boundary (peripheral).

Stop-liquid		Liquid-stop	
Internal	Peripheral	Internal	Peripheral
jɪt4*a*	mɛjɪt ɹanoʊ	jɪɹ4t*a*	mɛjɪɹ tanoʊ
kɛt4*a*	mɛkɛt ɹanoʊ	kɛɹt*a*	mɛkɛɹ tanoʊ
lɪtɹe	mɛlɪt ɹɛnoʊ	lɪɹte	mɛlɪɹ tɛnoʊ
mɛtɹ*a*	mɛmɛt 4anoʊ	mɛ4t*a*	mɛmɛɹ tanoʊ
nɛkl*a*	mɛnɛk lanoʊ	nɛlk*a*	mɛnɛl kanoʊ
sɛkl*a*	mɛsɛk lanoʊ	sɛlk*a*	mɛsɛl kanoʊ
tɪkl*a*	mɛtɪk lanoʊ	tɪlk*a*	mɛtɪl kanoʊ
vɪkle	mɛvɪk lɛnoʊ	vɪlke	mɛvɪl kɛnoʊ

The stimuli were embedded in the frame sentence, *Say___again*, and presented to speakers in writing as a randomized list of phrases. Each of these phrases was produced as a coherent utterance with one low-high-low intonational contour that focused the nonsense word stimuli and tended to peak on the first nonsense word when the stimulus was a nonsense word pair.

To ensure consistent vowel pronunciation across speakers, the stimuli were represented using the International Phonetic Alphabet (IPAs) i.e., as a phonemic transcription, e.g., *nɛkla*). Speakers were reminded of or taught the sound correspondence of all non-English graphemes (e.g., /ɛ/ and /ɹ/) and were asked to practice producing all of the stimuli before the recording session began. Feedback during practice and recording was given only when a symbol was clearly misinterpreted (e.g., /e/ for /ɛ/). With respect to pronunciation, it is important to emphasize that the experimenter and speakers treated the IPA representation as mere orthography: speakers produced whatever allophonic variation they found natural given the different contexts and conditions.

Each of the stimuli was repeated three times for a total of 96 phrases per speaker, that is, 24 tokens with word-internal stop-liquid sequences, 24 tokens with word-peripheral stop-liquid sequences, 24 tokens with word-internal liquid-stop sequences, and 24 tokens with word-peripheral liquid-stop sequences. All 96 phrases were spoken twice—once with the weak-strong stress pattern typical of English verbs, and once with the strong-weak stress pattern typical of English nouns. Stress was marked as vowel accent, and all stimuli were produced with the same weak-strong or strong-weak pattern in a given recording session (e.g., session 1: *nɛklá,mɛnɛ́k lanó*; session 2: *nɛ́kla,mɛ́nek láno*). The two sessions were conducted on separate days, with the different stress patterns completed in different random orders by different speakers. Speakers practiced the stress pattern of the day at the same time they practiced pronouncing the stimuli.

2. Measurement and analyses

The utterances were recorded in a sound-attenuated room using a Shure BG 5.1 microphone, and saved in digital format. They were later displayed in Praat (Boersma and Weenink, 2002) as oscillograms and spectrograms, and the acoustic durations of the intervocalic stop and liquid were measured. The measurement criteria were as follows. Stop boundaries were defined by a sudden drop/rise in the amplitude of a periodic wave form. Closure duration and burst plus aspiration duration (henceforth aspiration) were measured separately for clearer insight into the relationship between acoustic duration and articulatory changes in the vocal tract. Liquid boundaries were defined by amplitude and frequency changes in the periodic wave form on one side, and by the stop boundary on the other. In particular, the boundary between an /l/ and a vowel was located at the point where energy in the midfrequencies abruptly changed and wave form amplitude increased. The boundary between an /ɹ/ and an adjacent vowel was also placed wherever there was an abrupt change in energy. If no such changes occurred, then the boundary was placed at the temporal midpoint of the transitioning third formant (Klatt, 1976). Ambiguity in defining segment boundaries was resolved by repeated listening to different sections of the wave form and by establishing the boundary where spectral and amplitude changes were observed that corresponded to an auditory segmentation of the wave form.

To test for measurement consistency, 10% of the tokens produced by all the speakers were randomly selected for remeasurement and then the intrarater reliability was assessed using Cronbach's coefficient alpha. In all cases, the coefficient alpha was high [stop closure, $\alpha=0.91$; aspiration, $\alpha=0.96$; liquid duration, $\alpha=0.93$], indicating excellent internal measurement consistency.

Since stress was expected to affect consonantal duration, the acoustic correlates of stress were also measured; namely, vowel pitch, amplitude, and duration. The goal of measuring the stress correlates was to establish that the stress manipulation had worked and that there were differences in the relative strength of the vowels on either side of the intervocalic consonants in question. Pitch (f0) and amplitude (dB) measures were made at the midpoint of each of the flanking vowels, whether these were in a single word (e.g., *nɛkla*) or whether these spanned a word boundary (e.g., *mɛnɛk lano*). The duration of these vowels was measured from vowel onset to offset using the segmentation criteria discussed above for consonants. The pitch, amplitude, and duration measures were then converted to obtain a single dependent variable for the repeated measures analyses of variance (ANOVA) that would reflect the relative strength of the two flanking vowels. For the pitch and duration measures, this value was the ratio of the pre- and postconsonantal pitch or duration values. For the amplitude measure, it was a difference value.

All measures were evaluated as a function of the within-subject factors in two- or three-way repeated measures ANOVAs. The relative strength of vowels was evaluated as a function of two within-subjects factors: stress (weak or strong) and word position (internal or across a word boundary). Consonantal durations were evaluated as a function of three within-subjects factors: stress, sequence order (stop-liquid and liquid-stop), and position. In both the vocalic and consonantal analyses, the levels in the factor stress represented the nominal stress pattern of the pre- and postconsonantal vowels, not the nominal stress pattern of the disyllabic nonsense word or word pairs. This means that items with a weak-strong pattern, for example, included the single nonsense words produced with weak-strong stress and nonsense word pairs produced with strong-weak stress (S<u>W#S</u>W). Stress was coded in this way for the analyses to facilitate interpretation of the results.

B. Results

1. Stress

The analyses showed that the relative strength of the pre- and postconsonantal vowels varied with stress, as expected [$f0$ ratio, $F(1,287)=467.21, p<0.01, \eta_p^2=0.62$; dB difference, $F(1,287)=786.71, p<0.01, \eta_p^2=0.73$; duration ratio, $F(1,287)=793.08, p<0.01, \eta_p^2=0.73$]. The second vowel was longer, louder, and higher pitched than the first vowel when these were produced with weak-strong stress. Conversely, the first vowel was longer, louder, and higher pitched than the second vowel when these were produced with strong-weak stress. This effect of stress on relative pitch and amplitude was greatly reduced when a word boundary intervened, as is evident from Fig. 1 where the significant interaction between position and stress pattern is shown for the three dependent variables [$f0$ ratio, $F(1,287)=393.04, p<0.01, \eta_p^2=0.58$; dB difference, $f(1,287)=350.87, p<0.01, \eta_p^2=0.55$; duration ratio, $F(1,287)=157.26, p<0.01, \eta_p^2=0.35$]. Post hoc tests indicate that the stress pattern did not affect the relative pitch of vowels on either side of a word boundary [$F(1,287)=3.09$, NS], though it did affect their relative amplitude [$F(1,287)=56.43, p<0.01$]. In contrast to relative pitch and amplitude, stress had a large effect on the relative duration of vowels on either side of a word boundary, as shown in the bottom panel of Fig. 1.

The different patterns of interaction between stress and position for the three dependent variables is likely due to an interaction with focal accent. The six speakers placed the focal accent of the phrase on the nonsense word. If the stimulus consisted of two disyllabic words, the focal accent was typically placed on the first word rather than on the second. Focal accent likely increased the pitch and amplitude of the stressed syllable on words where accent peaked, but not on those outside the accent peak (van Sluijter and van Heuven, 1996). Vowel duration, which appeared to be unaffected by peak focal accent, is therefore the more reliable indicator of the nominal stress pattern.

