

Durational cues to stress, final lengthening, and the perception of rhythm

Anya Lunden*

Abstract. Binary stress languages have a well-known asymmetry between their tolerance of initial versus final lapse; the former being extremely rare and the latter being quite common. Lunden (to appear) proposes that final lengthening plays a role in this asymmetry, as the additional inherent phonetic duration of the final syllable can contribute to the continuation of a perceived rhythm, even in the absence of actual final stress. She notes this effect of final lengthening should only be available in languages that use duration as a cue to stress. However, some languages are described as having different cues to primary and secondary stress, and it is not clear which is more important for this perceptual effect. The results of four new studies show that final lengthening contributes to the perceptual rhythm of the word even when only one level of stress is cued with duration.

Keywords. stress; stress cues; stress perception; duration; rhythm; final lengthening

1. An asymmetry and an explanation. Among binary stress languages (i.e. those with roughly alternating stresses), we commonly find stress lapse tolerated at the right edge of the word, but not at the left edge (Kager 2001). It is known that word-final syllables have increased duration due to the phonetic effect of final lengthening (e.g. Oller 1973, Klatt 1976). Lunden (to appear) proposes that rhythmic prominence can come from (at least) two different sources: phonological prominence due to stress, and phonetic prominence due to final lengthening. Under this theory languages tolerate stress lapse at the right edge because an alternation in prominence nevertheless occurs.

Lunden (to appear) gives two types of evidence in support of this hypothesis. The first is a shown correlation between languages that tolerate a final stress lapse and those that use duration as a correlate to stress, among binary stress languages in Lunden and Kalivoda's (2017) stress correlate database. The second is a series of perception experiments which find that rhythmic patterns that involve an unstressed final syllable with final lengthening are confusable with fully rhythmically-alternating patterns. The methodology of these studies in general will be discussed further, and then the question at hand introduced.

The studies discussed in Lunden (to appear) all follow a similar format. Strings of a single syllable repeated five times are given different rhythmic patterns and listeners are asked to categorize them, in a forced-choice task, as either "alternating [in prominence]" or "not alternating [in prominence]." Two levels of syllable strength were used: unstressed syllables and stressed syllables. Syllable strings were made in five different varieties. Three were control cases: those that stressed the first, third, and fifth syllables (*fully alternating*, i.e. BAbaBAbaBA), those that stressed the first and third (*final lapse*, i.e. BAbaBAba), and those that stressed the the third and fifth (*initial lapse*, i.e. babaBAbaBA). Two were test cases: those that stressed the first and third syllables,

*Thanks are due to Nick Kalivoda for fruitful discussion of both the current studies and ongoing discussion of the research project as a whole. I am indebted to Marissa Messner for setting up both the ibex farm studies and the qualifying test through Amazon's Mechanical Turk. I thank the audience at LSA 2018 for their comments and feedback. Any mistakes are my own. Author: Anya Lunden, College of William & Mary (lunden@wm.edu).

and had an unstressed syllable with the characteristics of final lengthening (increased duration) in the fifth position (*final test*), and those that stressed the third and fifth syllables, and had an unstressed syllable with the characteristics of initial strengthening (increased pitch and intensity) in the first position (*initial test*). Each study contained equal numbers of alternating strings, lapse strings, and test strings (necessitating doubling the *fully alternating* strings, as the other groups have two of each type).

Different versions of the study were run, and it was found that when duration was a cue to stress, listeners frequently categorized *final test* strings (those with final lengthening) as “alternating:” 60.7 percent of *final test* strings were categorized as “alternating” (compared with 18.6 percent of control *final lapse* strings and 97.5 percent of *fully alternating* strings). On the other hand, when stressed syllables in the strings only had increased intensity and pitch but not duration, subjects were not particularly likely to categorize *final test* strings as “alternating:” 17.6 percent of *final test* strings were categorized as “alternating” (compared with 9.4 percent of control *final lapse* strings and 95.6 percent of *fully alternating* strings). While *final test* strings in this second study were statistically significantly more likely to be categorized as “alternating” than *final lapse* strings ($p < 0.001$), *final test* strings in these cases cannot be considered actually confusable with *fully alternating* strings. *Initial test* strings were included as there is also a known phonetic strengthening effect that corresponds with the left edge of the word, and it was found not to cause the same perception of rhythmic alternation that syllables at the right edge with final lengthening did.

