Origin of Consonant Duration Patterns

Melissa A. Redford
The University of Oregon

1. Introduction

Segment duration varies with syllable structure (Lehiste, 1970), resulting in word-level duration patterns. These patterns provide important phonetic cues to syllable and word boundaries in English (Christie, 1977; Boucher, 1988; Tuller & Kelso, 1991), and possibly in other languages (Maddieson, 1985). For instance, English-speaking listeners syllabify the same sequence differently depending on the relative duration of the participating segments. An obstruent-sonorant sequence is syllabified as an onset cluster when the obstruent is longer than the sonorant, but split when the consonants are similar in length (Christie, 1977). The relationship between segment duration patterns and syllable perception could indicate that syllables are units of speech production. Such a conclusion is consistent with most theories of speech production, which invoke the syllable explicitly or implicitly as a unit in speech motor programming (e.g., Stetson, 1951; Kozhevnikov & Chistovich, 1965; Lehiste, 1970; MacNeilage, 1970; Fujimura, 1981; Shattuck-Hufnagel, 1983; Fowler, 1987; Dell, Burger, Svec, 1997).

A major reason that the syllable has been chosen as an appropriate unit in speech motor programming is because suprasegmental sound patterns, such as the aforementioned consonant duration patterns, are best characterized with respect to the syllable. But the duration patterns themselves—long for consonantal onsets, short for consonantal offsets, long-short for onset clusters, and so on—are arbitrary. That is, there is nothing about the phonological syllable that can explain why an onset cluster is produced with a long-short duration pattern instead of short-long, long-long or any other kind of pattern. Since we know that speakers control duration to convey other sorts of boundary information, e.g., lexical boundaries (Turk & Shattuck-Hufnagel, 2001), it is reasonable to suggest that the duration patterns are so produced because they convey syllable structure information. But if the claim is that the same duration patterns are universally associated with the same syllable structures (Lehiste, 1970; Maddieson, 1985), then we might wonder how such arbitrary patterns have become language universal.

The seeming arbitrariness of the segment duration patterns and their universality might be explained, however, if the link between the patterns and the syllable is broken, at least in speech production. If segment duration patterns are delinked from syllables in production, then the patterns are free to be explained according to other, more basic, production constraints on the sequential articulation of segments. The patterns may then be relinked to the syllable in perception. Syllable boundaries can be extrapolated from word boundaries, which are also characterized by the same suprasegmental phonetic and phonological patterns.

For illustration of a basic production explanation for segment duration patterns, consider the long-short pattern typically associated with onset clusters. Cross-language frequency data indicate that onset clusters are usually constructed of an obstruent-sonorant sequence that precedes a vowel (Bell & Hooper, 1978). This type of sequence suggests a coarticulatory origin for the long-short duration pattern based on jaw movement. In general, obstruents are articulated with a more closed jaw than sonorants, and sonorants are articulated with a more closed jaw than vowels (Lindblom, 1983; Keating, Lindblom, Lubker, Kreiman, 1994). An obstruent-sonorant-vowel sequence is therefore articulated with the jaw moving from a closed position to an open position. The sonorants may be shortened when the segments are sequentially articulated because jaw movement is continuous in speech. Sonorants are only initiated after obstruent release, when the jaw is already in a lower position, and must be terminated when the continually downward-moving jaw makes for inefficient consonantal

 Proceedings of the 2003 Texas Linguistics Society Conference, ed. Augustine Agwuele, Willis Warren, and Sang-Hoon Park, 54-61. Somerville, MA: Cascadilla Proceedings Project.

articulation, whereupon the vowel is initiated. Such a hypothesis agrees with a theory of coproduction (Fowler, 1980; Fowler & Saltzman, 1993), and can be restated in its terms: the articulation of the internal consonant in a consonant cluster is truncated due to its coproduction with the following vowel. Similar coproduction explanations have been advanced to account for closed-syllable vowel shortening (Munhall, Fowler, Hawkins, Saltzman, 1992; Shaiman, 2002). The jaw-movement-first view advocated here adds a reason for why the segment is truncated when it is.