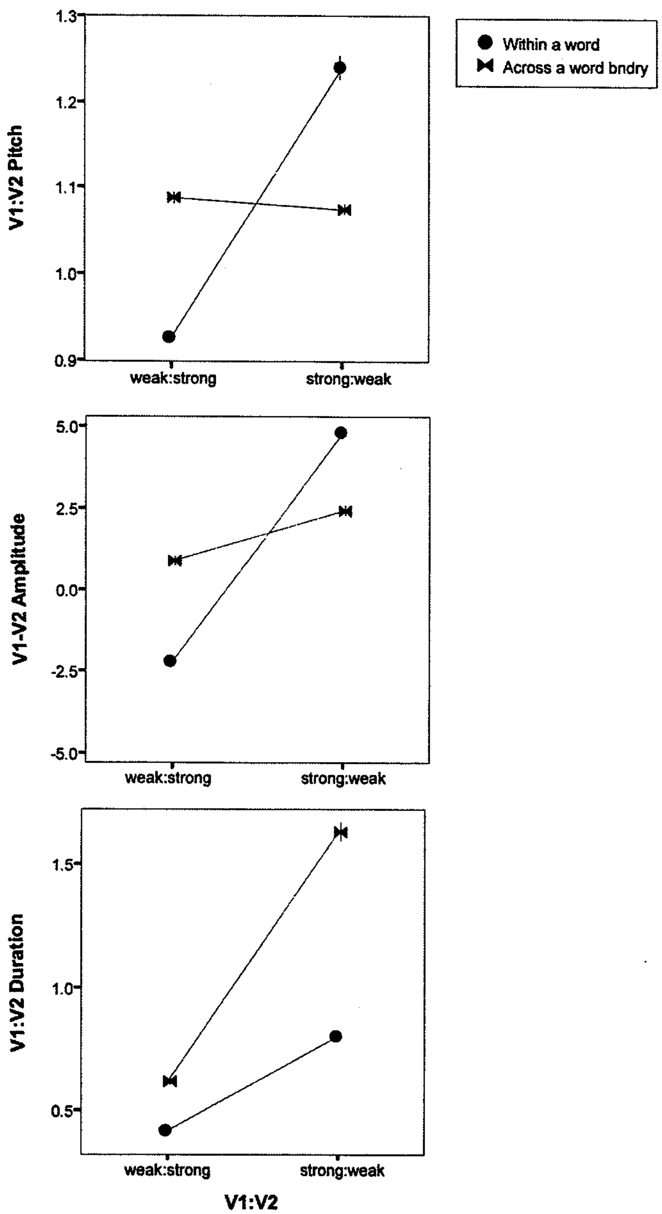

FIG. 1. The relative pitch, amplitude, and duration is shown for the vowels on either side of the consonant sequence that was the focus of the present experiment. The three panels of the figure show from top to bottom the $f0$ ratio, the dB difference, and the duration ratio for the pre- ($V1$) and post-consonantal ($V2$) vowels as a function of their word position and stress specifications ($V1:V2$).

2. Consonantal durations

The analyses showed that the effects of stress and sequencing on consonantal duration varied with word position and was different for the different consonantal articulations under consideration. The results on stop closure, stop aspiration, and liquid duration are described in turn with respect to the hypotheses.

Closure duration varied systematically with stress and sequencing [$F(1,143)=83.99$, $p<0.01$, $\eta_p^2=0.37$], but the simple effects of stress and sequencing were not statistically significant. Figure 2 shows that postvocalic stop closure is longer when the stop follows a stressed vowel than when it follows an unstressed vowel. Similarly, prevocalic stop closure is longer when the stop precedes a stressed vowel than an unstressed vowel. Thus, the pattern of interaction between stress and sequencing suggests that stop closure duration is influenced only by adjacent vowel strength. Stress does not influence closure duration when a liquid consonant intervenes between the stop and the vowel.

Figure 2 also shows that the pattern of interaction between stress and sequencing varied with word position [$F(1,143)=8.24, p<0.01, \eta_p^2=0.06$]. In particular, stress had a greater influence on closure duration at the edges of words than within a word (i.e., the interaction between stress and sequencing was even stronger at the edges of words). This result is consistent with the juncture hypothesis, which predicts that consonantal duration patterns are amplified at a word boundary. The result also parallels the finding that stress-induced differences in vowel duration were larger across a word boundary than within a word.

Aspiration duration also varied systematically with stress and sequencing [$F(1,143)=26.01$, $p<0.01$, $\eta_p^2=0.15$], but the simple effects were also significant [stress, $F(1,143)=208.86$, $p<0.01$, $\eta_p^2=0.59$; sequencing, $F(1,143)=307.4$, $p<0.01, \eta_p^2=0.68$], indicating that the pattern of interaction between the factors was different than that observed for closure duration. In particular, Fig. 3 shows that aspiration duration was influenced by the strength of the following vowel, rather than by adjacent vowel strength. Aspiration duration was longer before a stressed vowel than before an unstressed vowel regardless of whether the stop was first or second in the consonant sequence.

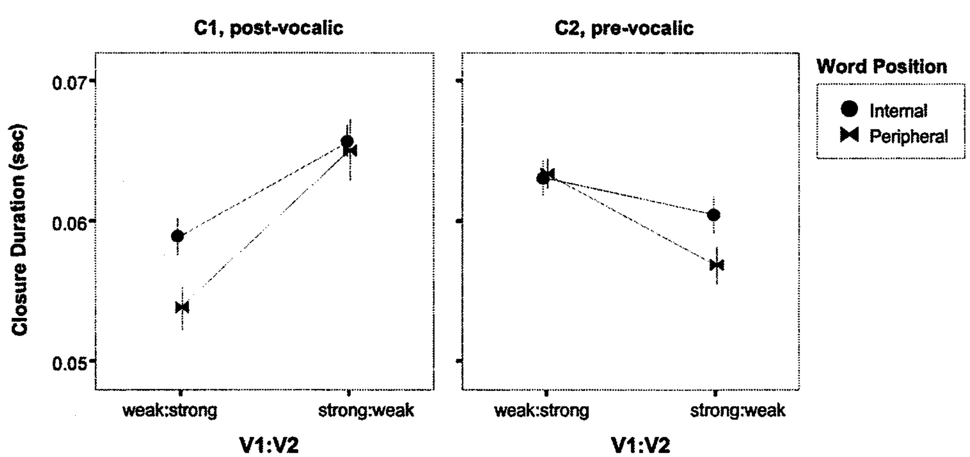

FIG. 2. Stop closure duration is shown as a function of sequence order ($C1$ or $C2$), word position (internal or peripheral), and stress specification of the adjacent pre- and postconsonantal vowels (weak:strong or strong:weak).

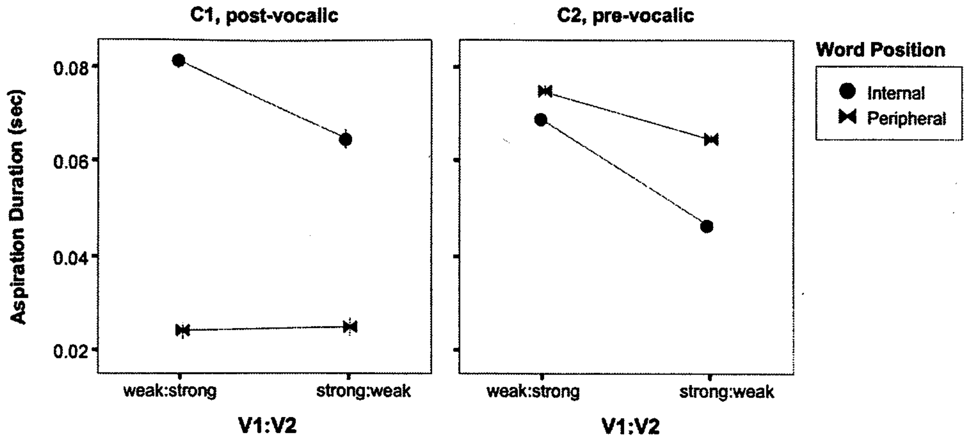

FIG. 3. Stop aspiration duration is shown as a function of sequence order ($C1$ or $C2$), word position (internal or peripheral), and stress specification of the adjacent pre- and postconsonantal vowels (weak:strong or strong:weak).

