Phonetic Reflexes of Morphological Boundaries at a Normal Speech Rate

Mariko Sugahara & Alice Turk

Department of Theoretical and Applied Lingusitics
University of Edinburgh, UK
sugahara@ling.ed.ac.uk;turk@ling.ed.ac.uk

Abstract

Our production experiment in Scottish English revealed that the duration of a *rhyme* immediately followed by a Level II suffix such as *–s* (the 1st person singular/plural/possessive suffix) and *–t* (the past tense suffix) was significantly longer than that of a monomorphemic counterpart. Such a durational difference between suffixed forms and monomorphemic forms was absent when the Level II suffix was *–er* (the agentive suffix) or *–ing* (the progressive suffix). Those results may indicate that morphological boundaries (i.e. stem-suffix boundaries) are not directly influencing acoustic duration adjustment and support a prosody-phonetics interface hypothesis that the phonetic component is only accessible to prosodic structure (but not to morpho-syntactic structure). We also found that the *rhyme* duration of suffixed forms was shorter than the duration of a comparable *rhyme* in two-word sequence forms. This result, however, does not necessarily refute a hypothesis that there is a lexical/prosodic word boundary at the stem-suffix boundary because the presence of the word boundary at the stem-suffix boundary still allows "polysegmental" shortening to be applied to a higher word that dominates both the stem and the suffix.

1. Introduction

One of the universal aspects of speech is that speakers manipulate acoustic parameters to mark linguistic boundaries. For example, it is well-documented that word-final segments and syllables are longer when immediately followed by higher order linguistic boundaries such as sentence/utterance and phrase boundaries than when followed by no such boundaries ([5], [12], among many others). Recent work by Turk and colleagues ([9]. [10], [11]) and [3] have further shown that segmental duration adjustment is also affected by the presence or absence of a word-level constituent boundaries (i.e. pre-word boundary syllables are longer than word-medial syllables). For example, in [9], they used pairs of two target words forming a morphological compound word or a phrase such as *tune acquire* and *tuna choir*. There is a word boundary immediately after the stressed initial syllable [tun] in *tune acquire* while the same syllable in *tuna quire* is not contiguous to such a word boundary. They found the duration of the first stressed syllable [u] in *tune acquire* was significantly longer than that in *tuna choir* when the initial stressed syllable [tun] was pitch accented. It is, however, not yet clear whether the near-boundary duration adjustment is confined to word sequences or more generally to any forms with a morphological boundary. The main goal of this study is to examine the presence or absence of boundary duration adjustments in forms containing Level II suffixes in English and to ask whether there is any duration-based evidence for a prosodic boundary between the preceding stem and the following Level II suffix.

In English, affixes belong to two classes, the less productive Level I affix class (*-al, -ion, -ic*, etc) and the more productive Level II affix class (*-ing, -s, -er, -ness*, etc). [8], [7], [4], [1], among others argue based on phonological evidence that a linguistic boundary between the stem and the Level II suffix is stronger than that between the stem and the Level I suffix. More concretely, they agree that there is a word-boundary between the stem and the Level II suffix.

[6] compared the duration of a segment immediately followed by a Level II suffix (a plural and a first person singular *–s*) and that of a corresponding segment in a monomorphemic context. For example, they compared the duration of the stem-final consonant [k] in *mack-s* and that of the penultimate segment [k] in *max*, and found that the former was slightly but significantly longer (about 2% ~ 6% longer) than the latter. This result may be interpreted as phonetic evidence for a "strong" linguistic boundary between the preceding stem and the following Level II suffix. However, it is not yet clear whether the near-suffix boundary lengthening is true of all forms with other Level II suffixes such as the agentive suffix *–er* and the progressive suffix *–ing*. Also we still do not know whether the magnitude of the pre-boundary lengthening in those suffixed forms is the same as that in two-word sequences with a clear lexical/prosodic word boundary. If they are quantitatively similar, then that will be used as phonetic evidence for the presence of a word boundary between the stem and the following Level II suffix.

2. The Experiments

In our experiments, we compared the duration of (a) a rhyme immediately followed by a Level II suffix (i.e. a singular person and a possessive suffix *–s*, a past tense suffix *–t* and an agentive suffix *–er*, a progressive suffix *–ing*.) and (b) that of a corresponding rhyme within a monomorphemic word and (c) that of a rhyme immediately followed by another word. For one set of materials we found slightly longer rhyme duration in the Level II suffix boundary case than in the monomorphemic boundary case. This result was consistent with the results obtained in [6].