To summarize, segment duration patterns provide a phonetic cue to syllable boundaries (Christie, 1977; Boucher, 1988), and have been used to support the syllable as a unit of speech production (e.g., Kozhevnikov & Chistovich, 1965; Lehiste, 1970). So one explanation for segment duration patterns is that they are tied to the syllable in production, which means that the timing patterns themselves are explicitly specified in the speech plan. Such an explanation does not account for the specific patterns that are associated with different syllable structures. To explain the patterns themselves, an alternative explanation is proposed that delinks consonant duration patterns from the syllable in production, and relinks them only later in perception. In this account, consonant duration patterns are due to segment articulation during jaw movement. The timing patterns are not specified in advance, but instead emerge from the biomechanical constraint of the jaw on segment articulation. The present experiment was conducted to differentiate between these two possible explanations for the long-short duration pattern associated with onset clusters.

Specifically, we reasoned that if syllables specify the segment duration patterns, then these patterns should coincide with syllabification judgments even in languages where such judgments contradict cross-language preferences. Finnish speakers were asked to produce possible Finnish non-words with medial obstruent-sonorant sequences (e.g., *ketra*). In Finnish, these sequences are split (Berg & Niemi, 2001) and so should be produced with a long-long or short-long duration pattern. Such a pattern would be consistent with Finnish syllabification, but contra cross-language norms, which are to produce the sequence with a long-short duration pattern and syllabify it as an onset cluster to the second syllable.

If Finnish speakers produce the sequence in a manner consistent with cross-language norms, then it is likely that segment duration patterns are not tied to the syllable, but emerge from a production constraint. To test the specific production hypothesis proposed to explain the long-short duration pattern of onset clusters, where sonorants are shortened due to coarticulation with the jaw, the speakers were asked to produce the segments with a fixed jaw. The rationale was that if consonant duration patterns emerge from biomechanical factors, then the patterns should change if the biomechanics are altered.

2. Method

2.1. Design

Three adult, native-Finnish speakers produced intervocalic consonant sequences embedded in nonsense words under four speaking conditions: (1) normal speech; (2) bite-block speech; (3) clenched-jaw speech; and (4) a normal speech control. Auditory feedback was eliminated in conditions 2 and 3 to avoid perception-induced compensation. Condition 4 controlled for the effect of no auditory feedback on normal production. Conditions 1 and 4 were always completed first and last, respectively. Conditions 2 and 3 were completed in different orders by different speakers. The nonsense word tokens were produced in the same random order by the speakers and in all four speaking conditions.

Auditory feedback was eliminated in conditions 2–4 by having the speakers listen to a continuous stream of pink noise. Speakers listened to the noise over headphones, and had control over its level. They were instructed to adjust the level so that they could not hear themselves speak. This instruction appeared to be effective since the experimenter often had to indicate to the speakers during the experiment that they were speaking either too loudly or too softly.

2.2. Stimuli and Materials

The stimuli were obstruent-sonorant and, for comparison, sonorant-obstruent sequences that occurred word-internally (e.g., *ketra* and *kerta*) and, for comparison, at the edges of words (e.g., *meket rano* and *meker kano*). The stimuli were constructed so that the sonorant-obstruent sequences mirrored the obstruent-sonorant sequences, and so that mirror sequences occurred between the same vowel sounds. There 8 nonwords or nonword pairs that were repeated 3 times (x 4 stimuli types = 96 tokens total). The nonwords and nonword pairs conformed to Finnish phonology, and were vetted by a native-Finnish speaker. Since word-final consonants and word-medial obstruent-sonorant sequences are relatively rare in Finnish, only 2 obstruents, /t/ and /k/, and 2 sonorants, /r/ and /l/, were used, and /t/ only ever occurred with /r/ and /k/ only ever occurred with /l/. All nonsense words and nonsense word pairs were spoken in the frame sentence “Sano ____ taas,” which translates as “Say ____ again.”

The bite blocks were 15 millimeters (mm) sections of a hard rubber belt, which was 10 mm thick and 12 mm wide. Speakers clenched two bite blocks (one for each side of the mouth) between their premolars. If the bite blocks had been clenched closer to the front of the mouth, as in other such experiments (e.g., Fowler & Turvey, 1980; Lindblom, Lubker, Gay, 1979), they would have directly impeded the anterior consonantal articulations under study.

2.3. Recording and Acoustic Measures

The utterances were recorded in an acoustically-insulated experiment room using a Shure BG 5.1 microphone, and saved directly into a computer. They were later displayed as oscillograms and spectrograms, and the acoustic durations of the intervocalic obstruent and sonorant were measured. Obstruent boundaries were defined by a sudden drop/rise in the amplitude of a periodic waveform and by aperiodicity. Sonorant boundaries were defined by amplitude and frequency changes in the periodic waveform on one side, and by the obstruent boundary on the other. Ambiguity in defining sonorant boundaries was resolved by repeated listening to different sections of the waveform

3. Results

The aim of the experiment was to determine whether consonant duration patterns are specified by syllable structure, or emerge from more basic production processes, such as a coarticulation constraint imposed by the jaw. Overall the results favor the latter explanation. Finnish speakers produced obstruent-sonorant sequences with an onset-cluster-like long-short duration pattern, contradicting the Finnish syllabification of these sequences as coda-onset. Further, examination of segment durations across stimuli types suggests that the duration patterns emerge from passive processes.