The effect of sequence cannot be discussed without describing the interaction between sequencing and position that is also evident in Fig. 3 [$F(1,143)=500.04$, $p<0.01$, $\eta_p^2=0.78$]. Word-internally, the effect of sequencing was consistent with the prediction from coarticulation: aspiration duration was greater when it preceded a liquid ($C1$, postvocalic) than when it preceded a vowel ($C2$, prevocalic). At the word boundary, however, only prevocalic ($C2$, word-initial) stops were significantly aspirated. Postvocalic ($C1$, word-final) stops were produced with little to no aspiration. Prevocalic stops also had longer aspiration durations overall at a word boundary than they did word-internally.

The observed interaction between sequence and position is consisent with the juncture hypothesis, but not with the syllable frame hypothesis. In particular, the syllable frame hypothesis predicts aspiration in syllable-onset position and no aspiration in syllable-offset position, consistent with position-dependent allophonic variation of stop aspiration in English. This prediction was clearly upheld at the word boundary. However, word-internal stops were aspirated before a liquid regardless of stress even though the strong-weak stress manipulation was expected to result in offset-onset stop-liquid sequences. Nonetheless, word-internal syllable boundaries are difficult to establish according to independent criteria. So, the word-internal pattern could indicate that speakers syllabified word-internal stop-liquid sequences as onset clusters to the second syllable regardless of lexical stress. If this is the case, then the results can be interpreted to support a syllable frame hypothesis. If it is not the case, then the results are only consistent with the juncture hypothesis.

The problem of distinguishing between the syllable frame and the juncture hypotheses is further exacerbated by the results on liquid duration. Liquid duration varied with stress [$F(1,143)=20.16$, $p<0.01$, $\eta_p^2=0.12$], sequencing [$F(1,143)=27.39$, $p<0.01$, $\eta_p^2=0.16$], and position [$F(1,143)=291.02$, $p<0.01$, $\eta_p^2=0.67$]. All interactions between these factors were also significant [stress $\times$ sequencing, $F(1,143)=222.12$, $p<0.01$, $\eta_p^2=0.61$; stress $\times$ position, $F(1,143)=29.91$, $p<0.01$, $\eta_p^2=0.17$; sequencing $\times$ position, $F(1,143)=69.21$, $p<0.01$, $\eta_p^2=0.33$; stress $\times$sequencing $\times$ position, $F(1,143)=11.97$, $p<0.01$, $\eta_p^2=0.08$]. This complex pattern of results is best understood with reference to Fig. 4.

Figure 4 shows two principal results: (1) liquid duration is greater at a word boundary than it is word-internally and (2) adjacent vowel stress affects liquid duration except when liquids follow a stop word-internally. These results are consistent with the juncture hypothesis, which predicts boundary highlighting and unimpeded coarticulation of word-internal stop-liquid sequences. The word-internal result may also be consistent with the syllable frame hypothesis, but only if syllable boundary location is unaffected by lexical stress. In particular, short prevocalic liquids are expected if the word-internal stop-liquid sequence forms an onset cluster to the second vowel (Haggard, 1973).

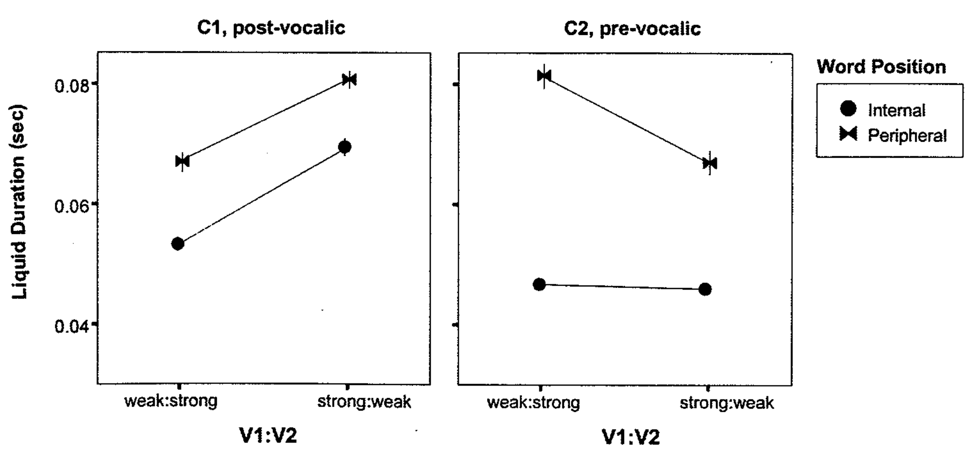

FIG. 4. Liquid duration is shown as a function of sequence order ($C1$ or $C2$), word position (internal or peripheral), and stress specification of the adjacent pre- and postconsonantal vowels (weak:strong or strong:weak).

C. Discussion of stress, sequencing, and juncture effects on segmental duration

Overall, the results show that segmental duration patterns within a word differ from the patterns that occur across words, and that stress and sequencing may be sufficient for explaining word-internal consonantal duration patterns. The duration ratio that captures the flanking weak-strong and strong-weak vowel patterns is greater across a word boundary than within a word. Closure duration was also affected by stress and boundary strength: adjacent vowel stress affected stop closure duration, but the relative increase or decrease was greater across a word boundary than within a word. These findings support the juncture hypothesis, which predicts durational patterns that highlight word boundaries. The hypothesis also receives some support from the results on stop aspiration and liquid duration. While postvocalic stop aspiration was nearly absent in word-final position, prevocalic stop aspiration was greater at a word boundary than within a word. Liquid duration was longer overall at a word boundary than word-internally, suggesting boundary-dependent lengthening and within unit coarticulation.

The latter two results on stop aspiration and liquid duration are also consistent with the syllable frame hypothesis, but only if syllable boundaries are defined in a post hoc fashion. In particular, the effect of sequence on word-internal stop aspiration and liquid duration suggests that stop-liquid sequences were always coarticulated with the following vowel. Coarticulation of the stop-liquid sequence with the following vowel is consistent with an onset cluster syllabification. However, in the strong-weak condition this syllabification is at odds with the psycholinguistic literature, which indicates that listeners would syllabify the sequence as an offset-onset when the preconsonantal vowel is stressed (Treiman and Danis, 1988; Treiman and Zukowski, 1990; Derwing and Nearey, 1991; Redford and Randall, 2005). Nonetheless, it is possible that listener boundary perception is irrelevant to production. If this is the case, then the results on word-internal patterns cannot be used to disconfirm the syllable frame hypothesis.

On the other hand, the possibility that syllable structure is better defined in speech according to output patterns rather than according to independent criteria (e.g., listener judgments) suggests that the syllable may not define articulatory dynamics so much as emerge from them. Such a view of syllables is consistent with the juncture hypothesis, which presumes that the word-internal consonantal duration patterns produced by American English speakers reflect basic coarticulatory processes rather than the effects of a linguistic boundary. This possibility is considered further below.

III. COMPARISONS OF ENGLISH CONSONANTAL DURATION PATTERNS WITH RUSSIAN AND FINNISH PATTERNS

Whether or not we assume that the stress manipulation resulted in two different syllabifications of the word-medial stop-liquid sequence, the acoustic results indicate that word-internal patterns in American English can be explained to result from sequencing and the implementation of stress. In particular, stop aspiration may be greater before a liquid than before a vowel for aerodynamic reasons; namely, coarticulation with the following liquid will mean greater closure in the vocal tract during the stop release gesture, which may make it difficult to re-establish the transglottal pressure differential needed for voicing (Docherty, 1992). Liquids that occur after stop consonants have shorter acoustic duration for a similar reason; namely, because the onset of voicing is deferred after a stop consonant. In other words, the dynamics of articulatory sequencing create different consonantal duration patterns word-internally. Further differences are induced by adjacent vowel strength in the case of stop closure duration and liquid duration, and with subsequent vowel strength in the case of stop aspiration. Stressed vowels are hyperarticulated, that is, they are produced with greater jaw opening and more accurate target attainment (de Jong, Beckman, and Edwards, 1993; de Jong, 1995). The longer articulatory trajectories result in vowel lengthening and, it would seem from the present results, in the lengthening of consonantal gestures that are coarticulated with the vowel.