The studies discussed by Lunden all use syllables of only two strengths: unstressed and stressed, where the stressed syllables had pitch, intensity, and duration consistent with the correlates of primary stress in English (unless modified to exclude increased duration). There are multiple languages in Lunden and Kalivoda’s (2017) stress correlate database, however, which describe somewhat different correlates of primary and secondary stress, where duration is reported to be a correlate of one but not the other. Correlates of primary versus secondary stress are not separated in the database, as the distinction is only made in a handful of sources. Maidu (a member of the Maidu family, spoken in what is now northeastern California), for example, is described as having high pitch, loudness, and tenseness in the realization of primary stress, and lower pitch, loudness, and lengthening in the realization of secondary stress (Rice 2010:225). Similarly, stress correlates in Menominee (an Algonquian language) are said to be pitch in primary stress and duration in secondary (Milligan 2005:83). Pintupi (an Australian language) is described as using loudness, pitch, and duration to cue primary stress, but only loudness for secondary stress. (Hansen and Hansen 1969:162-3).

Cases such as Maidu, Menominee, and Pintupi raise the question of whether final lengthening would perceptually contribute to the rhythm if duration were only a cue to one level of stress but not the other. A series of perceptual studies were designed to answer this question.









2. Two-stress level experiments. Four versions of the rhythmic string experiments were created using syllables that had characteristics of primary stress, secondary stress, and stressless syllables. The syllables were created with the speech synthesizer MBROLA (Dutoit et al. 1996), using American male voice us3. Previous versions of the rhythmic string experiments have been run with concatenated syllables taken from actual speech in a production study and with synthesized syllables bearing the acoustic characteristics of the recorded-speech syllables. The same levels of pitch,

intensity, and, where relevant, duration were used here, with the additional creation of secondary-stressed syllables which showed intermediate levels of the correlates. All primary-stress syllables in the study peaked at 140Hz 20 percent of the way through the syllable, secondary-stress syllables peaked at 130Hz, and unstressed at 120Hz. In the same vein, all primary-stress syllables were increased in loudness through Praat (Boersma and Weenink 2015), secondary-stress syllables were kept at the baseline level of the MBROLA output, and unstressed syllables, including unstressed syllables with final lengthening, were made softer through Praat.¹ This created three degrees of baseline prominence, which were consistent throughout all four studies. As with previous studies, stressed syllables and those with final lengthening had full vowels: [i], [u], or [ɑ], and unstressed syllables (except those with final lengthening) had the corresponding reduced vowel: [ɪ], [ʊ], or [ə]. As with the suprasegmental correlates of stress, this division was based on measurements of actual speech from a production study.² Duration-cued stressed syllables had 120ms vowels, and non-duration-cued stressed syllables, as well as unstressed syllables, had 60ms vowels.

The first two studies replicated previous experiments (although with the addition of two levels of stressed syllables), having either both primary and secondary stresses showing increased duration, or neither showing increased duration. The second two studies used the duration-enhanced stressed syllables from Study 1 for one level of stress (primary or secondary), and the non-duration-enhanced stressed syllables from Study 2 for the other level of stress.

The stress pattern used was consistent with left-aligned trochees; that is, in *fully alternating* strings, the initial syllable bore primary stress, and the third and fifth bore secondary stress. The same was true in *final lapse* and *final test* strings, except the final syllable was either a plain unstressed syllable, or an unstressed but lengthened syllable. In *initial lapse* and *initial test* strings, the third syllable bore primary stress and the final syllable bore secondary stress. Waveforms for the *fully alternating* and *final test* strings from each study are shown in (1).

(1) Waveforms from *fully alternating* and *final test* syllable strings from each study

Study	duration is	<i>fully alternating</i> waveform	<i>final test</i> waveform
1	a stress correlate		
2	not a stress correlate		
3	a correlate of \acute{o}		
4	a correlate of \grave{o}		

¹As several different vowels were used, actual intensity varied. Primary stress syllables were created by multiplying the baseline by 1.5, and unstressed syllables, including syllables with final lengthening, were multiplied by 0.5.

