The Effects of Polysyllabic Shortening and Word/Phrase Boundary on Duration Patterns of English

Bingru Chen

The Chinese University of Hong Kong
Vivian_Chan@link.cuhk.edu.hk

ABSTRACT

This study is intended to examine the duration patterns of English by Mandarin speakers in comparison with native British speakers. 11 native British English speakers and 20 Mandarin speakers of English were asked to produce 4 sets of tokens consisting of a mono-syllabic base form, disyllabic, and trisyllabic words derived from the base by the addition of suffixes, and a set of short sentences with a particular combination of phrase size, stress pattern, and boundary location. The results suggest that the amount of polysyllabic shortening and the effect of word or phrase position are likely to affect duration patterns of Mandarin speakers of English, leading to their Chinese accent of English. This study can benefit both L2 leaners and language teachers by increasing their sensitivity to the duration differences and difficulties experienced by L2 learners of English.

Keywords: duration patterns, polysyllabic shortening, word/phrase boundary, Chinese accent

1. INTRODUCTION

Foreign accents are often perceived in the speech of Chinese speakers who speak English as an L2 (ESL) (e.g., [7], [14]). Many scholars have examined the factors affecting overall degree of perceived accent. However, most previous studies on the phonological features of Chinese-accent English focus on the segmental features (e.g. [7]); the suprasegmental elements receive much less attention. As is stated by Flege [8], articulatory habits and phonological knowledge of the first language (L1) are often applied by L2 learners when they speak a foreign language. According to [14], cross-linguistic variation in duration presents a potential problem for L2 learners attempting to learn the sound system of a new language. Although certain research (e.g., [3], [17]) shows that L2 productions may be affected by L1 duration patterns, little work has been done, especially for Chinese ESL speakers (e.g., [4], [17]).

English and Mandarin Chinese are different in terms of duration patterns. Languages have been traditionally classified as stress-timed (e.g., English, German), syllable-timed (e.g., Chinese), and mora-timed (e.g., Japanese) ([1], [10]). This classification is based upon the concept of isochrony, meaning that there are units of equal duration within the speech signals. However, no clear evidence for such isochrony has been found ([6], [13]). Wang et al., [16] compared the relationship between segment and syllable duration for English and Mandarin. They found that English segments are not compressible for the sake of equal syllable duration. Mandarin Chinese, on the other hand, is perceived to be syllable-timed. Although this distinction may remain debatable, the rhythmic pattern of English is nonetheless quite different from that of Chinese.

In the process of exploring duration factors in speech, Lehiste [11] discovered that the durational structure in English is conditioned by the number of syllables: as the number of syllables in a word or phrase increases, the duration of all individual syllables shortens. This effect is referred to as polysyllabic shortening and has been tested in a number of studies (e.g., [2], [15], [18]).

According to [21], polysyllabic shortening is equally relevant in Mandarin, but the shortening effect is greater. In Mandarin, the final syllable compresses considerably when changing from a monosyllabic word to disyllabic word pattern (Figs. 3c, 4a in [21]), which is different from English (Figs. 4 and 5 in [12]). Moreover, in Mandarin, syllables that are newly added in the medial position are considerably more compressed than initial ones when words are changing from disyllabic to tri- and quadra-syllabic words (Figs. 4a in [21]), whereas in English this considerable difference is much more subtle: medial syllables are only slightly shorter than initial syllables (also Figs. 4 and 5 in [12]).

Klatt [9] stated that a word pronounced independently from others has a relatively identical duration as if it is at the final position of an utterance. This phenomenon is described as a pre-pause elongation. The details include word-final lengthening and phrase-final lengthening. The duration of a word-final syllable is longer than that of their counterparts found in different word positions ([2], [15], [18]). The effect of phrase-final lengthening was stated by Klatt [9] that eloquent pauses take place in sentences. Such pauses are often found at the closing of an embedded section or prior to a prepositional word which does not alter the word

positioned before it. Pre-pause elongation is also noticed in such cases. This type of lengthening also takes place within phrase and clause boundaries when physical intervals are absent from the acoustic signal.

