Anticipatory lengthening

Diana Van Lancker

Neuropsychology Program, Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles.

Jody Kreiman

Phonetics Laboratory, Department of Linguistics, University of California, Los Angeles, U.S.A.

and

Dwight Bolinger

Harvard University

Received 19th October 1987 and in revised form 1st May 1988

We test the hypothesis that one stable factor in English in the length of a full syllable is the status of the vocalic nucleus (full or reduced) in the next following syllable in continuous speech: in a sequence of full plus full the first full will be longer than in a sequence of full plus reduced, regardless of other phonological, syntactic or morphological conditions.

1. Introduction

1.1. Background

A well-known effect on syllable length in English is the presence or absence of a syllable-final consonant: length increases with voicing, most markedly in pre-pausal monosyllables, e.g. rib-rid-rig versus rip-rit-rick. Is there a similar effect on the syllable in question given conditions in the following syllable?

In a series of impressionistic studies (Bolinger, 1963, 1965, 1981, 1986) and in one limited experimental study (Dasher and Bolinger, 1982), Bolinger made the claim that (1) a syllable that contains a full vowel followed by a syllable likewise containing a full vowel will be longer than (2) a syllable containing a full vowel followed by a syllable containing a reduced vowel. To abbreviate, we mark (1) (S)trong and (2) (W)eak, and we refer to a syllable containing a full vowel as a full syllable and one containing a reduced vowel as a reduced syllable.

This lengthening effect is alleged to hold, in continuous speech, regardless of word length or morpheme division. It is thus predicted that in the sentences:

(a) A workman like(S) Jack can lift a heavy weight.
(b) A workmanlike(S) jack can lift a heavy weight.
(c) A workmanlike(W) jacuzzi is a nice thing to have.

the syllable like will be longer in (a) and (b) than in (c); that in
She yelled, “What kind of sandwich do you want?,” and I yelled back,
(d) “Ham(S) on rye!”
(e) “Ham(W) and rye!”
ham will be longer in (d) than in (e); and that in
(f) Hey coroner, when you finish that au(?)topsy put your au(W)tograph on it.
those who pronounce autopsy with medial schwa will not lengthen the au- of either autopsy
or autograph, but those who use medial [a] in autopsy will make that au-
longer, contrasting with both their own autograph and the schwa pronunciation of autopsy.

1.2. The vowel series

The vowels, full and reduced, are sufficiently defined, for our purpose, by enumerating
those of the reduced series, which are fewer in number. They include the syllable
sonorants of words like battle, batten, batter, and chasm and the three contrasting final
vowels in Calley, calla, callow, of which the central member, the schwa of calla, is
the most frequent. We assume that all the reduced vowels function similarly and the full
vowels (which include diphthongs) likewise, in respect of the point we are investigating.

1.3. Simple and complex contexts

In the one situation, namely that of successive monosyllables, the length difference we
hypothesize is uncontroversial. Thus Van Draat (1910, p. 14) pointed out that in Money
makes the mare go, mare is longer than in Money makes the mare to go. Our purpose is
to test the hypothesis under more general syntactic and morphological conditions. If, as
Lehiste (1972, p. 2021) observes, “the temporal structure of the utterances seems to
depend most of all on their syllable structure”, the effect should be relatively constant
regardless of such other factors.

2. Methods

2.1. The test sentences

Two sets of sentences were devised, typed on cards, and randomized (see Appendices).
Words and phrases were chosen to provide minimal pairs, one member of which
contained a target syllable (S) and the other a target syllable (W). The aim of the first
set was to expose the target syllables to a wide variety of environments, particularly
involving accent, for example, initial accent versus initial nonaccent (e.g. the ban- of
bandit versus the ban of bandanna), initial accent versus accented monosyllable (rotate
versus wrote it), initial accent versus initial accent (mandragon versus mandolin), and final
nonaccent versus final nonaccent (electron lineup versus electron alignment). As this was
exploratory, in a few cases there were variables other than the one to be tested that might
unduly affect length. For example, in the pair
It’s (S) quite a grotesque idea.
It’s (W) a quite grotesque idea.
there is potential interference in the immediate phonetic environment: the complex
syllable structure of quite as against the simple structure of a exaggerates the strong-weak
contrast, producing a bias toward the hypothesis. On the other hand, in the pair
The ban(S)danna is green.
The ban(W)dit is mean.

the emotive content of the second member, which came across quite obviously in the speaker’s productions, was conducive to the lengthening of the already accented ban- of bandit, creating a bias against the hypothesis. Although the results suggest that these effects canceled each other out, we nevertheless designed the second test to correct them as far as possible within the limits of natural speech.