²The vowel quality in syllables with final-lengthening falls between [ə] and [ɑ] (see Lunden (2017)) but must be synthesized with one vowel or the other. It is important to the perception effects that the vowel in syllables with final lengthening not be completely reduced (discussed in Lunden (to appear)).

The same five types of syllable strings as previously described were used. They are given in (2) for ease of reference.

(2) Syllable string types

Type	string
<i>fully alternating</i>	óσòσò
<i>initial lapse</i>	σσóσò
<i>initial test</i>	σσóσò, with phonetic effects of initial strengthening on initial syllable
<i>final lapse</i>	óσòσσ
<i>final test</i>	óσòσσ, with phonetic effects of final lengthening on final syllable

Note that in fact both the *initial test* and the *final test* strings also include a stress lapse. But they crucially have a phonetically-enhanced edge syllable: initial strengthening with a pitch peak of 130Hz (20% into the syllable) and with a level of intensity intermediate between stressed and unstressed syllables; final lengthening with a 120ms vowel (rather than 60ms for other unstressed syllables).

Syllable strings consisted of either /bi/, /bu/, or /ba/, and equal numbers of alternating, lapse, and test strings were included (so the number of *fully alternating* strings was doubled). These three sets of six strings were repeated four times in the study, for a total of 72 strings.

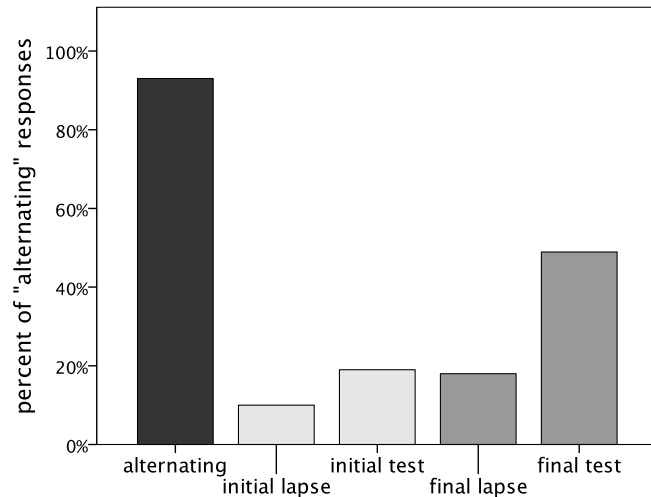
The studies were run using ibex farm (Drummond 2014) via Amazon’s Mechanical Turk. Participants in each study were unique. All participants had to pass a qualification test in which they heard alternating and non-alternating syllable strings based on the syllable ‘da.’ Subjects had to categorize ten strings as “alternating” or “not alternating.” In the qualification test they could replay each string as many times as they wanted and were told the answers to the first two as examples. A score of at least nine correct out of the ten categorizations was needed in order to access an actual study. Subjects who completed a study received \$1.00. Data was excluded from subjects who were not native English speakers. Further, subjects had to get more than two-thirds of the *fully alternating* cases in the study correct in order for their data to be used.

A logistic regression (run as a Generalized Linear Model) was run in SPSS with *response* as the dependent variable, and *string_type* and *study* as the independent variables. Post-hoc pairwise comparisons using Fisher’s least significant difference adjustment were run for the interaction term (*string_type*study*), and all significance levels reported come from these pairwise comparisons.

2.1 STUDY 1: DURATION CUE TO PRIMARY AND SECONDARY STRESS. Having duration as a correlate of both primary and secondary stress in one sense replicates previous work, but here is being carried out with levels of pitch and intensity that distinguish primary and secondary stresses. The syllables with primary stress used in this study were the same as the primary stress syllables used in Study 3, where duration was a cue to primary (but not to secondary) stress. Likewise, the syllables in this study with secondary stress were the same as the secondary stress syllables used

in Study 4, where duration was a cue to secondary (but not primary) stress. Twenty-eight subjects completed this study (aged 18–65, average age = 32, 13 females).