Although the results suggest that consonant duration patterns emerge from production processes, they do not support the specific production hypothesis proposed. The specific hypothesis was that duration patterns emerge from the biomechanics of sequential segment articulation, but the duration patterns were not disrupted when the biomechanics were changed.

These overall results are presented in more detail below. The first section concentrates on the relationship between Finnish syllable structure and consonant duration patterns, and so only considers those tokens produced in the normal speaking condition (condition 1 above). The second section concentrates on the question of whether or not these duration patterns emerge naturally from the biomechanics of the system, and so considers the similarities and differences of the duration patterns across the different speaking conditions.

3.1. Consonant duration patterns in Finnish

In order to gain an overview of the duration patterns for the different sequence types, the duration of the second consonant in the sequence (C2) was subtracted from the first (C1) yielding a duration difference. A positive duration difference indicated that C1 was longer than C2, a negative difference indicated that C2 was longer than C1. This overall measure of the C1C2 duration pattern was used as a dependent variable in a univariate Analysis of Variances (ANOVA) with Speaker, Sequence Type (sonorant-obstruent versus obstruent-sonorant), and Boundary Condition (word-internal versus word-edge) as factors. The analysis showed a significant 2-way interactions with Speaker and Consonant Sequence [$F(2, 272) = 17.66$, $p < .001$], but the duration pattern only differed quantitatively, not qualitatively, across the 3 speakers.

The interaction between Sequence Type and Boundary Condition was not significant, but the simple effects were (Sequence Type [$F(1, 272) = 542.61$, $p < .001$]; Boundary Condition [$F(1, 272) = 25.64$, $p < .001$]). These results are shown in Figure 1, where it can be seen that obstruent-sonorant sequences were produced with a long-short duration pattern that was especially pronounced at the word-internal boundary, and the sonorant-obstruent sequences were produced with a short-long pattern that was especially pronounced when these sequences were divided by a word-boundary.

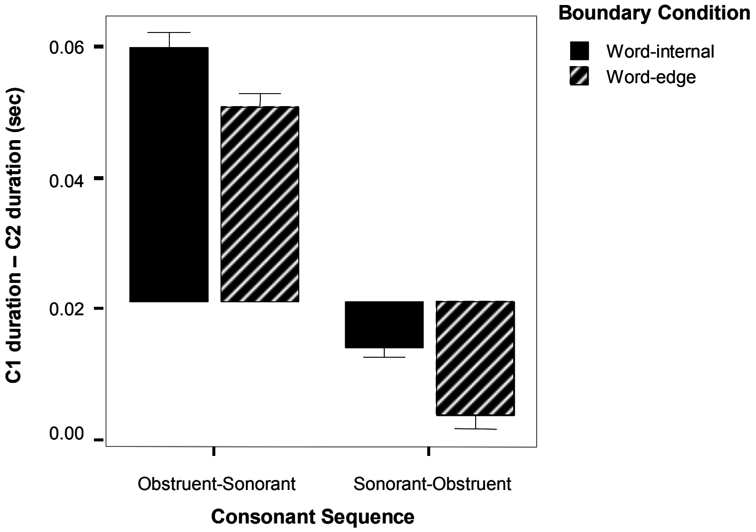

Figure 1: The duration difference between C1 and C2 is shown for the different sequence types and different boundary conditions. A positive difference corresponds to a long-short duration pattern, and a negative different to a short-long pattern.

To determine how and whether the consonant duration pattern was controlled, a second analysis was performed using consonant duration as a dependent variable in a univariate ANOVA with Speaker, Boundary Condition, Consonant Position, and Consonant Type as fixed factors. The 4-way interaction was not significant, but the 3-way interaction between Boundary Condition, Consonant Position, and Consonant Type was [$F(1, 544) = 30.67$, $p < .001$]. Figure 2 shows this interaction.

Figure 2 shows that although segment duration differs as a function of boundary condition, the effect was not consistent across segment type and position within a consonant sequence. The effect of position appears to be more systematic, at least for obstruents. Obstruents that occur after sonorants in a consonant sequence are shorter than those that occur before sonorants.