According to [19], the duration patterns in Mandarin are influenced by the effect of final lengthening as well. It was reported that a disyllabic word has a short-long duration pattern regardless of whether it is utterance-initial or utterance-final. Similar patterns can also be found in Chen [5]. Xu and Wang [21] also found that there is a shortening of word- or phrase-medial syllables in Mandarin that is virtually absent from English.

In this study, we aim to examine the duration patterns of English by Mandarin ESL learners. A comparative study is further conducted with native British English speakers, expecting to see whether the accented pronunciation of Chinese ESL speakers of English, if any, is dependent on their native duration patterns.

2. METHOD

2.1. Subjects

The overall experiments had a between-subjects design. One group consists of 11 native British English speakers (7 females, 4 males) who were all from the Greater London area and self-declared to speak Southern British English, Received Pronunciation, or BBC English. Subjects in this group were average 33 years old. The other group includes 20 native Mandarin speakers (17 females, 4 males) who were all from mainland China and spoke standard Mandarin. They were average 24 years old, and have been learning English for average 14.5 years. Their English proficiency was overall band score in IELTS 6.5 or above (speaking test above 6.0). The subjects had 0-3 years (average 1.5 years) of experience living in the United Kingdom and had no experience of living in other English-speaking countries. All subjects were recorded under standard conditions in a sound-treated booth by using an Audio Technica AT2020 USB Microphone. All participants were recruited from the city of London and paid for their participation.

2.2. Recording Procedure

Subjects were told to read stimuli in a constant speaking rate and loudness. The stimuli were presented on the screen of a notebook computer via a java script in a random order, and a different order was used for each subject. All stimuli repeated 3 times. If the subjects made a mistake, they were instructed to repeat the word or sentence.

2.3. Measurements

The acoustic sound of each word was analyzed by using ProsodyPro, a Pratt script for prosody analysis [20]. Segmentation and labeling for each experiment were done in the TextGrid window. The onset and offset of each target sequence was manually labeled. The segmentation of the acoustic signal was carried out by reference to spectrographic evidence of the transition between consonants and vowels. The overall data was analyzed by two-way ANOVAs.

2.4. Experiments

2.4.1. Experiment 1

This experiment represented a replication and extension of [11]. The stimuli consisted of 4 sets of words (16 tokens in total), built around the base words *stick*, *sleep*, *shade*, and *speed*. Three derivational suffixes are added to the base words separately, one of them (-y) monosyllabic and two (-ily, -iness) disyllabic.

2.4.2. Experiment 2

This experiment used three sets of phrases in carrier sentences. The stimuli consisted of two-word adjective-noun phrases which were all 3 syllables in size. It is known that a particular combination of phrase size, stress pattern, and boundary location specified a phrase paradigm. Only the typical stress level "01+1# " and "The + adjective + noun+ verb phrase" phrase type was used in this study. The digits 0 and 1 denote no stress and primary stress. The + symbol and # symbol demote a word boundary and a phrase boundary, respectively. The three sets of phrases were put in a standard syntactic context (early position in the sentence, e.g. in 1a, 2a, and 3a), and then in a nonstandard syntactic context (later position in the sentence, e.g. in 1b, 2b, and 3b).

(1a) The *remote stream* was perfect for fishing.
(1b) Fishing in the *remote stream* was perfect.
(2a) The *absurd day* made many ideas seem strange.
(2b) Many ideas on this *absurd day* seemed strange.
(3a) The *confused girl* found the boy full of secrets.
(3b) The boy was found by the *confused girl* full of secrets.

2.5. Results

2.5.1. Result of Experiment 1

There is a clear correlation between the duration of the base and the number of syllables for both groups of speakers. That is, as the number of syllables in one

word increases, the durations of base or the durations of the same segments in the derived word decreases, as is graphically shown in Figure 1. Similar duration patterns are also seen in the syllable nuclei (vowel), as displayed in Figure 2, which is true of both British speakers and Mandarin ESL speakers. The present data confirms Lehiste's hypothesis [11] that durational structure is conditioned by the number of syllables in a word for English. A similar tendency can also be found for Mandarin ESL speakers when they read the same materials.