(3) Duration cue to both primary and secondary stress



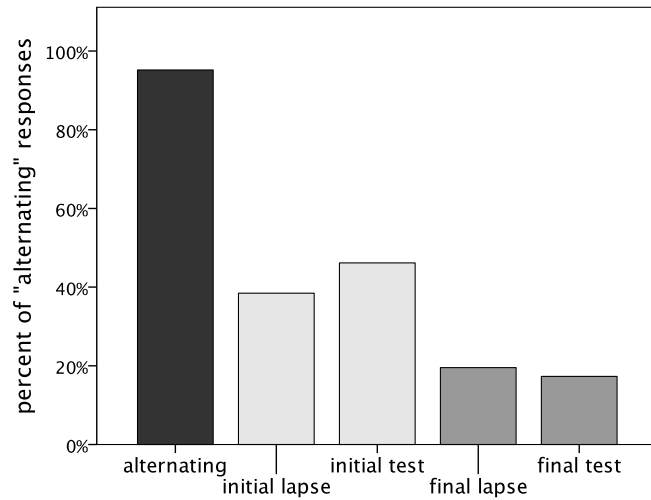
As was previously found by Lunden (to appear), a final syllable that is not stressed (i.e. has the pitch and intensity of unstressed syllables) but that has final lengthening, is sufficient to frequently cause a *final test* string to be categorized as “alternating.” While the percent of “alternat-ing” responses for the *final test* strings (48.9%) is slightly lower than what was found in previous studies, we see it is both near confusability (50%) and significantly more likely to be categorized as “alternating” than *final lapse* strings ($p < 0.001$).

There is also a marginally-significant increase in “alternating” categorizations of *initial test* strings compared with *initial lapse* strings ($p = 0.051$). But as *initial test* strings were only categorized as “alternating” in 19.0 percent of cases, they cannot be said to be confusable with *fully alternating* strings. The results of the *initial lapse* and *initial test* strings will continue to be shown for completeness/comparison, but will not be discussed.

2.2 STUDY 2: DURATION NOT A CUE TO PRIMARY OR SECONDARY STRESS. As with Study 1, this study replicates previous work except that two levels of stress, in addition to unstressed syllables, were used. The syllables with secondary stress used in this study were the same as the secondary stress syllables used in Study 3, where duration was a cue to primary (but not to secondary) stress. Likewise, the syllables in this study with primary stress were the same as the primary stress syllables used in Study 4, where duration was a cue to secondary (but not primary) stress.

In this study, both primary and secondary stress were indicated with increased pitch and intensity (to differing degrees), but neither was longer than the unstressed syllables. The vowel of the syllable with final lengthening was the only duration-enhanced syllable in this version of the task. When duration is not a cue to either primary or secondary stress, we do not expect final lengthening to contribute perceptually to the rhythm of the word, and indeed, it does not, as we see in (4). Twenty-six subjects completed this study (aged 18–59, average age = 35, 14 females).

(4) Duration cue to neither primary nor secondary stress

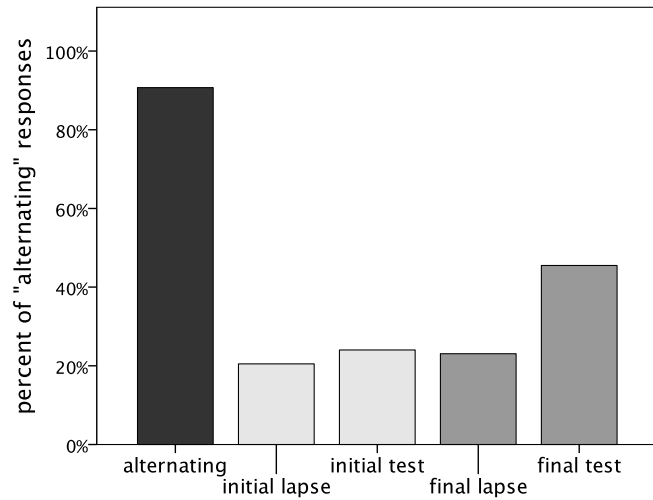


In this case, we see *final test* strings, which were categorized as “alternating” in 17.3 percent of cases, are not confusable with being truly alternating, and are no more likely to be categorized as “alternating” than *final lapse* strings are ($p = 0.470$).

The results of the first two studies replicate the findings of previous studies, indicating that final lengthening is confusable as part of the rhythm alternation of a word only if duration is a cue to stress. These two studies establish a baseline to compare with the results of the two following studies, in which duration is a correlate of one level of stress but not the other.