2.1. Reading Materials

Three sets of reading materials were used. The first set (Set 1) was created to compare forms containing a singular/plural/possessive suffix *–s* or a past tense suffix *–t* with monomorphemic words and sequences of two words. The second set (Set 2) was created to compare forms containing agentive *–er* and monomorphemic forms. The third set (Set 3) was created to compare forms containing the progressive suffix *–ing* with monomorphemic forms and sequences of two words. The list of all the reading materials

in those three sets is shown below. Each triad or pair of the target forms consisted of identical or near-identical phonetic strings which differed in the presence or absence of Level II suffixes or word boundaries. In the Level II suffix cases, there was always a stem-suffix boundary immediately after the first coda of the initial syllable. In the monomorphemic case, there was no morphological boundary between the first coda of the initial syllable and the following strings. The two-word sequences, on the other hand, had a word boundary immediately after the first coda of the first syllable of the target forms. Target forms are underlined in the following list of reading materials.

Table 1: *The list of reading materials*

	Monomorph	**Stem+Suffix**	**Word Sequence**
Set 1	***max*** *tapes*	***Mac's*** *tapes*	***Mac S****tapes*
	tax *paper*	***tacks*** *paper*	***Tack S****pider*
	tux *party*	***tucks*** *corners*	***Tuck S****parta*
	mix *targets*	***Mic's*** *targets*	***Mick S****tarling*
	lapse *rate*	***laps*** *rain*	***Lap S****lain*
	mist *rain*	***missed*** *rain*	***miss t****rain*
	past *rolls*	***passed*** *rolls*	***pass t****rolls*
	paste *rails*	***paced*** *rails*	***pace t****rails*
Set 2	***taper***	***taper*** *"tape+er"*	
	poker	***poker*** *"poke+er"*	
	minor	***miner*** *"mine+er"*	
	Eton	***eater*** *"eat+er"*	
	former	***former*** *"form+er"*	
Set 3	***Bacon*** *Force*	***baking*** *bread*	***Bake En****force*
	Cannon *Spire*	***canning*** *fish*	***Can In****spire*
	Padden *Vest*	***padding*** *trails*	***Padd In****vest*

2.2. Speakers

Six native speakers of Scottish English served as paid subjects in the experiment, five female and one male. None reported any hearing or speaking difficulties.

2.3. Elicitation and recording

Each triad or pair was presented to the speaker in a printed frame sentence(s), which was preceded by a priming sentence/phrase. The frame sentence and the priming sentence were manipulated so that speakers put a pitch accent on the initial syllable of the target forms. Placing a pitch accent on the initial syllable whose duration adjustment is the focus of the experiment is important because previous studies ([9], [3]) have shown that the pre-word boundary lengthening was most reliably observed when the syllable immediately followed by a word boundary was pitch-accented. We expect that the same is true of our stem-suffix cases if there is any pre-suffix lengthening.

The priming sentence/phrase carried information that explained the meaning of the target forms. For example, the priming phrase said "a person who mines" when the target word was a suffixed form "*miner* (*mine*+agentive *-er*). Then the priming phrase was followed by either one framed sentence (*Say ___ again/for me*) or two framed sentences (*Don't say ___ again/for me. Say ___ again/for me*). When there were two framed sentences, the target form was always in the second one and contrasted with the expression embedded in the preceding one.

Each of 6 speakers produced each stimulus 7 times (once in one recording session and there were 7 recording sessions for each speaker). There were 9 triads of reading materials in Set 1, 7 pairs in Set 2 and 5 triads in Set 3 together with additional 164 filler reading materials presented to each speaker in each recording session. Reading materials in those three sets and the fillers were randomized in each recording session and presented to each speaker on individual cards. Speakers were instructed to fully understand the intended relationship between the priming phrase/sentence and the following target forms, and they were asked to produce those prime sentences, too.

The total number of utterances relevant to this study measured was: 864 utterances from Set 1 (8 triads* 3 members of the triads * 6 sessions * 6 speakers), 360 utterances from Set 2 (5 pairs * 2 members * 6 sessions * 6 speakers), and 324 utterances from Set 3 (3 triads * 3 members * 6 sessions * 6 speakers).

