Domain-Final Lengthening and Foot-Level Shortening in Spoken English

Brad Rakerd[a], William Sennett[a], Carol A. Fowler[b]

[a] Department of Audiology and Speech Sciences, Michigan State University, East Lansing, Mich., USA;
[b] Department of Psychology, Gerry Hall, Dartmouth College, Hanover, N.H., USA, and Haskins Laboratories, New Haven, Conn., USA

Abstract. The literature describes two kinds of durational influence on the syllables of an utterance. They are a lengthening of syllables before many syntactic boundaries ('domain-final lengthening') and a shortening of stressed syllables followed by unstressed syllables ('foot-level shortening'). In the present study we examined the relationship between these two timing phenomena. In particular, we examined the possibility that syntactic boundaries at which lengthening occurs delimit the domains over which foot-level shortening is realized. To test this hypothesis, we varied the syllabic structure of metrical feet spanning word boundaries that either coincided with a noun-phrase (NP)/verb-phrase (VP) boundary or did not. Comparison of stressed syllable durations in these conditions failed to confirm the hypothesis. Instead, unstressed syllables shortened stressed syllables by the same duration across an NP/VP boundary as within a phrase. Our findings suggest that the two effects are independent. The finding that foot-level shortening spans finally-lengthened syntactic boundaries is discussed in relation to theories of the shortening effect.

We report an unexpected outcome in experiments that explore the relationship between two durational regularities of spoken English, domain-final lengthening and foot-level shortening.

'Domain-final lengthening' [e.g., Cooper and Paccia-Cooper, 1980] refers to the tendency for words before a major syntactic boundary to be longer than the same words produced within a syntactic unit. The lengthening (often accompanied by pausing) increases with the size of the syntactic unit, suggesting articulatory sensitivity to the tiered nature of sequences of grammatically structured words and suggestive of a system that slows at the edges of performance units, syntactically defined (see note 1).

'Foot-level shortening' refers to a shortening of stressed syllables by unstressed syllables in the same metrical foot. Syllables of spoken English, and of other languages identified as stress-timed, are said to be organized metrically into feet [Abercrombie, 1964]. In English, a foot consists of a stressed syllable and zero, one, or two unstressed syllables following it. One index of

the foot structure of English is an asymmetrical shortening of stressed syllables by unstressed neighbors [Fowler, 1981]. Unstressed syllables shorten stressed syllables in the same foot (that is, stressed syllables that precede them), but not, or to a much lesser extent, stressed syllables across a foot boundary (that is, stressed syllables that follow them).

The literature offers different accounts of this organization of speech into feet. By some accounts [e.g., Classe, 1939], it reflects a stress timing tendency. Due to the shortening of stressed syllables when an unstressed syllable is added to a foot, the foot grows in duration by less than the duration of the added syllable. This may reflect an effort by talkers, for whatever reason, to maintain equal intervals between stresses. If it does, however, the effort is far from successful. Shortening by an unstressed syllable is substantially less than the duration of the syllable itself; hence a foot's duration correlates well with its number of component syllables [Huggins, 1978; Lehiste, 1972].

A different account [Fowler, 1981] has been that the shortening of stressed syllables reflects an (unexplained) asymmetrical articulatory cohesion among stressed and unstressed syllables such that stressed syllables serve as an articulatory backdrop on which production of following, but not preceding, unstressed syllables is superimposed. Consistent with this view is the finding that a stressed syllable exerts a more marked coarticulatory influence on following unstressed syllables than on preceding ones [Bell-Berti and Harris, 1976; Fowler, 1981]. Foot-level shortening would be explained as the measured shortening of a stressed syllable effected by the coarticula-

tory superimposition of following, but not preceding unstressed syllables.

Questions arise as to how domain-final lengthening and foot-level shortening originate in speech production and, in particular, how each originates in relation to the other [see also Beckman and Edwards, 1986]. From the perspective of either account of foot-level shortening – as stress timing or as articulatory cohesion – a likely possibility is that the intervals between lengthenings define the domains within which foot-level shortening occurs. That is, lengthening and pausing at syntactic boundaries will prevent an unstressed syllable across the boundary from participating in a foot on the other side. If that is so, however, the intervals within which foot-level shortening is realized must be quite limited because syntactic boundaries, such as the NP/VP and the NP/PP boundaries, which typically delimit just a few words, show lengthening [Cooper et al., 1977].