2.3 STUDY 3: DURATION CUE TO PRIMARY BUT NOT SECONDARY STRESS. In this study, duration was a cue to primary stress but not to secondary stress. Looking at the sample waveforms for study 3 in (1), it can be seen that since the primary stress is leftward in the stress pattern used, in the *final test* cases, the duration-based rhythm is in fact interrupted by the intervening secondary stress that is not cued with duration. Therefore this study is the stronger test case for the question of whether final lengthening can perceptually contribute to the rhythm of a word if it is a correlate of only one level of stress. In fact, we see results that are almost identical to the first study, where duration was a correlate of both primary and secondary stress, as shown in (5). Twenty-six subjects completed this study (aged 21–59, average age = 32, 13 females).

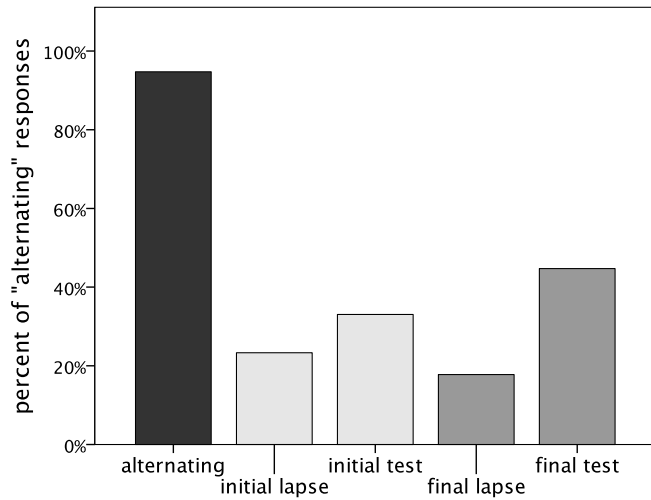
(5) Duration cue to primary, not cue to secondary stress



In this study *final test* strings were categorized as “alternating” 45.5 percent of the time, which is almost as frequently as they were in Study 1 (in which duration was a cue to both primary and secondary stress). There is not a statistically significant difference between the rates of *final lapse* strings categorized as alternating between Study 1 and Study 3 ($p = 0.444$).

2.4 STUDY 4: DURATION CUE TO SECONDARY BUT NOT PRIMARY STRESS. In the final study, duration was a cue to secondary stress, but not to primary stress. Because of the left-aligned trochaic nature of the stress pattern used, a secondary stress is the most recent stress to a syllable with final lengthening. The results of Study 3 in the preceding subsection, however, indicate that the duration-cued stress does not need to be proximal to the syllable with final lengthening for *final test* strings to be confused with truly alternating strings. And again, the results of this study are essentially identical to those of Study 1, as shown in (6). Twenty-six subjects completed this study (aged 22–65, average age = 35, 14 females).

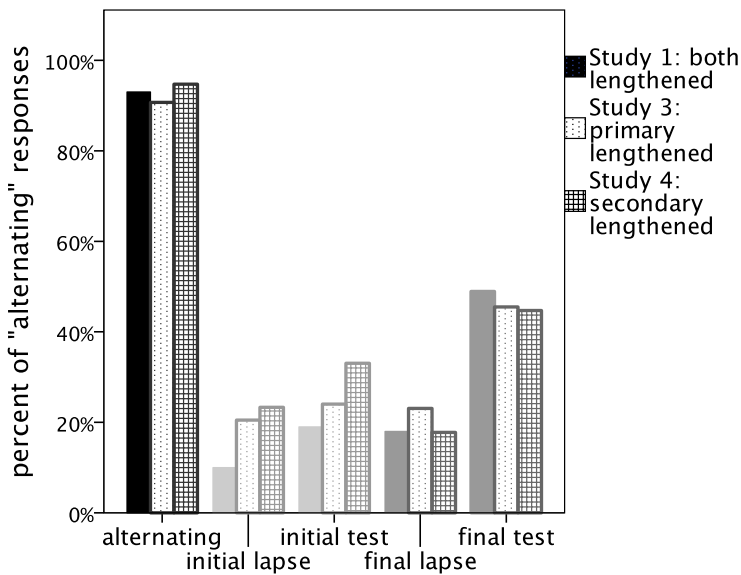
(6) Duration cue to secondary, not cue to primary stress



In this case, *final test* strings were identified as “alternating” in 44.7 percent of cases, which is not significantly different from the categorizations of *final test* strings in Study 1 ($p = 0.316$), nor is it significantly different from the categorizations of *final test* strings in Study 3 ($p = 0.837$).