Figure 2 shows that the effect of consonant type is very clear and very systematic. Obstruents are longer than sonorants. This effect explains the long-short and short-long duration patterns of the different sequence types. In the next section, we explore whether this effect emerges from a biomechanical constraint of the jaw on segment articulation.

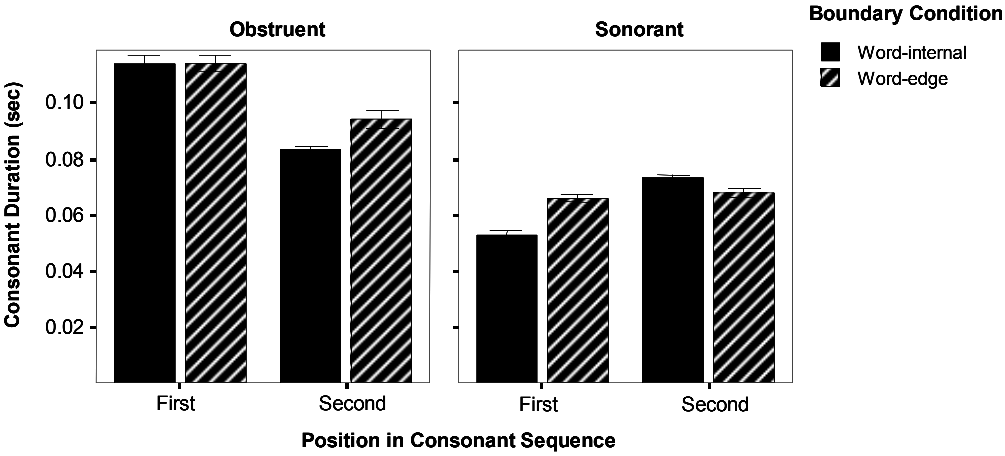

Figure 2: Obstruent and sonorant durations are shown as a function of position within a sequence and the boundary condition.

3.2. The Biomechanical Hypothesis

By hypothesis, jaw movement is the key contributor to the long-short duration pattern, and so it was predicted that fixed-jaw productions of intervocalic consonant sequences would not exhibit typical duration patterns. A 4-way (Speaker, Sequence Type, Boundary Condition, Speaking Condition) analysis of the duration patterns (captured by the C1-C2 duration difference variable) revealed a significant interaction between Speaker, Sequence Type, and Speaking Condition [$F(6, 1100) = 6.39$, $p < .001$], indicating that articulatory and perceptual manipulations had an effect on production. The interaction was otherwise not revealing in that there was no systematic effect of speaking condition on the consonant duration pattern across speakers. In general, the speakers produced obstruent-sonorant sequences with a long-short duration pattern, and sonorant-obstruent sequences with a short-long pattern. These patterns were more or less pronounced depending on the speaking condition and on the speaker (Figure 3).

Analyses of consonant durations indicated that the same 3-way interaction held for both obstruents and sonorants (obstruents [$F(6, 1096) = 2.3 < .05$]; sonorants [$F(6, 1096) = 7.03$, $p < .001$]), and there was no systematic effect of speaking condition on duration across all speakers. Such results disconfirm the specific biomechanical hypothesis proposed. The duration pattern associated with the sequences under normal speaking conditions was preserved in spite of articulatory and perceptual disruptions to the system.