Two experiments appear to support the idea that lengthening delimits stress timing domains. In one, Huggins [1975] asked talkers to produce members of a family of sentences generated from the prototype:

The cheese(s) (a)bound(ed) (ab)out.

Across family members, the stressed syllables 'cheese' and 'bound' were followed by one or two unstressed syllables, and if they were followed by just one unstressed syllable, that syllable either was part of the same word as the target syllable or else it was across a word boundary. Huggins measured vowel durations for the stressed syllables 'cheese' and 'bound' and found that 'bound' was shortened by unstressed syllables in the same word and across a word boundary, but 'cheese' was shortened

only by the plural suffix in the same word. One difference between the two target syllables is the status of the syntactic boundary that follows them. Whereas 'bound' and its following word are in the same (verb) phrase, 'cheese' and 'bound' span a boundary between noun phrase and verb phrase. Huggins proposed that metrically governed shortening is interrupted at this major syntactic boundary.

Findings compatible with those of Huggins [1975] are reported by Cooper et al. [1977]. Their experiments were designed to examine shortening by unstressed syllables within and across a variety of syntactic boundaries. To assess shortening, Cooper et al. examined sentence pairs such as the following:

(a) The police kept *Clinton* 'til three o'clock.
(b) The police kept *Clint* until three o'clock.

Both of these sentences contain a target stressed syllable (italicized) followed by an unstressed syllable. In (a), the unstressed syllable is in the same syntactic constituent as the target and in (b) it is in a different constituent. Cooper et al. found that the target syllable was significantly shorter in productions of sentence (a) than in productions of (b). They concluded that the syntactic boundary in (b) had blocked application of the so-called 'trochaic shortening rule' that, in their view, leads to foot-level shortening. Cooper et al. found shorter durations of the target syllables in sentences of type (a) as compared to sentences of type (b) across a variety of syntactic boundaries including NP/VP, VP/NP, NP/PP, PP/PP and many others.

The goals of the present study were to replicate and extend the findings of Huggins [1975] and Cooper et al. [1977] in ex-

periments of a somewhat different design. Our efforts were motivated, in part, by Huggins' failure to replicate his own findings with added sentences and talkers [Huggins, 1974]. (In the attempt at replication, Huggins found very little shortening at all.) We also corrected certain design characteristics affecting the interpretability of the findings of Cooper et al. [1977]. One was that their experiments confounded presence and absence of a syntactic boundary with presence and absence of a word boundary. This can be seen in the sample sentences above. In sentences of type (a), unstressed syllables following the target were not only in the same syntactic phrase, they were also within the same word as the target syllable. In sentences of type (b), the unstressed syllables were both in a different phrase and in a different word. Both Huggins [1975] and Fowler [1977] have found less shortening of a stressed syllable across a word boundary than within a word; accordingly, some or all of the shortening effects observed by Cooper et al. [1977] may have been due to word boundaries, not to syntactic boundaries (see note 2). A second difficulty with their design was that it did not allow them to tell whether shortening is eliminated across a syntactic boundary (or, in their terms, whether the trochaic shortening rule is, in fact, blocked), but only whether shortening is reduced. The experiments lacked comparison conditions either within a phrase or between syntactic phrases in which the target stressed syllable was followed immediately by another stressed syllable.

In the present study, we avoided any confounding of word and syntactic boundaries, and did so in a way that allowed us to ask whether a major syntactic boundary

Table I. Test sentence families used in the experiments

MSB

1. His first *date* (a)roused some anxiety [for obvious reasons].
2. That young *duke* (dis)armed his subjects [against the advice of his counselors].
3. [I have heard that] the new *coach* (dis)trusts his players.
4. [Contrary to expectations] those new *bees* (re)acted with fury.

WPH

5. The strong *peach* (de)light was unpleasant [but better than nothing].
6. John must *bike* (a)round the block [because it is too far to walk].
7. [The coaches say] he can *fake* (un)reasonably well.
8. [As you might expect] the young group's *deep* (di)visions did disrupt them.

---

eliminates shortening. In addition, we examined a larger number of sentences and talkers than in the experiments of Huggins [1975] in order to increase the power of our analyses and, hence, to enable us to detect the small shortening effects characteristic of foot-level shortening.

