A PROSODIC ACCOUNT OF ENGLISH VOWEL LENGTHENING

Geoffrey S. Nathan
Southern Illinois University at Carbondale

ABSTRACT
The famous rule of English Vowel Lengthening may not be directly attributable to the voicing of the final consonant, but may rather be related to the rhythmic organization induced by stress-timing in English.

INTRODUCTION
One of the classic problems in English phonology/phonetics has been the rule lengthening vowels before syllable-final voiced consonants. As is well-known, at least in American English, vowels are much longer before syllable final voiced consonants than they are before syllable-final voiceless consonants. The reason that this is a problem is that the lengthening is far too much to be accounted for by the universal phonetic effect found in other languages. For example, French shows lengthening, but the difference is on the order of 10% or so, while American English lengthening may approach 100%, at least in utterance-final position. For example, the Klatt synthesizer calculates a value for /ai/ before a voiced stop as 286ms, and before a voiceless stop as 167, a ratio of 1.71:1 [1]. So the question is, why does English have this rule?

Some have suggested that this is an instance of phonologization—the exaggeration of a pre-existing tendency for phonological purposes. The problem with this proposal is that voicing lengthening is a purely allophonic, or post-lexical process. It is a typical instance of something that native speakers are not consciously aware of, but which can be brought to consciousness in a phonetics classroom. As numerous phonologists have said, rules at this level are not normally available for conscious manipulation. Consequently, it is unlikely to have been 'seized upon' by the language for exaggeration.

In addition, no other instance of a phonetically-motivated allophonic rule that I am aware of has this property—namely that a universal, speech implementation tendency is exaggerated, stretched or otherwise distorted, resulting in an allophonic rule. For example, languages normally front velar stops before front vowels: the point of articulation of the stop in 'key' is different from that of 'caw'. However, I know of no language in which this fronting has been extended to front velars to, say palato-alveolars allophonically (although subsequent language change may make alternations between velars and palatals a morphologically-conditioned rule in the language). Similarly, we find that aspiration is longer after velars than after labials, but no language makes velar aspiration longer still (or conversely, deaspirates labials). In languages that have aspiration it is generally the same length cross-linguistically. It appears to be only English vowel-lengthening that is so extreme. This leads us to wonder whether the length alternations found connected with voicing contrasts in English are in fact caused by the voicing of the following consonants at all, and are in stead due to other features of English.

BISYLLABIC SHORTENINGS
There are other principles governing vowel length in English, but they are not related to segmental factors at all, but rather deal with metrical structure. Syllable length in English is sensitive to foot type. I am suggesting here, instead, that vowel lengthening is a rhythmic phenomenon, and is somehow related to the mapping of syllables onto timing beats in speech production.

It has been argued for a long time that English is a stress-timed language (Classé, [2] is the first definitive discussion). In a stress-timed language the same amount of time is assigned to every foot, where a foot consists of a stressed syllable and optionally one or more unstressed syllables. Some have said ([3]) that, based on measurements of spoken English, this dichotomy is an auditory illusion. It goes beyond the scope of this paper, but I believe that the reason Dauer and others have been unable to find stress timing is related to their definition of 'foot', which they normally define independent of the words being measured. The data to be reported below, both that found in the literature ([4],[5]) and collected for this paper, show that the stress-timing effect is found at least for isolated words.

LENTHENING IS RHYTHMIC
Let us suppose that every stressed syllable in English is associated with at least one beat. Unstressed syllables form, with the preceding stressed syllable, a single beat as well, being roughly equivalent to a trochaic foot in contemporary Metric Phonology (see, e.g. [6]). Let us also assume that the real time length of the beat can vary depending on such extralinguistic factors as speed of speech, but that the ratio of stressed to unstressed syllables will remain relatively constant under variation for tempo and other extralinguistic factors.

Let us suppose, further, that segments are mapped onto syllables following language specific implementations of universal principles, governed overall by something very much like the traditional sonority hierarchy. Thus, a full vowel will receive a single beat, but (for English) a coda consonant will not. Thus English will differ from, say, Japanese, where coda consonants do in fact receive beats.

If we assume that beats receive roughly the same amount of time, given a similar rate of speech, there should be rough isochrony in English among one and two syllable feet. Thus stead and steady should occupy roughly the same amount of time. [4], [7] investigated this with words like stick:sticky, sleep:sleepy, speed:speedy, shade:shady.

If it were the case that every foot received an identical amount of speech time (which is what we mean by assigning a beat to each foot), then we should expect that tight, tied, tie, and tidy should each receive the same amount of time. This is however, not what we find. Specifically, the stressed vowel lengths differ, and not in the way that one would expect. The vowel length in tight and tidy are (roughly) the same, and short, while those in tie and tide are also roughly the same, but much longer than in the preceding pair. Given these facts it is surprising that the voicing of the syllable-final consonant should be posited as the cause of the differences in vowel length: This can be seen in the following chart:

Short Vowel	Long Vowel
tight	tied
tidy	tie

We can explain the difference, however, if we assume that vowel length is determined by foot structure—if every foot gets an equal measure of time, tie and tidy should receive an equal measure. Since the latter word has two syllables, each must be much shorter than any single syllable by itself. Borrowing from musical principles, if we assume that each beat is worth a quarter note, tie would be assigned a quarter note, while tidy would be assigned two eighth notes. As a consequence, the /tai/ of the former should be much longer than the /tai/ of the latter. For example, in [7], Lehiste found the following average values for 'sleep', 'sleepy', 'speed' and 'speedy':