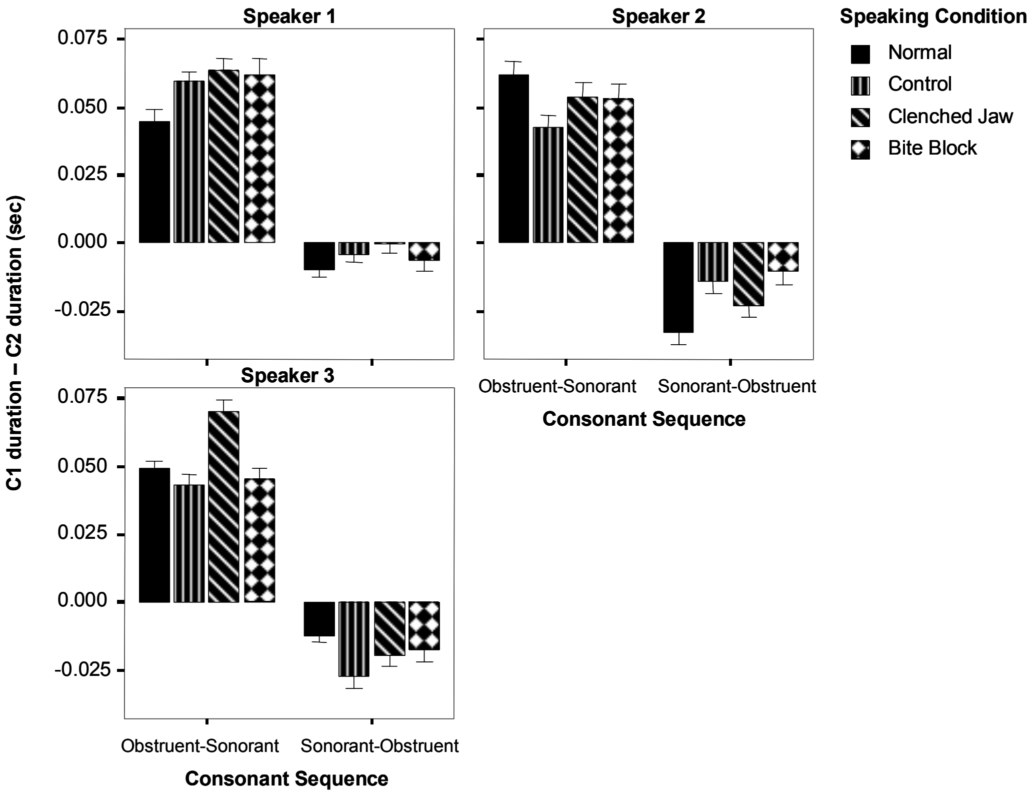

Figure 3: The consonant duration pattern is captured by a single variable (C1 duration – C2 duration) and is shown as a function of Sequence Type, Speaking Condition, and Speaker.

4. Discussion

Overall the results support the view that consonant duration patterns arise from basic production constraints, and are not tied to the syllable per se. Finnish speakers produced obstruent-sonorant sequences with duration patterns that were akin to the pattern associated with onset clusters in English and other languages, even though Finnish phonology does not allow syllables with onset clusters. The long-short pattern was even observed across word boundaries, suggesting a basic articulatory strategy that may be repressed in languages, such as English, that co-opt the duration patterns as a cue to lexical boundaries.

The basic long-short pattern of obstruent-sonorant sequences was mirrored by the short-long pattern of sonorant-obstruent sequences, typical of coda-onset syllabification. The analyses of consonant duration suggest that such patterns had their origins in a duration asymmetry: obstruents were always longer than sonorants. Future research will explore whether this asymmetry holds across different consonant types and different languages.

The origin of the duration asymmetry is unclear, and is not necessarily consistent with the biomechanical hypothesis that was proposed to account for sonorant shortening, since this hypothesis was meant only to explain sonorant shortening in onset clusters. The hypothesis focused on the contributions of the jaw to sequential segment articulation, and proposed that the duration patterns emerge from a biomechanical constraint. The preservation of the long-short and short-long patterns during bite-block and clenched-jaw speech undermined this explanation. Although it is possible that an alternative biomechanical explanation exists to account for the observed duration patterns, and that the emphasis on jaw movement was misplace, this author is unsure of how to formulate that alternative.

Some may look to the duration asymmetry results and advocate abandoning the idea that consonant duration patterns emerge from articulatory relationships between segments. These researchers may explain that the asymmetry, and hence the consonant duration patterns, are due to intrinsic duration differences between obstruents and sonorants. Such an explanation requires that the different intrinsic durations of the segments be further explored and themselves explained, which provides yet another avenue for future research.

As an alternative to having the duration patterns emerge from ill-defined intrinsic properties of segments, one might consider that the patterns emerge during the planning stages of speech production when the serial order problem is being solved. In this alternative, articulatory coordination with jaw movement may still be one of the processes conditioning consonant duration patterns, and the patterns would still be emergent, but they would emerge further upstream from the realization of the actual segments. The duration patterns would be instantiated in spite of perceptual and articulatory disruptions to the system because the coordinative structures of speech are immediately adaptable to these kinds of disruptions (Fowler & Saltzman, 1993). This hypothesis, though more resilient than the biomechanical hypothesis, is more difficult to test directly. Future work will aim to test the hypothesis by examining the sequential articulation of consonants in the developing speaker.