2.2. Procedure

Each stimulus set was recorded by two different speakers. All four speakers were naive as to the purpose of the experiment. The speakers read each sentence first silently, and then aloud into a dynamic microphone attached to an Ampex tape recorder in a sound-attenuated booth. They were instructed to say all test sentences as naturally as possible. Wide-band spectrograms were made of the tape-recorded utterances using a Kay digital spectrograph. Boundaries of the target syllables were marked independently by two phoneticians; for each syllable, a measurement was made to the nearest 0.5 mm (where 1 mm = 5.3 ms of speech). Measurements were rechecked twice by two different phoneticians, and any discrepancies were resolved by referring to the established measurement criteria.

Identification of speech segments from spectrograms requires an algorithm to be applied consistently throughout the measurement process. As virtually all measurements were made on minimal pairs, criteria could easily be applied equally to both test syllables. Several different approaches would serve equally well. The following criteria for establishing syllable boundaries in our data were adopted and applied consistently across all measurements:

Sentence-initially, for voiced segments (voiced consonants and vowels), beginnings of syllables were measured from the onset of voicing. (No voiceless target segments appeared sentence-initially in these stimuli.) Medially in the sentence, onsets of stop consonants were measured from the beginning of the closure, and nasals from the onset of the nasal formant. The points of terminus of test syllables were measured for stop consonants at the end of the closure/beginning of the burst (release); for all other segments, ends of syllables were found at the onset of the following segment (usually a stop).

3. Results

3.1. Collective effects

Mean durations for test syllables—(S) vs. (W)—are given in Table I (in millimeters) for the two sets of sentences spoken by the two pairs of speakers.

Two-way ANOVAs were performed on data from the two speakers for each stimulus set to determine whether the durational differences were independent of speaker. For set one, no significant main effect of speaker was observed \((F(1,59) = 0.32)\), indicating that the patterns of absolute durations were comparable for both speakers. A significant main effect of following syllable was seen \((F(1,59) = 8.83, p < 0.01)\) but no significant speaker by (S) vs. (W) interaction was observed \((F(1,59) = 0.07, \text{ ns})\). This indicates that (S) syllables were significantly longer than (W) syllables for both speakers.
D. Van Lancker et al.

Table I. Means and standard deviations of durations, comparing (S) and (W) test syllables (in mm: 1 mm = 5.3 ms)

<table>
<thead>
<tr>
<th>Stimulus set</th>
<th>(S)</th>
<th>(W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>set 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker 1</td>
<td>30.00 (9.9)</td>
<td>27.50 (8.9)</td>
</tr>
<tr>
<td>Speaker 2</td>
<td>29.2 (8.9)</td>
<td>26.5 (8.2)</td>
</tr>
<tr>
<td>set 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker 3</td>
<td>35.39 (7.09)</td>
<td>31.31 (7.39)</td>
</tr>
<tr>
<td>Speaker 4</td>
<td>39.06 (8.08)</td>
<td>37.18 (11.79)</td>
</tr>
</tbody>
</table>

3.2. Partial effects

An objection to this set of sentences was that some contrasting pairs differed in syllable count around the target syllable. For example, *The ban(S)danna is green* and *The ban(W)dit is mean* differ in that the word carrying (S) has 3 syllables but the word carrying (W) has two. To find whether this difference affected the results, we extracted from the set of stimuli recorded by Speakers 1 and 2, 15 pairs having the same number of syllables counting both target and following syllables (e.g., *mon(W)ey talks, mun(S)dane talks; I tried (W) a dozen, I tried (S) out a dozen*) and analysed them separately. Mean values were calculated for both speakers for both kinds of syllables. For speaker 1, the mean syllable duration for (S) syllables was 34.14 mm, while the mean duration for (W) syllables was 30.22 mm; for speaker 2, the contrasts were 31.38 mm for (S) and 27.22 mm for (W). An ANOVA indicated that the target syllable durations differed significantly depending on whether they were (S) or (W) (*F*(1,29) = 8.50, *p* < 0.01). The speaker and interaction effects did not achieve statistical significance.