Although stress and sequencing are sufficient to explain the word-internal patterns in isolation, they are not sufficient to explain (1) why the effects of adjacent vowel strength on closure duration are stronger at a word boundary than word-internally or (2) why stops aspiration is nearly absent in postvocalic (word-final) position and especially robust in prevocalic (word-initial) position or (3) why liquid duration is longer overall at a word boundary than it is word-internally. The juncture hypothesis assumes that all of these differences exist to highlight word boundary location because English phonology and morphology interact to make such boundaries opaque: phonotactics are insufficient for establishing boundary location because different types of syllable onsets and offsets are possible (e.g., *an aim versus a name, odd rum versus a drum*); and word beginnings and endings are not consistently marked with a set of recurring sounds or sound strings (cf. *What is your last name? Kakaja tvoja familija?*). It follows from the English case that languages with simpler syllable structure or more synthetic-inflexional structure will have more transparent word boundaries. It is possible that in such languages stress and sequencing will suffice to describe the consonantal duration patterns both word-internally and at word boundaries. This prediction is tested here in Russian and Finnish: two languages that are more highly synthetic and inflected than English, and one of which—Finnish—also has a simpler syllable structure.

A. Methods

Six native Russian speakers—three male and three female—and six native Finnish speakers—one male and five female—participated in the experiment. Russian and Finnish speakers produced the same stimuli as the English speakers, namely, disyllabic stimuli with intervocalic stop-liquid and liquid-stop sequences that occurred either within a word (e.g., *nɛkla* and *nɛlka*) or across a word boundary (e.g., *mεnek lano* and *mεnek lano*). Again, there were eight of each type—half with -kl- or -lk- sequences and half with -tr- or -rt- sequences (see Table I).

The Russian and Finnish stimuli were embedded in frame sentences, *Skazhi___snova* and *Sano___taas*, which means *Say___again* in Russian and Finnish, respectively. As in the English experiment, each of the stimuli was repeated three times for a total of 96 phrases per speaker. Russian and Finnish speakers were presented with the phrase in randomized order on a sheet of paper in normal orthography—Cyrillic for Russian and Roman for Finnish. Normal orthography could be used in place of the IPA because, unlike English, Russian and Finnish orthography is fairly transparent. Russian speakers produced the stimuli with weak-strong lexical stress, a common pattern in Russian. Finnish speakers produced the stimuli with strong-weak lexical stress, the obligatory pattern in Finnish.

The segmentation criteria and measurements were made as before. The principal difference between the sound patterns of the English and the Russian and Finnish stimuli was in the rhotic, which was trilled and sometimes tapped in both Russian and Finnish. The boundary between this segment and an adjacent vowel was easily discovered by an abrupt decrease in the amplitude of the periodic waveform. As before, intrarater reliability was assessed using Cronbach's coefficient alpha computed over 10% of the data (randomly selected and remeasured). In all cases, the coefficient alpha was high [Russian: stop closure, $\alpha=0.90$; aspiration, $\alpha=0.96$; liquid duration, $\alpha=0.92$; Finnish: stop closure, $\alpha=0.98$; aspiration, $\alpha=0.93$; liquid duration, $\alpha=0.92$], indicating excellent internal measurement consistency.

Once again the design was within-subjects so the effects of the within subjects factors (sequence order and/or position) on the vowel measures and consonantal durations were assessed in one- and two-way repeated measure ANOVAs. Instead of treating Russian and Finnish as a between-subjects factor, we analyzed the data from the different languages separately. The separate language analyses were in keeping with the English-only analyses reported in the first experiment (II above). A more fundamental reason for splitting the data, though, was the fact that decades of research on language-specific phonetics give us every reason to expect language-specific differences in phonetic detail, such as segmental duration. Thus, a qualitative comparison of the durational patterns across languages was deamed more meaningful than a statistical demonstration of difference. The results from the analyses are reported below.

B. Results

1. Stress

As expected, the relative pitch, amplitude, and duration of vowels varied with lexical stress placement. This effect was manifest as pitch, amplitude, and duration differences in the vowels that flanked the consonantal sequence. In Russian, the *f*0 ratios, amplitude differences, and duration ratios of the preconsonantal to postconsonantal vowel were higher across a word boundary than they were within a word [*f*0 ratio, $F(1,287)=34.71$, $p<0.01$, $\eta_p^2=0.11$; dB difference, $F(1,287)=65.59$, $p<0.01$, $\eta_p^2=0.19$; duration ratio, $F(1,287)=734.35$, $p<0.01$, $\eta_p^2=0.72$]. This is because the preconsonantal to postconsonantal vowel pattern in Russian was weak-strong within a word and strong-weak across a word boundary. In Finnish, the vowels that flanked the consonantal sequence had the opposite pattern: strong-weak within a word and weak-strong across a word boundary. Accordingly, the *f*0 ratios, amplitude differences, and duration ratios were higher within a word than across a word boundary in Finnish [*f*0 ratio, $1,287F(1,287)=113.88$, $p<0.01$, $\eta_p^2=0.28$; dB difference, $F(1,287)=423.99$, $p<0.01$, $\eta_p^2=0.60$; $F(1,287)=221.02$, $p<0.01$, $\eta_p^2=0.44$]. The Russian and Finnish results on the vowel-to-vowel duration ratio are shown as a function of vowel location in Fig. 5. The American English results from the weak-strong and strong-weak stress conditions are included for comparison.

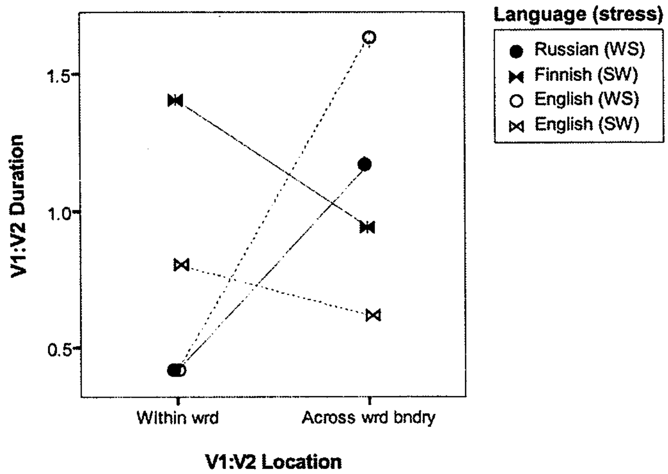

FIG. 5. The relative duration of the pre- and postconsonantal vowels is shown for the different languages and word positions. Note that the stress specification of these vowels changes across a word boundary within a language (e.g., #WS# is compared with #WS#WS#). The American English patterns are shown for comparison. These patterns as the same as those displayed in Fig. 1, but in Fig. 1 the duration ratios were displayed as a function of the stress specification for the flanking vowels (e.g., #WS# was compared with #SW#SW#).

Again, it is important to note that the pitch, amplitude, and duration patterns are different within a word than they are between words. This is because stress was always placed on the second syllable in Russian and on the first syllable in Finnish. When these stress patterns are realized on sequential disyllabic words, the result is a shift in stress across a word boundary. For example, Russian *nεklá* with weak-strong stress has a weak *V*1 and a strong *V*2, but the same vowels in Russian *mεnék lanó* are strong-weak (WS#WS). This stress shift is also apparent in the American English data, when these are graphed by stress condition rather than by the stress pattern associated with the vowels on either side of the focal consonant sequence (viz. Figs. 1–4).