5. Conclusion

The study examined the role of the syllable and of jaw movement in determining syllable-related consonant duration patterns. The results clearly showed that these patterns are not tied to the syllable in production, which suggests that these patterns originate from more basic production processes. The study also explored a coarticulatory origin for the patterns, conditioned by a biomechanical constraint of jaw movement on segment articulations. The results did not support this explanation for the patterns either. Future work will therefore be devoted to following up on several interesting alternative articulatory explanations for the origin of syllable-related consonant duration patterns.

References

Bell, A., Hooper, J. 1978. Issues and evidence in syllabic phonology. In A. Bell and J. Hooper (eds.), *Syllables and Segments,* (pp. 3-22). Amsterdam: North Holland Publishing.

Berg, T. & Niemi, J. 2000. Syllabification in Finnish and German: Onset filling vs. onset maximization. *Journal of Phonetics*, 28, 187-216.

Boucher, V.J. 1988. A parameter of syllabification for VstopV and relative-timing invariance. *Journal of Phonetics*, 16, 299-326.

Christie, W. M. 1977. Some multiple cues for juncture in English. *General Linguistics*, 17, 212-222.

Dell, G., Burger, L., Svec, W. 1997. Language production and serial order: A functional analysis and model. *Psychological Review*, 104, 123-147.

Fowler, C. 1980. Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 8, 113-133.

Fowler, C. 1987. Consonant-vowel cohesiveness in speech production as revealed by initial and final consonant exchanges. *Speech Communication*, 6, 231-244.

Fowler, C., Saltzman, E. 1993. Coordination and coarticulation in speech production. *Language and Speech,* 36, 171-195.

Fowler, C., Turvey, M. (1980). Immediate compensation in bite-block speech, Phonetica, 37, 306-326.

Fujimura, O. 1981. Temporal organization of articulatory movements as a multidimensional phrasal structure. *Phonetica,* 38, 66-83.

Keating, P., Lindblom, B., Lubker, J., Kreiman, J. 1994. Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics*, 22, 407-422.

Kozhevnikov, and Chistovich, . 1965. Speech articulation and perception. *Joint Publications Research*, 30, 543.

Lehiste, I. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.

Lindblom, B. 1983. Economy of speech gestures. In P. MacNeilage (ed), *The Production of Speech* (pp. 217-246). New York: Springer-Verlag.

Lindblom, B., Lubker, J., Gay, T. (1979). Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation, *Journal of Phonetics*, 7, 147-161.

MacNeilage, P. 1970. Motor control of serial ordering in speech. *Psychological Review*, 77, 182-196.

Maddieson I. 1985. Phonetic cues to syllabification. In *Phonetic Linguistics. Essays in Honor of Peter Ladefoged* (pp. 203-221), V. Fromkin (Ed.). New York: Academic Press.

Munhall, K., Fowler, C., Hawkins, S., Saltzman, E. 1992. 'Compensatory shortening' in monosyllables of spoken English, *Journal of Phonetics*, 20, 225-239.

Shaiman, S. 2001. Kinematics of compensatory vowel shortening: The effect of speaking rate and coda composition on intra- and inter-articulatory timing. *Journal of Phonetics*, 29, 89-107.

Shattuck-Hufnagel, S. 1983. Sublexical units and suprasegmental structure in speech production planning. In P. MacNeilage (ed.), *The Production of Speech*, (109-136). New York: Springer-Verlag.

Stetson, R. (1951). *Motor Phonetics: A Study of Speech Movements in Action*. Amsterdam: North Holland Publishing.

Tuller, B. & Kelso, J.A.S. 1991. The production and perception of syllable structure. *Journal of Speech and Hearing Research*, 34, 501-508.

Turk, A. & Shattuck-Hufnagel, S. 2000. Word-boundary-related duration patterns in English. *Journal of Phonetics*, 28, 397-440.

Proceedings of the 2003 Texas Linguistics Society Conference: Coarticulation in Speech Production and Perception

edited by Augustine Agwuele, Willis Warren, and Sang-Hoon Park

Cascadilla Proceedings Project **Somerville, MA** **2004**

Copyright information

ISBN 1-57473-402-4 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Redford, Melissa A. 2004. Origin of Consonant Duration Patterns. In *Proceedings of the 2003 Texas Linguistics Society Conference*, ed. Augustine Agwuele et al., 54-61. Somerville, MA: Cascadilla Proceedings Project.

or:

Redford, Melissa A. 2004. Origin of Consonant Duration Patterns. In *Proceedings of the 2003 Texas Linguistics Society Conference*, ed. Augustine Agwuele et al., 54-61. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1067.