Figure 1: Mean durations with standard deviations of base and of the same segments in the derived words for the words *stick, sleep, shade* and *speed* by British English speakers and Mandarin ESL speakers.

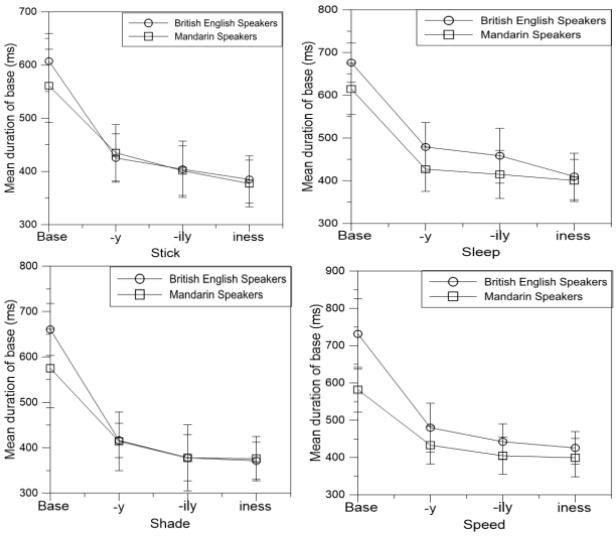
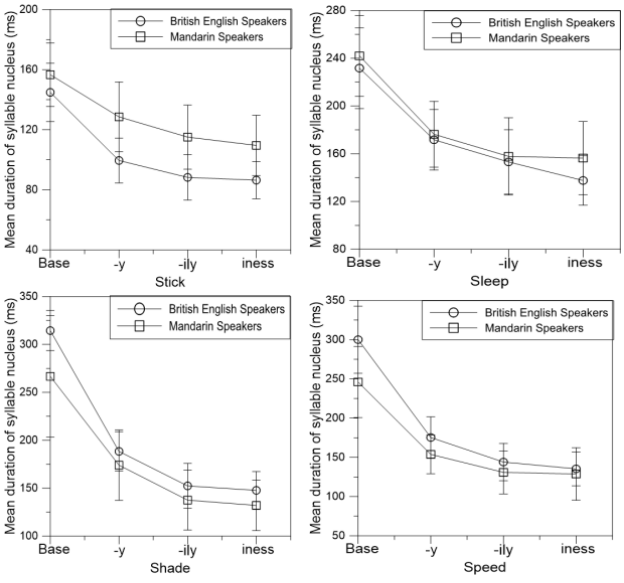


Figure 2: Mean durations with standard deviations of syllable nuclei for the words *stick, sleep, shade* and *speed* and their derivative words by British English speakers and Mandarin ESL speakers.

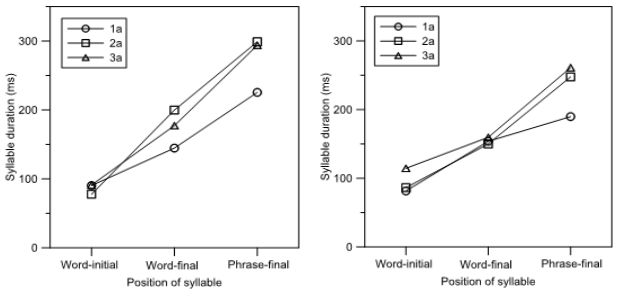


A two-way ANOVA was conducted and the result shows that there is a significant difference in the mean B/D ratio (the ratio of the duration of the base word to the duration of the same segments occurring in the derived word) of each set of words between two groups ($p<0.01$), except for the set of *sleep*. Besides, British English speakers show greater polysyllabic shortening in the base form of derived words as they have greater B/D ratio than Mandarin ESL speakers for each set of words. Significant differences can also be found in the mean B/D ratio within each set of words when three suffixes (-y, -ily, -iness) were added to the base words ($p<0.01$).