3. Discussion and conclusion. It was found that the rates of categorization of the key case, *final lapse* strings, were not different between the three studies that used duration as a cue to at least some level of stress. The comparisons for all three of these studies is shown in (7).

(7) Comparison of studies in which duration was a cue to some stress



As discussed in the previous section, *final test* strings did not differ significantly between the studies. (Some of the other string types showed some significant differences, which are not discussed here.) *Final test* strings in the three studies that used duration as cue to at least some

level of stress were all statistically significantly more likely to be categorized as “alternating” than they were in Study 2, which did not use duration to cue any level of stress ($p < 0.001$).

The results from Study 3 and Study 4 show that syllables with final lengthening cause the perception of continued rhythmic alternation when duration is a cue to only one level of stress. It thus appears to be the case that as long as duration is a cue to some level of stress, final lengthening can perceptually continue the rhythm of a word. This predicts that final lapses can be perceived as continuing the stress-based prominence of the word in any language that uses duration to cue some level of stress.

The findings are especially relevant for the case of Pintupi, discussed briefly in §1, as the stress pattern of the language is left-aligned trochees, with leftmost primary stress, and duration is described to be a correlate only of primary stress by Hansen and Hansen 1969. Lunden (to appear) counts Pintupi among the languages in which the tolerance of stress lapse word-finally is correlated with the use of duration as a cue to stress, although it was an open question whether that was sufficient to cause the perceptual rhythmic effect of final lengthening that she had found in studies run with only one level of stress. On the basis of the current findings, we can conclude that it is, and that Pintupi, along with other languages which use duration as a cue to only one level stress, can reasonably be included in the explanation of the tolerance of final lapse as occurring especially, or even specifically, in those languages that use duration as a stress correlate.

References

- Boersma, Paul & David Weenink. 2015. Praat: A system for doing phonetics by computer [Computer program]. Version 5.4.22, retrieved 8 October 2015 from www.praat.org.
- Drummond, Alex. 2014. Ibex: Internet based experiments. <http://spellout.net/ibexfarm/>.
- Dutoit, T, V. Pagel, N. Pierret, F. Bataille & O. van der Vreken. 1996. The MBROLA project: Towards a set of high-quality speech synthesizers free of use for noncommercial purposes. In *Proceedings International Congress of Spoken Language Processing*, vol. 3, 1393–6. Philadelphia.
- Hansen, K.C. & L.E. Hansen. 1969. Pintupi phonology. *Oceanic Linguistics* 8. 153–70.
- Kager, René. 2001. Rhythmic directionality by positional licensing. Paper presented at HILP-5, University of Potsdam, 11 January, 2001. ROA-514, <http://roa.rutgers.edu>.
- Klatt, Dennis. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59(5). 1208–1221.
- Lunden, Anya. 2017. Duration, vowel quality, and the rhythmic pattern of English. *Laboratory Phonology* 8. 1–20.
- Lunden, Anya. to appear. Explaining word-final stress lapse. In R. Goedemans, J. Heinz & H. van der Hulst (eds.), *The study of word stress and accent: Theories, methods and data*, Cambridge: Cambridge University Press.
- Lunden, Anya & Nick Kalivoda. 2016. Stress correlate database. Online database of stress correlates: <http://wmpeople.wm.edu/sllund/stresscorrelatedatabase>.

- Milligan, Marianne. 2005. *Menominee prosodic structure*: University of Wisconsin-Madison dissertation.
- Oller, D.K. 1973. The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America* 54(5). 1235–47.
- Rice, Keren. 2010. Accent in the native languages of North America. In Harry van der Hulst, Rob Goedemans & Ellen van Zanten (eds.), *A survey of word accentual patterns in the languages of the world*, 155–248. Berlin and New York: De Gruyter Mouton.