2. Consonantal durations

The stress shift across a word boundary means that the effects of adjacent vowel stress on consonantal durations in Russian and Finnish will be shown by an interaction between sequencing and word position. This type of interaction was

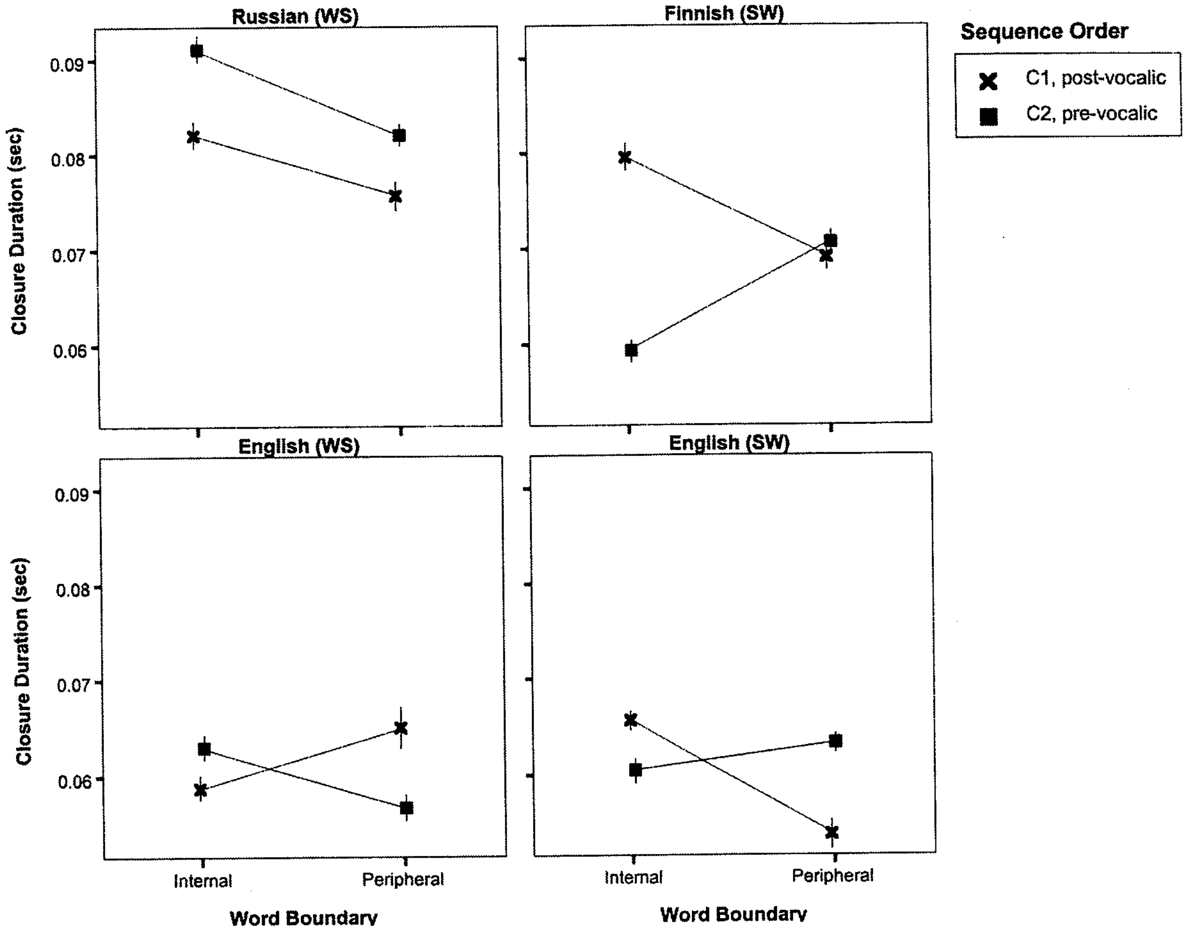

FIG. 6. Stop closure duration is shown as a function of language (stress pattern), sequence order ($C1$ or $C2$), and word position (internal or peripheral). The American English data from Fig. 2 are redrawn here to facilitate comparison with the Russian and Finnish data.

statistically significant for stop closure duration in Finnish [$F(1,143)=92.75$, $p<0.01$, $\eta_p^2=0.39$] and for liquid duration in both Russian and Finnish [Russian, $F(1,143)=20.46$, $p<0.01$, $\eta_p^2=0.13$; Finnish, $F(1,143)=85.22$, $p<0.01$, $\eta_p^2=0.37$]. Note that the American English data show a similar effect of stress on stop closure duration in both stress conditions and on liquid duration in the strong-weak stress condition. The results on stop closure duration in Russian and Finnish are shown in Fig. 6 and those on liquid duration are shown in Fig. 7. The English data are included for comparison.

Figures 6 and 7 are also instructive because they show that sequencing explains stop closure duration in Russian, even when stress does not; and that sequencing is generally an important factor in explaining stop closure and liquid duration across both languages [Russian stop closure, $F(1,143)=21.46$, $p<0.01$, $\eta_p^2=0.13$; Finnish stop closure, $F(1,143)=49.74$, $p<0.01$, $\eta_p^2=0.26$; Russian liquid, $F(1,143)=46.50$, $p<0.01$, $\eta_p^2=0.25$; Finnish liquid, $F(1,143)=131.34$, $p<0.01$, $\eta_p^2=0.48$]. In particular, prevocalic stop closure ($C2$) is longer than postvocalic stop closure ($C1$) in Russian regardless of word position, and postvocalic liquids ($C1$) are longer than prevocalic liquids ($C2$) regardless of position.

Sequencing also influences liquid duration in American English, but the effect of word position is even more obvious (see Figs. 4 and 7). Word-internal liquids are short, except when they occur after a stressed vowel. Liquids at a word boundary are long, though influenced by the strength of an adjacent vowel.

Finally, Figs. 6 and 7 show that the effect of word position in Russian and Finnish is different from their effect in English. In English, sequencing effects are amplified at a word boundary, that is, the difference between postvocalic and prevocalic stop closure is greater at a word boundary than word-internally. In contrast, position induced differences in closure duration are smaller at a word boundary in Russian and Finnish than they are within a word.

Finnish stops and liquids show the opposite pattern of Russian liquids, but this pattern is nonetheless quite different from English. In Russian, liquid duration is more influenced by adjacent vowel stress at a word boundary than word-internally, but the interaction with sequence order negates a simple effect of word position on duration [$F(1,143)=2.95$, NS]. This means that there is no net lengthening effect for liquids at a word boundary in Russian as there is in English. In Finnish, the effects of stress are stronger word-internally than they are at a word boundary. This negates a simple effect of word position on duration for stop closure [$F(1,143)=0.13$, NS], though the effect on liquid duration reaches significance—on average liquids are longer at a word boundary [$F(1,143)=4.04$, $p<0.05$, $\eta_p^2=0.03$]. Although liq-