2.5.2. Result of Experiment 2

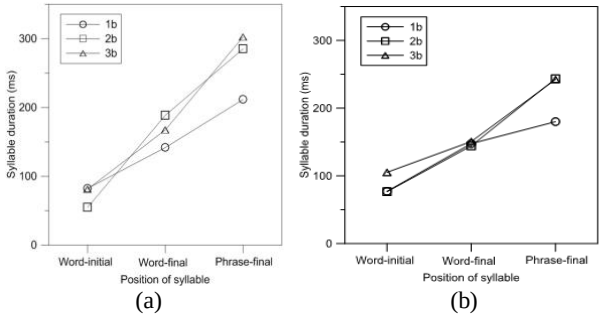
For both groups, word-final syllables are longer than word-initial syllables; phrase-final syllables are longer than word-final syllables; and monosyllabic words behave like word-final syllables when they are in word-final position, and like phrase-final syllables when they are in the phrase-final position. These are true of both standard context and nonstandard context, as shown in Figure 3 and 4.

Figure 3: Duration of syllables as a function of their positions within words and phrases (1a, 2a, 3a) in standard syntactic context for British English speakers (a) and Mandarin ESL speakers (b)



(a) (b)

Figure 4: Duration of syllables as a function of their positions within words and phrases (1b, 2b, 3b) in nonstandard syntactic context for British English speakers (a) and Mandarin speakers of English (b)

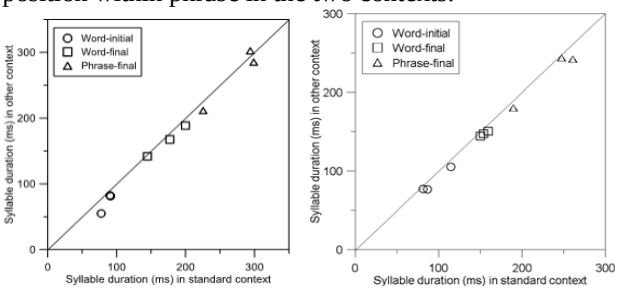


(a) (b)

The effect of syntax on syllable duration was also investigated by varying the syntactic contexts into standard syntactic context. Syllables with the same

positions in words and phrases have the same stress level "01+1#". If equivalent syllables have the same duration in both standard and nonstandard contexts, then the data points are likely to fall on the diagonal line. Figure 5 displays that equivalent syllables are consistently longer in standard context than in nonstandard context, especially for phrase-final syllables. This result shows that there is an apparent effect of syntactic context for syllable durations for both groups.

Figure 5: Duration of syllables in standard syntactic context versus nonstandard syntactic contexts (labelled "other" on the ordinate) for British English speakers (a) and Mandarin ESL speakers (b). Each point represents two syllables equated for stress-level, position within word, and position within phrase in the two contexts.



(a) (b)

A two-way ANOVA was conducted to show that under standard context, there is a significant interaction between the effect of subject groups and location in words and phrases ($p=0.01$) on the mean duration of syllables. Simple main effect analysis shows that there is no statistically significant difference between two groups when producing word-initial syllables ($p=0.39$). However, when these two groups produce word-final syllables ($p=0.03$) and phrase-final syllables ($p<0.01$), significant difference can be found. Duration of both word-final syllables and phrase-final syllables produced by British English speakers (mean duration 173.97ms and 272.82ms) is longer than that of Mandarin ESL speakers (mean duration 154.53ms and 232.94ms).

On the other hand, a significant interaction between the effect of subject groups and location in words and phrases on the mean duration of syllables ($p<0.01$) can also be found under nonstandard context. No significant difference exists between two groups for the mean duration of word-initial syllables ($p=0.09$), but it does exist between them for the mean duration of word-final syllables ($p=0.02$) and phrase-final syllables ($p<0.01$). British English speakers produce longer duration for syllables in word-final position (mean duration 165.94ms) than Mandarin ESL speakers (mean duration 147.58ms); and British English speakers also produce longer duration for syllables in phrase-final position (mean duration 266.50ms versus mean duration 222.03ms).