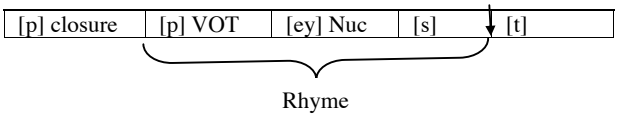
Recordings were made in a sound-proof recording studio at the Department of Theotretical & Applied Linguistics, University of Edinburgh, using a DAT tape recorder.

2.4. Segmentation and measurements

We measured the duration of the *rhyme* of the target strings embedded in the target forms. To extract the duration of those parts, we used the Praat speech analysis software (Boersma & Weenink).

The *rhyme* was the combination of the VOT region (i.e. the interval between the release of the preceding consonant and the voicing onset of the following vowel), the post-VOT vowel region and the following coda consonant region that is immediately preceded by a morphological boundary. For example, consider a triads of expressions "*paste*" ~ "*paced*" ~ "*pace t…*". The target *rhyme* of those three forms was the region that consists of the VOT of [p] and the following nucleus+coda (i.e. [es]) which is immediately. This is visually represented below.

(1) Suffix/Word Boundary



2.5. Results

We compared the duration of the *rhyme* immediately preceded by a stem-suffix boundary with that of the corresponding *rhyme* of the monomorphemic cases and of the two-word sequence cases. For those comparisons, we carried out by-subjects repeated measures analyses.

2.5.1. Set 1: –s and -t

The duration of the rhyme immediately followed by the 3rd person singular/plural/possessive siffix *–s* or the past tense suffix *–t* was significantly but only slightly longer (about 3 ms longer) than that of the corresponding monomorphemic cases ($F(1, 281) = 7.581$, $p < .01$). This result is consistent with the results obtained by [6]. In their experiment, too, the mean duration of the consonant immediately preceding the *–s* suffix was only about 3~5 ms longer than that of a monomorphemic counterpart (according to a bar graph provided in their report). The duration of the *rhyme* of the suffixed cases, however, was still significantly shorter (about 6 ms shorter) than that of the

corresponding *rhyme* of the two-word sequences ($F(1,281) = 23.304$, $p < .001$).

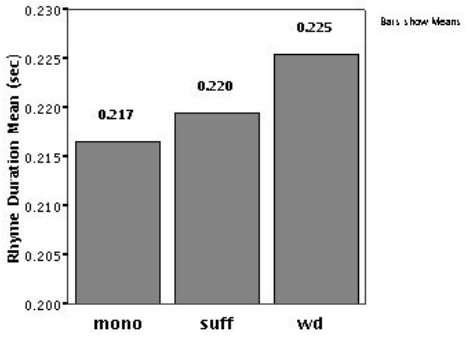


Figure 1: *The Mean duration of rhyme in the three morphological contexts (Set 1)*

One may suspect that the subtle difference between the *rhyme* duration of the suffix cases and the monomorpheme cases may not be of a structural/linguistic origin but rather of a paralinguistic effect of "frequency". It is already known that acoustic duration and the frequency/familiarity of words are inversely correlated, i.e. the more common a word is the shorter its duration is (see [2]). We, however, reject this frequency based account because the mean *rhyme* duration of suffixed words whose occurance was as frequent as or more frequent than their monomorphemic counterparts also turned out to be longer than that of the monomorphemic counterparts. A logarithmic frequency count of target words was obtained by the CELEX database. According to the database, *Mac* in the suffixed form *Mac's* was more frequent than the monomorphemic counterpart *max*; the suffixed form *tucks* and *paced* were as frequent as their monomorphemic counterparts *tux* and *paste* respectively. Nonetheless, the mean *rhyme* duration of those three suffixed forms were still 4 ms longer than that of their monomorphemic counterparts ($F(1, 101) = 4.530$, $p < .05$).

In the same way, one may argue that the difference between the *rhyme* duration of the two-word sequence cases and the suffixed cases may be attributed to paralinguistic factors. Five of the word-sequence cases were person's names and three of them were uncommon. In contrast most of the suffixed cases were either commonly used verbs or nouns. This sort of paralinguistic factors could have contributed to the durational difference. We again reject this paralinguistic account. It is because the *rhyme* duration of the word-sequences was longer than that of the suffixed cases even when both the two-word sequence cases and the suffixed cases consist of commonly used verbs or nouns only (i.e. *missed ~ miss t…, passed ~ pass t…, paced ~ pace t…*): the mean *rhyme* duration of those two-word sequence cases was about 18 ms longer than their suffixed form counterparts.