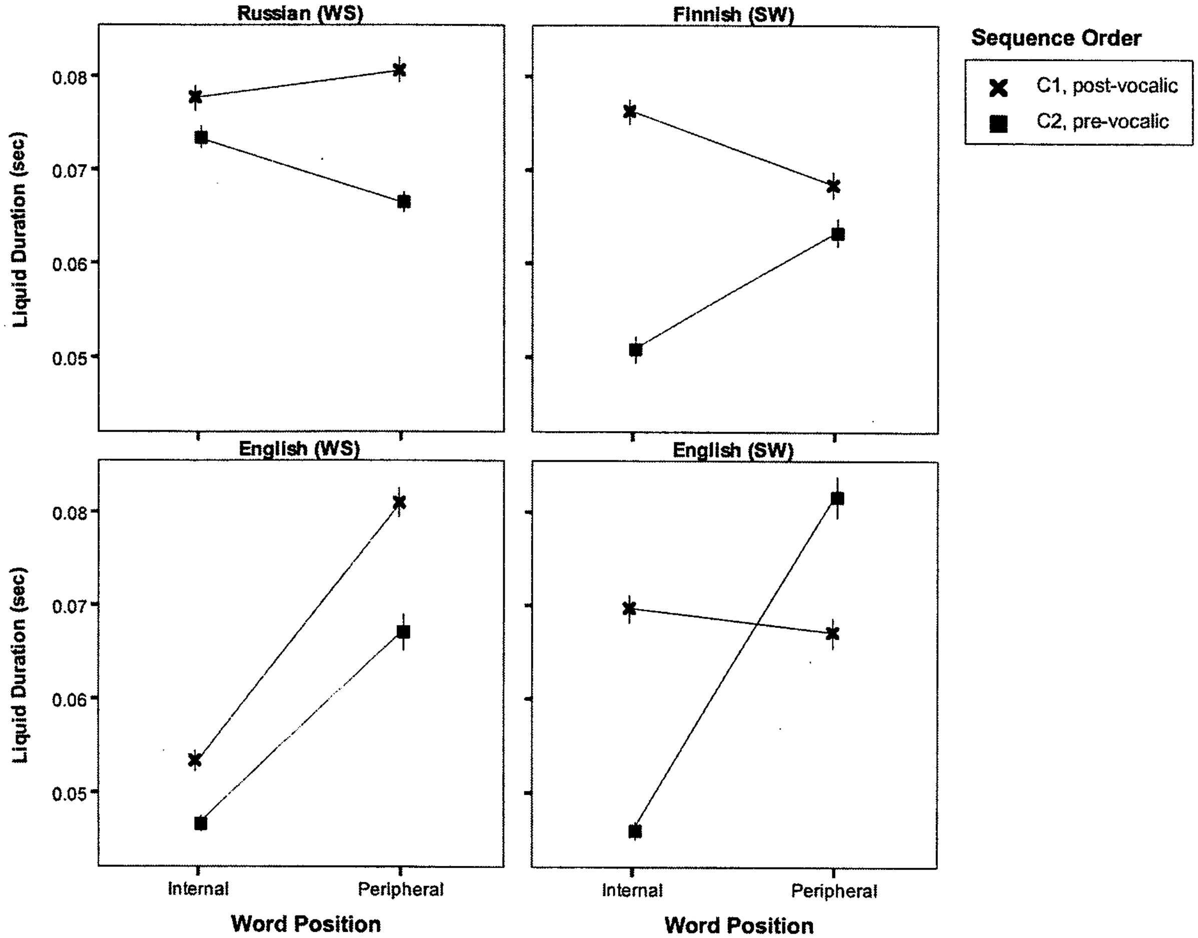

FIG. 7. Liquid duration is shown as a function of language (stress pattern), sequence order (*C*1 or *C*2), and word position (internal or peripheral). The American English data from Fig. 4 are redrawn here to facilitate comparison with the Russian and Finnish data.

uids may be longer on average at a word boundary in Finnish, examination of Figs. 6 and 7 leaves the impression that consonantal duration patterns are amplified word-internally in Finnish, which is opposite of the English results where such patterns are amplified at a word boundary. However, the Finnish pattern parallels the greater word-internal differences in vowel duration, and this is similar to the English pattern: stop closure and liquid duration differences are also greater in English where vowel duration differences are greater (albeit across a word boundary rather than word-internally as in Finnish).

The minimal effect of word boundaries in Russian and Finnish is most apparent in the results on stop aspiration duration. Figure 8 shows that Russian stop aspiration is unaffected by word position, and minimally affected by adjacent vowel stress (i.e., there is a nonsignificant trend towards an interaction between sequence and position [$F(1,143)=3.65$, $p=0.058$, $\eta_p^2=0.03$]). Finnish stop aspiration is affected by sequence [$F(1,143)=198.62$, $p<0.01$, $\eta_p^2=0.58$] and by word position [$F(1,143)=13.18$, $p<0.01$, $\eta_p^2=0.08$], which likely reflects the strength of the subsequent vowel as it does word-internally in English.

Figure 8 shows that the effect of sequencing in Finnish is the same within a word as it is across a word boundary. Stops that occur before liquids are produced with more aspiration than stops that occur before vowels. This result differs from English, where only word-internal stops are longer before liquids. At the word's edge, English stop aspiration is determined by whether the stop occurs word-finally or word-initially.

C. Discussion of the comparison between Russian, Finnish, and English

A comparison of Russian, Finnish, and American English consonantal duration patterns shows that stress and sequencing affect duration in similar ways across languages, though there are language-specific differences in the extent to which any one factor explains the durational patterns. For instance, vowel stress has a lengthening effect on adjacent consonants in both Russian and Finnish, but stress affects consonantal duration less in Russian than in Finnish.

Whereas the effect of stress and sequencing is similar across languages, the effect of word position is different. In English, differences in stop closure duration and aspiration are especially strong at a word boundary, and liquids are especially long. The opposite pattern is found in Russian and Finnish. In Russian, stop closure duration is longer within a word than at a word boundary, and there is no effect of word position on aspiration. In Finnish, the differences in stop closure and liquid duration are especially strong within a word compared to at a word boundary.

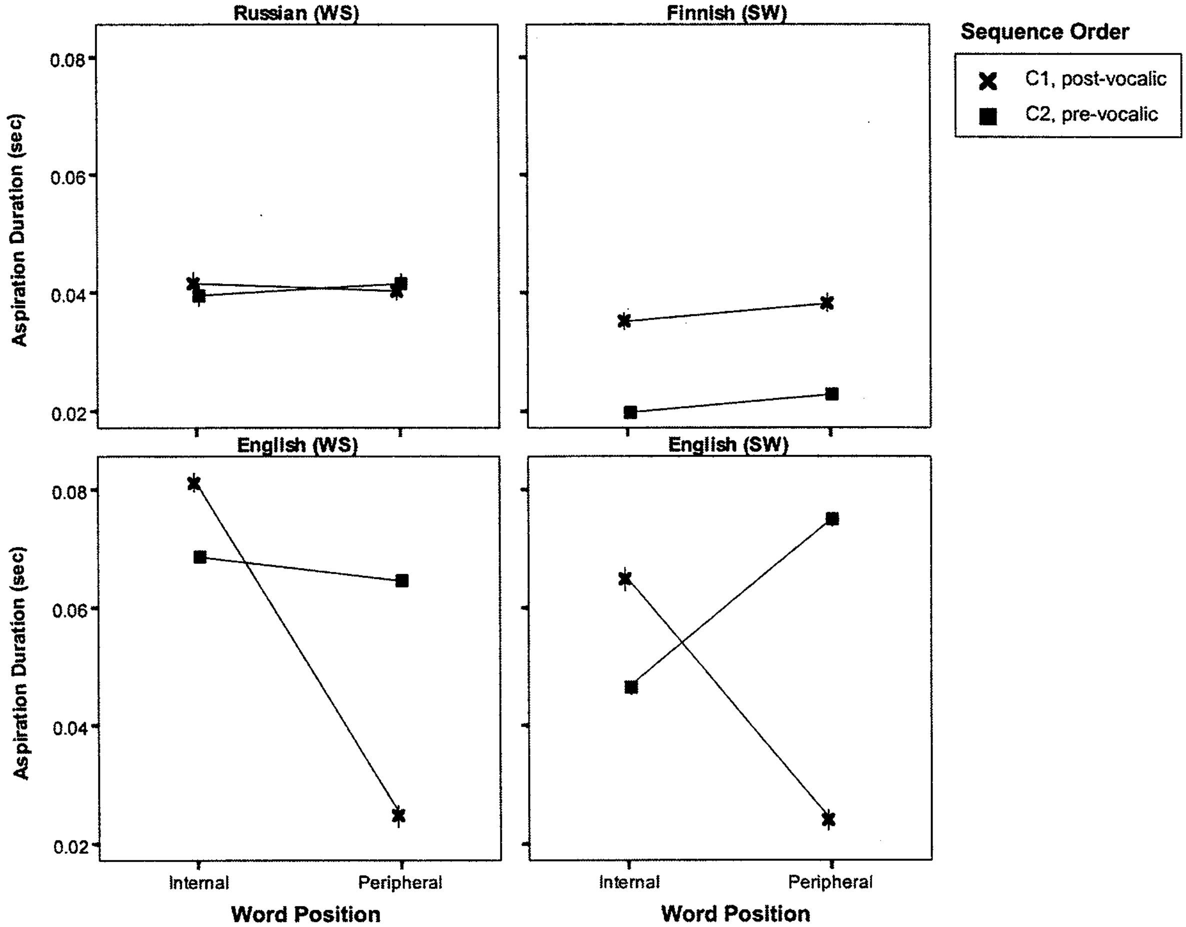

FIG. 8. Stop aspiration duration is shown as a function of language (stress pattern), sequence order ($C1$ or $C2$), and word position (internal or peripheral). The American English data from Fig. 3 are redrawn here to facilitate comparison with the Russian and Finnish data.