For the second set of sentence pairs spoken by speakers 3 and 4, a two-way ANOVA showed significant main effects of speaker and of (S) vs. (W): (speaker effect: *F*(1,74) = 8.54, *p* < 0.01; (S) vs. (W): (*F*(1,74) = 15.34, *p* > 0.01). The speaker effect is due to the absolute duration differences between speakers, seen in Table I. Again, no speaker by (S) vs. (W) syllable interaction was observed (*F*(1,74) = 1.48; ns). This finding matched the results obtained on the first stimulus set.

Given the lack of significant interaction terms in both ANOVAs, data from the two stimulus sets were combined for a final ANOVA. The combined data showed significant effects of speaker (*F*(3,133) = 11.51, *p* < 0.01) and of (S) vs. (W) *F*(1,133) = 32.34, *p* < 0.01), but no interaction term (*F*(3,133) = 0.49).

3.3. Effects of word boundary and accent class

To examine our claim that the durational differences demonstrated above would hold regardless of the location of word boundaries or accent, two further ANOVAs were undertaken. The first compared pairs where both target syllables were accented with pairs where neither was accented (e.g. *They ro(S)tate both ways/They wrote(W) it both ways vs. Who can overcome a handicap (S) like that one/who can overcome a handicap(S) as bad as that one*). Mean durational values for (S) and (W) syllables in accented and unaccented position are given in Table II. An ANOVA showed a significant effect of accent location on duration (*F*(1,95) = 12.55, *p* < 0.01) and a marginally significant accent X speaker interaction effect (*F*(1,95) = 15.03, *p* < 0.01), suggesting that the
Anticipatory lengthening

Table II. Mean durations for (W) and (S) stimuli comparing accented and unaccented syllables, and between and across word boundaries

<table>
<thead>
<tr>
<th></th>
<th>Strong (S)</th>
<th>Weak (W)</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accented</td>
<td>33.42</td>
<td>33.07</td>
<td>80</td>
</tr>
<tr>
<td>Unaccented</td>
<td>29.19</td>
<td>25.10</td>
<td>23</td>
</tr>
<tr>
<td>Within word</td>
<td>33.19</td>
<td>29.69</td>
<td>46</td>
</tr>
<tr>
<td>Across word</td>
<td>34.42</td>
<td>31.84</td>
<td>73</td>
</tr>
</tbody>
</table>

effects of accent and syllable type on durations reflect separate phonetic processes. (The main speaker effect did not reach significance, nor did any other interaction.)

A second ANOVA compared sentence pairs where the target and following syllable were separated by a word boundary with pairs where no word boundary intervened (e.g. I tried (S) out a dozen/I tried (W) a dozen vs. Bring me my man(S)dragon now/Bring me my man(W) dolin now). Presence vs. absence of a word boundary had no significant effect on syllable durations ($F(1,111) = 0.02$, ns); both the speaker ($F(3,111) = 6.12$, $p < 0.01$) and (S) vs. (W) ($F(1,111) = 25.04$, $p < 0.01$) were significant (no interaction terms achieved significance). Thus our initial hypotheses are confirmed: a target syllable followed by a full syllable is longer than one followed by a reduced syllable, whether that target is accented or unaccented and regardless of word boundaries which may intervene between the two syllables.

4. Discussion

The results give statistical backing to the auditory impression created by a sequence like All(S)sam(S)pan(S)hands(S)work(S)long(S)hours—that it has a “syllable-timed” rhythm, differing from the “accent-timed rhythm” of All(W)the sam(W)ple pan(W)eling han(W)dles work(W) alike. Although the -pan of sampan is unstressed, the fact that the vowel is full gives that syllable the same status as all the others in affecting the length of the preceding syllable and being affected in its own length by the following syllable.