2.5.2. Set1: -er

Set 2 only consists of the suffixed forms with the agentive *–er* and their monomorphemic counterparts. There was no significant difference between the mean *rhyme* duration of those two forms ($F(1, 174) = .006$, $p > .9$), which is visually shown in Figure 2.

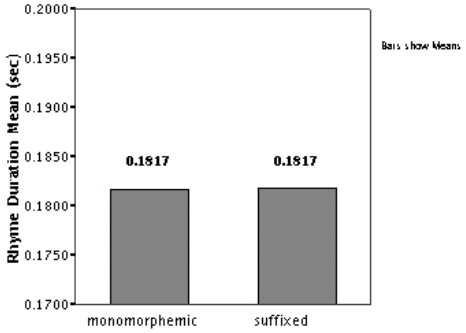


Figure 2: *The Mean duration of rhyme in the two morphological contexts (Set 2).*

2.5.3. Set 3: -ing

Set 3 consists of suffixed forms with *–ing* and monomorphemic forms as well as two-word sequence forms. The mean duration of the *rhyme* immediately followed by the suffix *–ing* was unexpectedly and significantly shorter (9 ms shorter) than that of the corresponding rhyme in the monomorphemic counterpart ($F(1, 102) = 19.431$, $p < .0001$). At the same time, the *rhyme* duration of the suffixed case was significantly shorter (20 ms shorter) than that of the corresponding *rhyme* of the two-word sequence counterpart ($F(1, 102) = 98.678$, $p < .0001$).

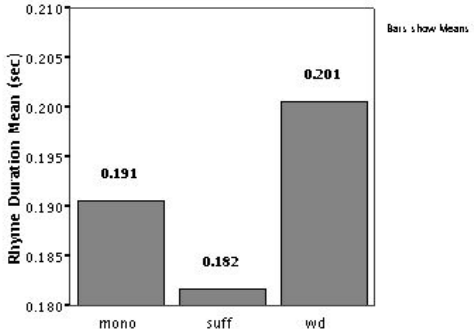


Figure 3: *The Mean duration of rhyme in the three morphological contexts (Set 3).*

The unexpected difference between the suffixed cases and the monomorphemic cases (i.e. the former was shorter than the latter) may be attributed to paralinguistic factors such as the frequency/familiarity factor. The suffixed cases were all relatively common verbs with the suffix *–ing* (i.e. *baking*, *canning*, *padding*). In contrast the monomorphemic cases were all person's first names (i.e. *Bacon*, *Cannon*, *Padden*), which are not necessarily common. It is likely that speakers produced the monomorphemic cases more carefully just because those words were less familiar, which resulted in a longer *rhyme* duration. Furthermore, the verb vs. noun difference and orthographic difference between the suffixed cases and the monomorphemic cases may be another factor to induce the unexpected difference. It is not inconceivable that nouns are psychologically more prominent than verbs and tend to be more carefully produced, which again results in a slower speech. Also in the reading material cards those monomorphemic person names were spelled with an initial

capital letter while the suffixed cases were not. It is not unlikely that speakers put more emphasis on words starting with a capital letter, which resulted in a slower speech rate.

3. Discussions

In this discussion section, we focus on (a) the inconsistent results obtained from Set 1 and Set 2 and (b) the durational difference between the suffixed cases and two-word sequence cases. (An explanation for the unexpected results of the comparison between the suffixed and the monomorphemic cases of Set 3 was already provided in Section *2.5.3*).

3.1. The inconsistent results from Set 1 and Set 2

The *rhyme* duration of the suffixed forms containing *–s/-t* in Set 1 was significantly (but only slightly) longer than that of the monomorphemic counterparts. However, the *rhyme* duration of the suffixed cases containing *-er* in Set 2 was not longer than that of their monomorphemic counterparts: the monomorphemic and the suffixed cases were of exactly the same duration.