Overall, the different effects of word position in English on the one hand and Russian and Finnish on the other support the juncture hypothesis for English. The results also suggest language-specific differences in how sequential action is organized for output. If the boundaries of planning units are signaled by robust durational patterns because longer consonantal durations indicate less articulatory overlap, then the word-internal consonantal duration patterns in Russian and Finnish suggest that a sublexical unit may be important for organizing action in these languages. Further research is needed to determine whether the Russian unit is an articulatory syllable, as defined by Kozhevnikov and Chistovich (1965), and the Finnish unit is the mora, as suggested by Suomi's work (Suomi and Ylitalo, 2004; Suomi, 2005).

IV. CONCLUSION

This study contributes to a long tradition of research that uses segmental duration patterns to understand what suprasegmental units are involved in motor speech planning. Since the beginning of this tradition, the linguistic syllable has been suggested as a basic unit in the plan. Here, this suggestion was restated in terms of a generic syllable frame hypothesis, which was then investigated with two problems in mind. The principal problem with the hypothesis is that word-internal syllable boundaries are difficult to define for particular sequences. Accordingly, the boundaries of the unit can either be defined post hoc and with respect to observed durational patterns (e.g., Kozhevnikov and Chistovich, 1965) or they can be defined a priori and with respect to word boundaries (e.g., word-final equals syllable-final), thereby confounding the role of the syllable and word in speech. In cases where word-internal syllable boundaries are less disputed (e.g., liquid-stop sequences), there is a secondary problem with the hypothesis; namely, that segmental duration is affected by intrinsic duration, coarticulatory dynamics, and the rhythm structure of a language. This means that durational patterns attributed to the syllable in English may be better understood as emergent from articulatory processes having to do with the realization of stress and particular segmental sequencing patterns.

In the current study, the syllable frame hypothesis would have been clearly supported and alternate explanations clearly refuted in the American English data if the stress manipulation on disyllabic words with medial stop-liquid sequences had resulted in distinct durational patterns: one that was consistent with an onset cluster pattern when the disyllabic word received weak-strong stress; and one that mirrored the word-final/word-initial pattern when the disyllabic word received strong-weak stress. Instead, the results on this class of words only support the syllable frame hypothesis if word-internal syllable structure is redefined to fit the durational data.

In the absence of clear support for the syllable frame hypothesis, the results were explained with reference to stress, sequencing, and the communicative function of language. The remainder of this section outlines a production model proposed to integrate these factors into an explanation for positional effects on consonantal duration—one that heeds the emergence of cross-language differences in durational patterns.

The proposed model assumes that the rhythm and sound patterns of a linguistic unit are linked in production via jaw movement. Such an assumption follows from the work on prosodic structure and jaw movement (Stone, 1981; de Jong, Beckman, and Edwards, 1993; Harrington, Fletcher, and Roberts, 1995; Erickson, 1998; Erickson, Fujimura, and Pardo, 1998) as well as from the work on sound sequencing and jaw movement (MacNeilage, 1998; MacNeilage and Davis, 2000; MacNeilage *et al.*, 2000). Specifically, the linguistic unit specifies through its rhythmic structure an abstract open-close pattern to be used as an articulatory frame within which the open (vowel) and closed (consonants) segments of a sequence can be located in space and time. The rhythmic structure of the unit dictates both the number of open-close movements and movement amplitude. Assuming a constant speech rate, cycle number and amplitude will have temporal consequences on segmental articulation. In the case of a stress-timed language, the larger amplitude cycles associated with a stressed beat will allow more time for target attainment, resulting in longer segmental durations. Since the lengthening effects apply only within the cycle, only those gestures that are executed during the downward or upwards trajectory of the cycle will have increased durations. For this reason, stress may affect only the duration of an opening or closing gesture in segments that contain both, such as stop consonants. For instance, whereas closure for postvocalic stops can be initiated during either the upwards or downwards trajectory of the vowel, an audible stop release cannot occur unless the jaw is opening. This example also illustrates how stress conditions the stop consonantal duration patterns observed in the present study; in particular, it explain why adjacent vowel stress affects closure duration, but why only subsequent vowel stress explains aspiration duration.

An additional example illustrates how stress and sequencing interact in the proposed model to condition the pattern of short word-internal liquids in stop-liquid sequences observed for American English. A liquid that follows a stop can be articulated on the downward trajectory of the jaw because they are comprised of two gestures—an anterior and posterior one, which are sequenced differently depending on whether the liquid occurs in pre- or post vocalic position (Sproat and Fujimura, 1993; Gick *et al.*, 2006). Liquid acoustic duration is thus limited on the stop side by difficulties in regaining a transglottal pressure differential, but it is also limited on the vowel side. This is because the liquid is being articulated while the jaw continues in its downward trajectory and, at a certain point, the open jaw configuration may become inconsistent even with the posterior constriction gesture that precedes the vocalic gesture.

At a higher level of control, the meaningful units of language create structure within the speech plan, which interact with the size-limits of the motor programs (i.e., executable chunks of planned speech action). Future work could determine the maximal size of such programs by investigating long-distance coarticulation, but work on vowel-to-vowel coarticulation already suggests that they can include more than one cycle (e.g., Öhman, 1966). In light of such work, and the evidence presented in the current study, it is reasonable to assume that the word could be the most basic unit—at least in English. The unit may be different in a more highly synthetic-inflexional language like Finnish, where more than one programming unit may be needed to execute the average word, which is significantly longer and morphologically more complex than the average English word.

Although future work will be aimed at testing and elaborating the proposed production model, the model is useful in that it highlights two broad conclusions drawn from the current study. First, positional effects on consonantal duration may or may not refer to linguistic units, which suggests that more basic articulatory explanations for such patterns should be excluded before linguistic units such as the syllable are invoked. Second, cross-language differences in durational patterns may reflect more than language-specific phonetic tuning: such differences may reflect interesting interactions between the different ways in which languages encode meaning and universal limits on the maximal size of the motor programs in speech motor control.

ACKNOWLEDGMENTS

The author is grateful to Tatiana Furrow and Lewis Notestine for help with subject recruitment and data collection. The author also thanks (in alphabetic order) Dani Byrd, Edward Flemming, Susan Guion, Scott Myers, Kari Suomi, Marija Tabain, and Doug Whalen for their constructive criticisms of earlier versions of this paper.

[1]For example, biomechanical dependencies between the articulators may influence action sequencing and thus intergestural timing. This idea is developed further in the conclusion. It is also possible that habitual sequencing patterns may be transformed through extensive practice into syllable-sized speech motor programs. Although such programs could inform judgments on syllable boundaries [e.g., when listeners subvocalize in a segmentation task (e.g., Sato *et al.*, 2006)], they need not be specified in the phonology. Instead, the phonology need only represent abstract language-specific knowledge that emerges at the interface between language production and comprehension, which is the lexicon in some models (e.g., Bybee, 2001).

Boersma, P., and Weenink, D. (**2002**). "Praat: Doing phonetics by computer (version 4.0.34)" (Computer program). Retrieved November 2002 from http://www.praat.org/.

Boucher, V. (**1988**). "A parameter of syllabification for VstopV and relative timing invariance," J. Phonetics **16**, 299–326.

Browman, C., and Goldstein, L. (**1988**). "Some notes on syllable structure in articulatory phonology," Phonetica **45**, 140–155.

Bybee, J. (**2001**). *Phonology and Language Use* (Cambridge University Press, Cambridge).

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

Campbell, W. N. (**1992**). "Syllable-based segmental duration," in *Talking Machines: Theories, Models, and Designs*, edited by G. Bailly, C. Benoit, and T. R. Sawallis (Elsevier, North Holland) pp. 211–224.