Experiments 1A and 1B

In two experiments we examined the duration of a target stressed syllable in the presence of three factors that may affect its duration. Two of these factors have been discussed already; they are the syntactic and metrical context of a target syllable. A third factor, sentence length in experiment 1A and rate in experiment 1B, was included to increase the range of target word durations over which the relationship between

shortening and lengthening would be observed.

Methods

Subjects

Two different groups of 10 subjects participated in experiments 1A and 1B. All 20 subjects were undergraduates enrolled in an introductory psychology course. They were native speakers of English with normal speech and hearing according to self-report, and they were naive as to the purpose of the experiment.

Materials

Table I lists the eight test sentence families used in the experiments. Each sentence included a target stressed monosyllable, italicized in the table. A following unstressed syllable, present in some versions of the sentence and absent in others, is in parentheses. The sentence length manipulation consisted of adding or deleting the bracketed words in the table. The additional phrase in the long version of each sentence preceded the target word in half of the sentences and followed it in the remaining sentences. Materials for experiment 1B were the same as those for experiment 1A except that the bracketed material was deleted in all productions of the sentences. In experiment 1B, subjects read the short versions of the sentences at both normal and fast rates.

In sentences 1–4 in table I, there is a major syntactic boundary (MSB), an NP/VP boundary, separating the target and following words. In sentences 5–8, target and following words are within the same phrase (WPH).

We avoided the confounding of word and syntactic boundaries present in the materials used by Cooper et al. [1977], but with some sacrifice to the careful equating of phonetic properties of the target syllables and their environments achieved in that study. To eliminate the confounding, following Huggins [1975], we compared sequences having no unstressed syllable following the target syllable with sequences having an unstressed syllable across the word boundary. The cost of deconfounding was that (again, following Huggins' design) the segment on the far side of the boundary now was different

depending on whether the unstressed syllable was present or absent. This is unlikely to affect our measure of target-word duration (see below); however, it is a difference in the immediate phonetic environment of the word. Likewise, our materials were matched, but not equated, for another factor, phonetic characteristics of the target words. All target words were CVCs beginning and ending with obstruent consonants and having diphthongal (hence long) vowels (/ow, uw, ey, iy/ for MSB sentences and /ay, ey, iy/ for WPH sentences). On average, the vowels in these conditions were associated with nearly identical durations in the measurements of CVC words by Peterson and Lehiste [1960].

Procedure

In experiment 1 A, the 32 test sentences (8 base sentences, 2 lengths, 2 metrical contexts) were randomly assigned to four equal-sized blocks with two constraints: (a) that one version of each sentence appeared in each block, and (b) that there were equal numbers of long and short sentences in each block. The 16 sentences of experiment 1 B were partitioned into equal-sized blocks according to constraint (a) only. The order in which the blocks were read (and, in experiment 1 B, the rates at which they were read) was randomized across subjects.

We used the sentence reading paradigm of Cooper et al. [1977]. Subjects first familiarized themselves with a block of sentences. Next they read each sentence aloud three times with at least 5 s between readings. If a subject was dissatisfied with the naturalness of any production of a sentence, he or she was instructed to repeat it. If the experimenter heard any special emphatic or contrastive stress being used during an utterance, the subject was asked to repeat the sentence. We did not indicate which word of a sentence should receive sentential stress, and, therefore, subjects were left free to give the reading they preferred (see note 3). Subjects first produced a block of eight practice sentences. After that, they were invited to ask questions of clarification about the procedure. They then proceeded through the test blocks.

Productions were recorded on audiotape. We measured the duration of the vocalic portion of the target syllables from wideband spectrograms. For those target syllables starting with a stop consonant, the stop burst was used as the measure of onset; for targets starting with a fricative, duration was mea-
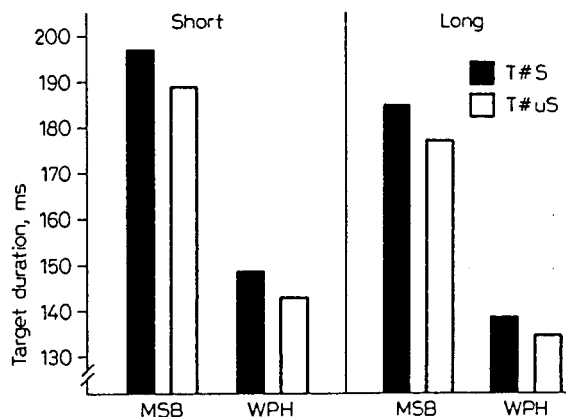

Fig. 1. Durations of target syllables in experiment 1A. Syllables were produced before a major syntactic boundary (MSB) or within a phrase (WPH), with a stressed (T#S) or unstressed (T#uS) syllable across the boundary in the short and long versions of sentences listed in table I.

sured from onset of voicing for the vowel. For target syllables ending in affricates or stops, measurement endpoints were offset of voicing and for fricatives, it was onset of frication. Both measures should approximate onset of the consonantal closure interval. Measurements were made separately by the first and third authors; any disagreements were resolved by consultation.