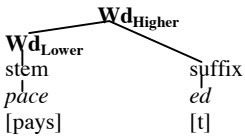
These inconsistent results show that the presence or absence of a morphological boundary does not have a direct influence on acoustic duration adjustments. Suffixed forms and monomorphemic forms in Set 1 and those in Set 2 are morphologically equivalent respectively: suffixed forms in both sets consist of a stem and a Level II suffix and monomorphemic forms in both sets contain no internal morphological boundary. In contrast, prosodic organizations of the members of those two sets could be quite different given that the suffixes *–s/-t* in Set1 and the suffix *–er* in Set 2 are different in their segmental configurations: the former consist of a single obstruent segment while the latter contains a syllabic sonorous segment. And we consider the difference in their prosodic organizations to be responsible for the inconsistent results between Set 1 and Set 2. Because of a space limit, we are not able to provide details of our prosodic account for the different results between Set1 and Set 2 in this paper. Please refer to our future work for more details.

3.2. The durational difference between the suffixed and the two-word sequence forms

In Section *2.5.1*, we have shown that the *rhyme* duration of the suffixed cases was shorter than that of the two-word sequence cases. This result is, however, still consistent with a hypothesis that there is a lexical/prosodic word boundary between the preceding stem and the following Level II suffix *–s/-t*.

Even if there is a Word boundary at the right edge of a stem, the stem and the following suffix constitute another higher Word as shown in the nested word structure in (2).

(2)



Given that the comparable Word in the two-word sequence counterparts has one segment less than the higher Word of the suffixed cases, e.g. *paced* [peys**t**] vs. *pace* [pays], it is not inconceivable that the higher Word of the suffixed cases as a whole undergoes a "polysegmental shortening", which results in shorter duration of the target *rhyme* duration. In this way, the presence of a word boundary at the stem-suffix boundary does not prevent the target *rhyme* of the suffixed cases from being shorter than that of the two-word sequence counterparts.

3.3. The future study

In our future study, we will compare the duration of suffixed and monomorphemic cases read in a slower speech rate. We assume following the observation made by [3] that a slower speech enhances durational contrast between different morphological/prosodic structures. If the subtle difference between the suffixed cases and the monomorphemic cases found in the current study is really induced by a structural difference, then we expect more consistent and larger durational difference to emerge under the slower speech rate.

4. Acknowledgement

This work is supported by a research grant from the Arts and Humanities Research Board (AHRB) to the second author.

5. References

[1] Aronoff, M. & Sridhar, S. N., 1983. Morphological Leels in English and Kannada; or Atarizing Reagan. *Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax,* Chicago Linguistic Society, 3-16.

[2] Aylett, M. & Turk, A, to appear. The Smooth Signal Redundancy Hypothesis: Relationships between Redundancy Prosodic Prominence and Syllable Duration in Spontaneous Speech, *Language & Speech*.

[3] Beckman, M. E. & Edwards, J., 1990. Lengthening and Shortening of Prosodic Constituency. In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech,* J. Kingston & M. E. Beckman (eds.). Cambridge: Cambridge University Press.

[4] Borowsky, T., 1993. On the Word Level, In *Phonetics & Phonology: Studies in Lexical Phonology*, S Hargus & E. M. Kaisse (eds.). Academic Press.

[5] Cooper, W. E. & Paccia-Cooper, J., *Syntax and Speech*. Cambridge, MA: Harvard University Press.

[6] Schwarzlose, R. & Bradlow, A. 2001. What happens to segement durations at the end of a word? iJournal of the Acoustical Society of America, 109(5, pt.2), 2292.

[7] Selkirk, E. 1982, *The Syntax of Words*. Cambridge, MA: MIT Press.

[8] Siegel, D. 1974, *Topics in English Morphology*. Doctoral dissertation, MIT.

[9] Turk, A. & Shattuck-Hufnagel, S., 2000. Word-Boundary-Related Duration Patterns in Enlish. *Journal of Phonetics,* 28. 397-440.

[10] Turk, A. & Sawusch, J., 1997. The Domain of Accentual Lengthening in American English, *Journal of Phonetics,* 25, 25-41.

[11] Turk, A. & White, L. 1999. Structural Influences on Accentual Lengthening in English, *Journal of Phonetics*, 27(2), 171-206

[12] Whiteman, C. W., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P.J, 1992. Segmental Durations in the Vicinity of Prosodic Phrase Boundaries, *Journal of the Acoustical Society of America,* 90(3), 1707-1717.