Campbell, W. N., and Isard, S. D. (**1991**). "Segment durations in a syllable frame," J. Phonetics **19**, 37–47.

Christie, W. M. (1977). "Some multiple cues for juncture in English," General Linguistics **17**, 212–222.

de Jong, K. J. (**2003**). "Temporal constraints and characterizing syllable structuring," in *Phonetic Interpretation: Papers in Laboratory Phonology 6*, edited by J. Local, R Ogden, and R. Temple (Cambridge University Press, Cambridge), pp. 253–268.

de Jong, K. J. (**1995**). "The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation," J. Acoust. Soc. Am. **97**, 491–504.

de Jong, K., Beckman, M. E., and Edwards, J. (**1993**). "The interplay between prosodic structure and coarticulation," Lang Speech **36**, 197–212.

Derwing, B. L., and Nearey, T. M. (**1991**). "The 'vowel-stickiness' phenomenon: Three experimental sources of evidence," in *Proceedings of the XIIth International Congress of Phonetic Sciences*, France (ICPhS, Aix-en-Provence); Vol. **2**, pp. 210–213.

Docherty, G. J. (**1992**). *The Timing of Voicing in British English Obstruents* (Foris, Berlin).

Edwards, J., and Beckman, M. E. (**1988**). "Articulatory timing and the prosodic interpretation of syllable duration," Phonetica **45**, 156–174.

Erickson, D. (**1998**). "Effects of contrastive emphasis on jaw opening," Phonetica **55**, 147–169.

Erickson, D., Fujimura, O., and Pardo, B. (**1998**). "Articulatory correlates of emotion control: Emotion and emphasis," Lang Speech **41**, 399–417.

Fougeron, C., and Keating, P. (**1997**). "Articulatory strengthening at the edges of prosodic domains," J. Acoust. Soc. Am. **101**, 3728–3740.

Fowler, C. A. (**1979**). "'Perceptual centers' in speech production and perception," Percept. Psychophys. **25**, 375–388.

Gick, B., Campbell, F., Oh, S., and Tamburri-Watt, L. (**2006**). "Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids," J. Phonetics **34**, 49–72.

Haggard, M. (**1973**). "Correlations between successive segment durations: values in clusters," J. Phonetics **1**, 111–116.

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

Kahn, D. (**1976**). "Syllable-based generalizations in English phonology," Doctoral dissertation, Massachusetts Institute of Technology, Distributed by Indiana University Linguistics Club.

Keating, P., Cho, T., Fougeron, C., and Hsu, C-S. (**2003**). "Domain-initial articulatory strengthening in four languages," in *Phonetic Interpretation: Papers in Laboratory Phonology 6*, edited by J. Local, R. Ogden, and R. Temple (Cambridge University Press; New York) pp. 143–161.

Keating, P., Wright, R., and Zhang, J. (1999). "Word-level asymmetries in consonant articulation," University of California Working Papers in Phonetics **97**, 157–173.

Klatt, D. (**1976**). "Linguistic uses of segmental duration in English: Acoustic and perceptual evidence," J. Acoust. Soc. Am. **59**, 1208–1221.

Kohler, K. J. (**1991**). "The phonetics/phonology issue in the study of articulatory reduction," Phonetica **48**, 180–192.

Kozhevnikov, V., and Chistovich, L. (**1965**). *Speech: Articulation and Perception* (Joint Publication Research Services, Washington, D.C.)

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

Lehiste, I. (**1970**). *Suprasegmentals* (MIT Press, Cambridge, MA).

MacNeilage, P. F. (**1998**). "The frame/content theory of evolution of speech production," Behav. Brain Sci. **21**, 499–546.

MacNeilage, P. F., and Davis, B. L. (**2000**). "On the origin of internal structure of word forms," Science **288**, 527–531.

MacNeilage, P. F., Davis, B. L., Kinney, A., and Matyear, C. L. (**2000**). "The motor core of speech," Child Dev. **71**, 153–163.

Maddieson, I. (**1985**). "Phonetic cues to syllabification," in *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, edited by V. Fromkin (Academic, New York), pp. 203–221.

Maddieson, I. (**2004**). "Timing and alignment: A case study of Lai," Language and Linguisticis **5**, 729–755.

Moon, S.-J., and Lindblom, B. (**1994**). "Interaction between duration, context, and speaking style in English stressed vowels," J. Acoust. Soc. Am. **96**, 27–56.

Munhall, K., Fowler, C., Hawkins, S., and Saltzman, E. (**1992**). "Compensatory shortening' in monosyllables of spoken English," J. Phonetics **20**, 225–239.

Öhman, S. E. G. (**1966**). "Coarticulation in VCV utterances: Spectrographic measurements," J. Acoust. Soc. Am. **39**, 151–168.

Prieto, M. (**2002**). "Acoustic correlates of the syllable: evidence from Spanish," Doctoral dissertation, University of Illinois at Urbana-Champagne, Dissertation Abstracts International **63**, 580.

Quené, H. (**1992**). "Durational cues for word segmentation in Dutch," J. Phonetics **20**, 331–350.

Redford, M. A., and Diehl, R. L. (**1999**). "The relative perceptual distinctiveness of initial and final consonants in CVC syllables," J. Acoust. Soc. Am. **106**, 1555–1565.

Redford, M. A., and Gildersleeve-Neumann, C. E. (**2006**). "The acquisition of two phonetic cues to word boundaries" (unpublished).

Redford, M. A., and Randall, P. (**2005**). "The role of juncture cues and phonological knowledge in English syllabification judgments," J. Phonetics **33**, 27–46.

Sato, M., Schwartz, J.-L., Abry, C., Cathiard, M.-A., and Loevenbruck, H. (**2006**). "Multistable syllables as enacted percepts: A source of an asymmetric bias in the verbal transformation effect," Percept. Psychophys. **68**, 458–474.

Selkirk, E. (**1982**). "The syllable," in *The Structure of Phonological Representation*, edited by H. van der Hulst and N. Smith (Foris, Dordrecht), pp. 337–383.

Sproat, R., and Fujimura, O. (**1993**). "Allophonic variation in English/l/and its implications for phonetic implementation," J. Phonetics **21**, 291–311.

Stone, M. (**1981**). "Evidence for a rhythm pattern in speech production: observations of jaw movement," J. Phonetics **9**, 109–120.

Suomi, K. (**2005**). "Temporal conspiracies for a tonal end: Segmental durations and accentual f0 movement in a quantity language," J. Phonetics **33**, 291–309.

Suomi, K., and Ylitalo, R. (**2004**). "On durational correlates of word stress in Finnish," J. Phonetics **32**, 35–65.

Tabain, M. (**2003**). "Effects of prosodic boundary on /aC/ sequences: Acoustic results," J. Acoust. Soc. Am. **113**, 516–531.

Treiman, R., and Zukowski, A. (**1990**). "Towards an understanding of English syllabification," J. Mem. Lang. **29**, 66–85.

Treiman, R., and Danis, C. (**1988**). "Syllabification of intervocalic consonants," J. Mem. Lang. **27**, 87–104.

Tuller, B., and Kelso, J. A. S. (**1991**). "The production and perception of syllable structure," J. Speech Hear. Res. **34**, 501–508.

Turk, A. (**1994**). "Articulatory phonetic clues to syllable affiliation: gestural characteristics of bilabial stops, in *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology 3*, edited by P. Keating (Cambridge University Press, New York).

Umeda, N. (**1975**). "Vowel duration in American English," J. Acoust. Soc. Am. **58**, 434–445.

Umeda, N. (**1977**). "Consonant duration in American English," J. Acoust. Soc. Am. **61**, 846–858.

van Santen, J. P. H., and Shih, C. (**2000**). "Suprasegmental and segmental timing models in Mandarin Chinese and American English," J. Acoust. Soc. Am. **107**, 1012–1026.

van Sluijter, A. M. C., and van Heuven, V. J. (**1996**). "Spectral balance as an acoustic correlate of linguistic stress," J. Acoust. Soc. Am. **100**, 2471–2485.