Results and Discussion

Sentence Length

Figure 1 summarizes the findings of the sentence length experiment. It shows that: (1) target syllables were substantially longer before a major syntactic boundary than within a syntactic phrase (by 45 ms on average); (2) they were shorter before an unstressed than a stressed syllable (by 7 ms on average), and (3) they were shorter in long than in short sentences (by 11 ms on average).
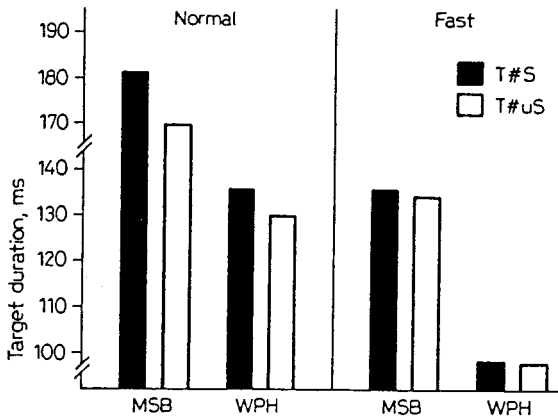
Fig. 2. Durations of target syllables in experiment 1B. Syllables were produced before a major syntactic boundary (MSB) or within a phrase (WPH), with a stressed (T#S) or unstressed (T#uS) syllable across the boundary in normal and fast productions of the short sentences listed in table I.

We performed two analyses of variance on these data with fixed factors foot structure, syntax, and sentence length. One of the analyses included subjects (F1) and the other included sentences (F2) as a random factor. In these analyses, we averaged over each talker's three tokens of each sentence type.

All of the main effects and none of the interactions were significant. The main effect of syntax (MSB vs. WPH) was highly significant in both analyses [F1(1,9) = 391.62, p < 0.001; F2(1,6) = 22.06, p = 0.004], and accounted for 68% of the total variance in the data. The main effects of foot structure [F1(1,9) = 14.84, p < 0.004; F2(1,6) = 22.14, p = 0.004] and of sentence length [F1(1,9) = 63.92, p < 0.001; F2(1,6) = 8.91, p = 0.002] were also significant, but much smaller than the effect of syntax.

The interaction of major interest was between foot structure and syntax. As figure 1 suggests, this interaction did not approach significance (both Fs < 1). Thus, shortening of a stressed target syllable by an unstressed syllable across a word boundary was not significantly less at an NP/VP boundary than elsewhere, even though the target word did show substantial domain-final lengthening at the boundary. To the contrary, there was a small but nonsignificant trend in the opposite direction; that is, a slightly greater shortening of targets occurred across the MSB (8 ms) than within a phrase (6 ms). No other interaction term was significant in either analysis (p > 0.05 in all cases).

Rate Experiment

Figure 2 presents the findings of experiment 1B. In general, target syllables were longer before NP/VP boundaries than elsewhere (34 ms on average), they were shorter followed by an unstressed than a stressed syllable (5 ms on average) and they were shorter spoken at a fast than at a normal rate (33 ms on average).

In analyses of variance with factors syntax, foot structure and rate, the main effect of syntax once again proved highly significant [F1(1,9) = 167.76, p < 0.001; F2(1,6) = 8.45, p = 0.03]. The effect of foot structure was significant in the analysis by subjects and marginal in the analysis by items [F1(1,9) = 8.48, p = 0.02; F2(1,6) = 4.60, p = 0.07]. The effect of rate was large and significant [F1(1,9) = 35.59, p = 0.0003; F2(1,6) = 131.65, p = 0.0001].