4.1. A function for length: quasi-rhythm rule

Many factors lead to lengthening. Unlike shortening, which responds more to overall speed, selective lengthening of a syllable is always an option, whether for emphasis, amplification (a l-o-n-g way), emotion (I h-a-t-e you!), or collecting one’s thoughts. Why then a non-selective, automatic lengthening conditioned by a following syllable? We can only speculate, but we would like to think that it is somehow functional, perhaps related to the so-called rhythm rule, which ensures that accentual “figures” will be provided with adequate “ground”. The rhythm rule is usually expressed as “avoidance of clash” (avoidance of accent on immediately successive syllables), exemplified by cases like a Taboo SUBject preferred to a taBOO SUBject or the TIMES TribUNE, a member of the ChiCAgo TRibune family of newspapers, preferred to the TIMES TRibune. Shifting the first accent leftward makes it stand out. But there is also the choice of not making the shift (especially under certain conditions of emphasis of lexical semantics), in which case another means of foregrounding the accentual figure is needed, and that plausibly comes from the extra length already provided in the (S) syllable that is to carry the
accent: the *Times* of *TIMES TRIbune* can more conveniently be given its own downglide after the peak frequency, something that is not needed in *TIMES TriBUNE*, where *Tri-* is the buffer (and not needed either in *ChicAgO TRIbune*, where *-go* is the buffer). If this is correct, we can say that the language creates a rhythmic regularity—anticipatory lengthening—to cover a potential need. It has become automatic in the interest of efficiency, whether or not actually needed in any given circumstance.

4.2. *A function for shortening: compactness*

As far as accentual figure and ground is concerned, the language would be served just as well if all syllables were full. What then is the advantage not just of syllable reduction itself but its effect on a preceding full syllable?

A possible answer lies with turning the tables and speaking not of lengthening in (S) but of shortening in (W): a reduced syllable or syllables “borrow time” from a preceding full syllable. As Lehiste (1972, p. 2018) puts it in summarizing two earlier studies:

... the words *stead*, *skid*, and *skit* were compared with *steady*, *skiddy*, and *skitty*. It might have been expected that the latter set would be longer than the former by the average duration of the derivational suffix. It turned out instead that the duration of the base part of the derived word was considerably shortened, so that even with the addition of a fairly long -y, the overall duration of the derived words was not much different from that of the base words.

The shortening does no harm to accentual structure: a contour consisting of full plus reduced syllable(s) is as effective as a contour consisting of one extended full syllable (or, for that matter, consisting of a redundantly extended full syllable followed by one or more full syllables without accent). There is even an advantage, given an established inventory of reduced vowel phonemes, in having certain syllables marked as unaccentable and therefore automatically recognizable as ground for accentual figures.

The upshot is increased compressibility of syllables and much more compact speech. Reduction becomes an ongoing process: an unaccented but full syllable no longer needs its fullness for intonational purposes, and its vowel can join the general drift to schwa.

5. *Conclusion*

Anticipatory lengthening (or its converse shortening) appears to be a constant among the determinants of syllabic duration in English, one that is often overridden but consistently reasserts itself. It is plausibly related on the one hand to intonation structure and on the other hand to compressibility, and probably represents an adjustment between the two.

Further research is needed to determine whether the length effects noted here depend entirely on the nature of the vowel, or whether the nature of the vowel is just one of several contributory factors in the “heaviness” of syllables, in which case one might expect the *Bow-* of *Bowditch* to be longer than the *Bo-* of *Bode* [bodi-] even though the second syllable in both words is reduced; and also to determine whether reduced syllables have any similar sensitivity to the nature of the syllables that follow them—a point that our examples were not set up to test.