Table 1. Mean values of the nucleus /i/ for sleep, sleepy, speed, speedy, *expressed in ms.*

sleep	sleepy	speed	speedy
180.3	131.45	297.85	163.3

But now we must ask, why is the vowel in tight so short, and the vowel in tied so long? If we continue our assumption that we are dealing with rhythmic principles here, perhaps we can make the same assumption as with the preceding pair. Suppose that there is something special about voiced consonants specifically that they are extrametrical. This is an assumption that is made about all final consonants in English nouns (see, e.g [8]) for early discussion). However, in the Metrical phonology literature the extrametricality assumption is made solely in order to place stress on the correct syllable in words like cannon, whose final syllable must be light. In these cases the extrametricality is posited solely to make stress assignment rules either regular (in some cases) or simpler (in others, such as penultimate stress). I am here assuming that extrametricality is a real rhythmic phenomenon, and that final voiced consonants do not 'count' for vowel length assignment, but that final voiceless consonants do.

The result of this set of assumptions is that, given the word tight, the entire syllable will be assigned a beat. Since final consonants take up real time, the vowel must shorten to permit the entire assemblage to occupy only a beat's worth of time. On the other hand, since final voiced consonants are extrametrical, only the vowel will be mapped onto the beat, and as a consequence, the vowels in tied and tie will be of roughly the same length.

Now, what justification, other than the fact that the results come out right, can we find for making voiced consonants extrametrical while voiceless ones are not? Overall I have no definitive answer. However, if we consider only stops, we can note that syllable final voiceless stops in English are normally accompanied by simultaneous glottal closure, while voiced stops are, of course, not. Thus it is often the case that voiced stops are released, while voiceless ones are not. It may be the case that this is somehow tied in to the timing relationships we have been discussing.

In any case, whatever the justification for assigning extrametricality to voiced obstruents, it ought to be possible to experimentally test this rhythmic account of VL in a number of ways. One would be simply to closely examine the differences between vowel allophones in final position and those before voiced vs. voiceless consonants. There are, of course, obvious cases like Canadian English where the prediction seems to be confirmed. Canadian English has an allophonic rule relating higher and lower nuclei of the diphthongs /ai/ and /au/, with the higher nucleus (normally close to [ə] occuring exclusively before syllable-final voiceless obstruents and the lower nuclei occurring elsewhere. The facts are sufficiently well known not to need rehearsal here, but it is the case that we find the raised vowels not only in classic cases such as *write* (vs. *ride*), but also in the shorter vowel contexts discussed above (such as *writer*). Exactly what happens with rider seems to be a matter of conjecture at this point, and the raising rule seems to be generalizing at this point to include not only voiceless stops but also /n/, which poses problems for any theory of phonology (including mine, incidentally) that believes that features rather than arbitrary classes of phonemes condition phonological rules. Other possible cases that bear investigation would include those dialects of English where only long allophones are diphthongs. For example, Northern Central US English seems to have monophthongal /e/ and /o/ in short contexts, with diphthongizing [ei] and [ou] only when either final or before voiced sounds.

In sum, while I have no definitive proof that word-final voiced consonants behave as if they were not located in the syllable they close, the length of the vowels preceding them indicates that they are not. As a consequence, we could also conclude that the supposedly `un'natural rule of English voicing lengthening might be somewhat more natural than was previously thought.

ACKNOWLEDGEMENT

I am indebted to David Stampe for the genesis of this idea.

REFERENCES

[1] Klatt, D.H. (1979), "Synthesis by rule of segmental durations in English sentences." *Proceedings of the 9th international congress of Phonetic Sciences.* Copenhagen.
[2] Classé, A. (1939), *The rhythm of English prose.* Oxford: Basil Blackwell
[3] Dauer, R.M. (1983). "Stress-timing and syllable-timing re-analyzed." *Journal of Phonetics* , vol. 11, pp. 51-62.
de Jong, Kenneth. 1991. An articulatory study of consonant-induced vowel duration changes in English. Phonetica 48:1-17
[4] Lehiste, I. (1971), "Temporal organization of spoken language." in *Form and Substance: Phonetic and Linguistic papers presented to Eli Fischer-Jorgensen.* L.L. Hammerich, R. Jakobson and E. Zwirner, eds. Copenhagen: Akademisk Forlag.
[5] Fowler, C.; K. Munhall, E. Salzman, and S. Hawkins. (1983), "Acoustic and articulatory evidence for consonant-vowel interactions." *Journal of the Acoustic Society of America,* vol. 80 (suppl. 1): S96

[6] Hayes, B., (1995) *Metrical stress theory* : principles and case studies Chicago : University of Chicago Press,
[7]Lehiste, I. (1972), "The timing of utterances and linguistic boundaries," *Journal of the Acoustic Society of America,* vol. 51 pp. 2018-2024.
[8] Hogg, R.M., and C.B. McCully, (1987), *Metrical Stress Theory: A coursebook.* Cambridge: Cambridge University Press.