The interaction of interest, between foot structure and syntax, was once more non-significant (both Fs < 1). Talkers shortened the target syllable nearly equally across the

two syntactic boundary types. Once again, however, the small numerical difference that did emerge suggested an increase, not a diminution, of shortening across the NP/VP boundary (6 ms across the boundary, 4 ms within a phrase).

Rate proved to be a more effective durational manipulation than sentence length (33 ms for rate versus 11 ms for sentence length) and, as a consequence, target words spoken at a fast rate may have approached their compression limits. Both the interactions of syntax by rate [F1(1,9) = 13.77, p = 0.005; F2(1,6) = 7.89, p = 0.03] and of foot structure by rate [F1(1,9) = 8.35, p = 0.017; F2(1,6) = 5.27, p = 0.06] were significant or marginally significant. And in both cases, the interaction reflected the larger effects of syntax and foot structure at the slow than at the fast rate.

General Discussion

Cooper et al. [1977] found, quite generally, that a stressed syllable followed by an unstressed syllable is longer if the boundary between them coincides with a major syntactic boundary than if it does not. We repeat the sample pair of sentences from their research that we presented in our introduction:

(a) The police kept *Clinton* 'til three o'clock.
(b) The police kept *Clint* until three o'clock.

We would now ascribe the difference in duration of the target syllable, 'Clint', primarily to the word boundary that is present in (b) but not in (a) and not to the syntactic boundary. Both Huggins [1975] and Fowler [1977] found a reduction in the shortening effect of an unstressed syllable across a word boundary from a target stressed syllable as compared to the effect of an unstressed syllable in the same word as the target. Moreover, according to our findings here [see also Lyberg, 1981], the target word, 'Clint', should be lengthened in both sentences (a) and (b) due to the syntactic boundary. In experiments 1 A and 1 B of the present study, we found no additional reduction in the shortening effect of an unstressed syllable due to an NP/VP boundary despite a lengthening effect on the target due to that syntactic boundary. The whole array of findings is captured, then, by generalizations that foot-level shortening is reduced across a word boundary, but is not further reduced if the word boundary corresponds to an NP/VP boundary, even though the syntactic boundary does lengthen a stressed syllable before the break.

Our findings are surprising from the standpoint of either account of foot-level shortening that we offered in the introduction. Once a stressed syllable has been finally lengthened, the tiny shortening effect of an unstressed syllable across a word boundary does essentially nothing to preserve isochrony among feet. Nor is it likely that carryover coarticulation from the stressed vowel extends to the unstressed syllable across the break. Further research is required before any explanation for this persistence in shortening can be offered. However, the result does help to explain how a 'stress timing tendency' was ever noticed in the first place based on the perceptual impressions of phoneticians [Abercrombie, 1964; Pike, 1968]. The shortening effect of a stressed syllable spans at least the syntactic boundaries that occur both within NPs and VPs and across the boundary between them as well.

Notes

[1.] Gee and Grosjean [1983] propose that the units governing pausing and lengthening reflect 'performance structures', rather than syntactic structures directly. Performance structures correspond to metrically defined phonological and intonational phrases [Selkirk, 1978, 1980].

[2.] A finding by Cooper et al. [1977] that may appear to disconfirm this idea is that shortening was not reduced across two syntactic boundaries, an NP/NP boundary and a V/NP boundary; yet these boundaries, too, of course, coincided with a word boundary. However, the sentence pairs in these syntactic conditions are also the only pairs in the study in which the target word was followed by two, rather than one, unstressed syllables and in which the first cross-word unstressed syllable was a monosyllabic word rather than the first unstressed syllable of a following word. (The sentences were: 'Horace bought Clint an enormous turtle'; 'The longshoreman must light an inflammable carton'.) Two unstressed syllables shorten a stressed syllable more than does one [Lehiste, 1972; Fowler, 1981], and possibly, due to that, durations were near their compression limits. Moreover, in some formulations, certain monosyllabic words, including articles, are considered, metrically, to be 'de-worded' [Selkirk, 1980], and, therefore, they may not be preceded by a word boundary in their production by talkers.