**References**


**Appendix A**

*Stimulus sentences, test one*

They proved a com(S)plex theorem
They proved a com(W)licated theorem
The ban(S)danna fits a bandit today
The bandanna fits a ban(W)dit today
The ban(W)dit wears a bandanna today
The bandit wears a ban(S)danna today
The ban(S)danna is green
The ban(W)dit is mean
Not all prac(S)titioners are very practical today
Not all practitioners are very prac(W)tical today
Not all prac(W)tical people are practitioners today
Not all practical people are prac(S)titioners today
All thirteen(S) men are ready
All thirteen(W) marines are ready
Is Nim(S)rod as nimble as Nim O’Leary used to be?
Is Nimrod as nim(W)ble as Nim O’Leary used to be?
Is Nimrod as nimble as Nim(W) O’Leary used to be?
Is Nim(W) O’Leary as nimble as Nimrod used to be?
Is Nim O’Leary as nim(W)ble as Nimrod used to be?
Is Nim O’Leary as nimble as Nim(S)rod used to be?
They ro(S)tate both ways
They wrote(W) it both ways
It describes the electron(S) lineup
It describes the electron(W) alignment
They run up Can(W)opy Hill every day
They run up Sun(W)apee Hill every day
I tried(S) out a dozen
I tried(W) a dozen
Wait till you’ve won(S) over a few
Wait till you’ve won(W) a few
I didn’t say An(W)dy’s, I said Andes the other day
I didn’t say Andy’s, I said An(S)des the other day
I didn’t say An(S)des, I said Andy’s the other day
I didn’t say Andes, I said An(W)dy’s the other day
There were ten(S) trustees looking after the affairs of the board
There were ten(S) trusties looking after the other prisoners
There were ten(W) defendants lined up in court
One(S) Dane talks, and another listens
Mon(W)ey talks, and everyone listens
Mun(S)dane talks are boring
Mon(W)day talks are boring
Those are reprobate(S) justifications
Those are reprobate (W) excuses
Those are substantive(S) justifications
Those are substantive (W) excuses
Bring me my man(S)dragon now
Who can overcome a handicap(S) like that one?
Who can overcome a handicap(W) as bad as that one?
Who can wear a cap(S) like that one?
Who can wear a cap(W) as big as that one?
It's(S) quite a grotesque idea
It's(W) a quite grotesque idea
She earns(S) starvation wages
She earns(W) a starvation wage
"Vor(W)ta sees" is truth eternal
Vor(W)tice is a plural number
Vor(S)tex stands for one
What are your favorite mountains?—Oh, I guess the Cascades(W) are the ones I like best
What are your favorite mountains?—Oh, I guess the Cascade(S) Mountains are the ones I like best

Appendix B

Stimulus sentences, test two

No tor(S)ture is worse than that
No to(W)mato is worth all that
They live near Manitou(S) Corners
They live near Manitou(W) Corrals
They wrote(W) it both ways
They ro(S)state both ways
Because the holiday(S) calendars are here
Because the holiday(W) collection is here
Why don’t you give us a lecture on the pontiff?—I’m not INTERESTED in the pon(W)tiff any more
Why don’t you give us a lecture on the pontoon?—I’m not INTERESTED in the pon(S)toon any more
I tried(W) a dozen
I tried(S) out a dozen
Who can wear a cap(S) like that one?
Who can wear a cap(W) as big as that one?
Wait till you’ve won(W) a few
Wait till you’ve won(S) over a few
They’re staying until Sun(W)day only
They're staying until sun(S)down again
They're staying until sundown(W) again
They're staying until sun(S)down only
They're staying until sundown(S) only
Bring me my man(W)olin now
Bring me my man(S)dragon now
Mon(W)day talks are boring
Mun(S)dane talks are boring
One attack(S) after another
One attack(W) at a time
You've picked the wrong(S) attache
You've picked the wrong(W) design
You've picked the wrong(W) assistant
You've picked the wrong(S) deacon
Was it Pom(W)pey you were talking about?
Was it Pom(S)peii you were talking about?
Mackinaw(S) prime beef
Mackinaw(W) prepared beef
John(S) cracked the walnuts
John(W) corrected the papers
Do you know John(W) O'Toole?
Do you know John(S) Olsen