[3.] A risk of this procedure is that it may result in a confounding of effects, if any, of the manipulation of metrical context with effects of variability in subjects' locations of sentential stress. That is, if it were to happen that the target word tended to receive sentential stress (and hence additional length) when it was followed by a stressed syllable, but not when it was followed by an unstressed syllable, differences in the duration of the target syllable across metrical contexts could not be ascribed unambiguously to the difference in metrical context. Rather, it might be an effect of the differential location of sentence stress. We asked whether such a confounding might be present in our data by using the following procedure. One of us (C.A.F.) listened to the utterances of a random sample of 5 of the subjects from experiment 1B and assigned sentence stress to the utterances. Because the subjects produced the four tokens of each of the 32 sentences one after the other, they tended to give the same reading for each token. Accordingly, just 32 assignments of sentence stress were made for each subject. Next, instances were examined in which the target word received sentence stress in one metrical context but not in the other. These instances were of two kinds. In one kind, the target word received sentence stress in the metrical context in which it was predicted to be long (that is, preceding a stressed syllable), but not in the context in which it was predicted to be short (that is, preceding an unstressed syllable). These instances, if sufficiently plentiful, would give spurious evidence for foot-level shortening. In the other kind, the target received sentence stress in the metrical context in which it was predicted to be short, but not in the condition in which it was predicted to be long. These instances would weaken evidence for foot-level shortening. Across the 80 pairs of sentences (eight normal-rate and eight fast-rate sentences for each subject), just four instances of the first kind of confounding were identified. Six instances of the second type were found. Accordingly, although subjects did vary their locations of sentence stress both across their own productions of a given sentence prototype (that is across metrical contexts and across rates) and among each other, just four pairs of sentences were read in such a way as to provide possibly spurious evidence of foot-level shortening. A slightly larger number of pairs worked against the hypothesis. We conclude, therefore, that the foot-level shortening we obtained in these experiments is not due to a confounding of metrical contexts and location of sentence stress.

Acknowledgments

References

Abercrombie, D.: Syllable quantity and enclitics in English; in Abercrombie, Fry, Maccarthy, Scott, Trim, in honour of Daniel Jones (Longman, London 1964).

Beckman, M.; Edwards, J.: The phonological domains of final lengthening. Meet. Linguistics Soc. Am., 1986.

Bell-Berti, F.; Harris, K.: Some aspects of coarticulation. Haskins Lab. Status Rep. Speech Res., SR 45/46, pp. 197–205 (Haskins Laboratories, New Haven 1976).

Classe, A.: The rhythm of English prose (Blackwell, Oxford 1939).

Cooper, W.; Lapointe, S.; Paccia, J.: Syntactic blocking of phonological rules in speech production. J. acoust. Soc. Am. *61:* 1314–1320 (1977).

Cooper, W.; Paccia-Cooper, J.: Syntax and speech (Harvard University, Cambridge 1980).

Fowler, C.: Timing control in speech production (Indiana University Linguistics Club, Bloomington 1977).

Fowler, C.: A relationship between coarticulation and compensatory shortening. Phonetica *38:* 21–24 (1981).

Gee, J. P.; Grosjean, F.: Performance structures: A psycholinguistic and linguistic appraisal. Cognitive Psychology, *15:* 411–458 (1983).

Huggins, A. W. F.: An effect of syntax on syllable-timing. Res. Lab. Electronics QPR-114, pp. 179–184 (1974).

Huggins, A. W. F.: Isochrony and syntax; in Fant, Tatham, Auditory analysis and synthesis of speech, pp. 455–464 (Academic Press, London 1975).

Huggins, A. W. F.: Speech timing and intelligibility; in Requin, Attention and performance, pp. 279–297 (Erlbaum, Hillsdale 1978).

Lehiste, L.: The timing of linguistic utterances and linguistic boundaries. J. acoust. Soc. Am. *51:* 2018–2024 (1972).

Lyberg, B.: Temporal properties of spoken Swedish. Monogr. Inst. Ling., No. 6 (University of Stockholm, Stockholm 1981).

Peterson, G.; Lehiste, L.: Duration of syllable nuclei in English. J. acoust. Soc. Am. *32:* 693–703 (1960).

Pike, K.: Phonemics; 11th printing (University of Michigan, Ann Arbor 1968).

Selkirk, E.: On prosodic structure and its relation to syntactic structure (Indiana University Linguistics Club, Bloomington 1978).

Selkirk, E. O.: The role of prosodic categories in English word stress. Linguistic Inquiry, *11:* 563–605 (1980).

Brad Rakerd, Ph. D.
Department of Audiology and Speed Sciences
378 Communication Arts and Sciences Building
Michigan State University
East Lansing, MI 48824–1212 (USA)