3. DISCUSSION AND CONCLUSION

The goal of the present study is trying to see if we can find any Mandarin-specific durational characteristics that are carried to Mandarin ESL speakers, leading to their foreign accent when they speak English.

First, it confirms Lehiste's hypothesis [11] that durational pattern is conditioned by the number of syllables. However, although the similar tendency occurs to both groups, British English speakers show greater reduction in the base form of derived words, and there is a significant difference between two groups in the degree of that reduction. Therefore, the amount of polysyllabic shortening is likely to be one of the duration-related factors leading to the foreign accent of Mandarin ESL speakers. But it is interesting that the result contradicts what has been predicted based on the finding of [21] that Mandarin has more substantial polysyllabic shortening. In terms of the reason of such a phenomenon, further investigation is worthy to be conducted.

Second, duration of syllables is dependent on the position of the syllable in a word and phrase. British English speakers show a consistently increasing tendency for duration of syllables from word-initial position to phrase-final position. This finding confirms the hypothesis of [12]. Similar tendency can also be found in Mandarin ESL speakers. In both standard and nonstandard contexts, two groups are significantly different in the mean duration of word-final and phrase-final syllables. That is, British English speakers produce longer duration for syllables in both word-final position and phrase-final position.

Thus, the amount of polysyllabic shortening and the effect of position in words and phrase on syllable duration are likely to be two duration-related factors affecting the accentedness of English for Mandarin L2 speakers. Even though the vowel types and syllable structures involved in this study seem not comprehensive enough, we do hope that this study can offer some preliminary findings for later research on duration patterns of languages. Mastering a good understanding of the effect of polysyllabic shortening and position in words or phrases may be also helpful for L2 teachers to set priorities when they teach pronunciation to L2 learners.

4. ACKNOWLEDGEMENTS

The author would like to express her gratitude to Professor Yi Xu for his supervision and offering valuable suggestions.

5. REFERENCES

[1] Abercrombie, D. 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

[2] Beckman, M. E., Edwards, J. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In J. Kingston, & M. E. Beckman (Eds.), Papers in *Laboratory Phonology* (pp. 152–178). Cambridge: Cambridge University Press.

[3] Bent, T., Bradlow, A. R., Smith, B. L. 2008. Production and perception of temporal patterns in native and non-native speech. *Phonetica*, 65(3), 131-147.

[4] Chen, H. C. 2015. Acoustic analyses and intelligibility assessments of timing patterns among Chinese English learners with different dialect backgrounds. *Journal of Psycholinguistic Research*, 44(6), 1-25.

[5] Chen, M.Y. 2006. Durational adjustment under contrastive focus in Standard Chinese. *Journal of Phonetics*, 34, 176-201.

[6] Dauer, R. M. 1983. Stress-timing and syllable-timing reanalyzed, *Journal of Phonetics*, 11, 51–62.

[7] Deterding, D. 2006. The pronunciation of English by speakers from China. *English World-Wide*, 27, 175–98.

[8] Flege, J. E. 1984. The detection of French accent by American listeners, *Journal of the Acoustical Society of America*, 76, 692–707.

[9] Klatt, D. H. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America*, 59(5), 1208–1221.

[10] Ladefoged, P., Johnson, K. 2011. *A course in phonetics*. Wadsworth, Ohio: Cengage Learning.

[11] Lehiste, I. 1972. The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America*, 51, 2018.

[12] Nakatani, L. H., Connor, K. D., Aston C. H. 1981. Prosodic Aspects of American English Speech Rhythm. *Phonetica*, 38, 84-105.

[13] Roach, P. 1982. On the distinction between 'stress-timed' and 'syllable-timed' languages, *Linguistic Controversies*, 73-79.

[14] Tajima, K., Port, R., Dalby, J. 1997. Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics*, 25(1), 1-24.

[15] Turk, A. E., Shattuck-Hufnagel, S. 2000. Word-boundary-related duration patterns in English. *Journal of Phonetics*, 28(4), 397–440.

[16] Wang, C., Zhang, J., Xu, Y. 2018. Compressibility of Segment Duration in English and Chinese. *Proceeding of Speech Prosody 2018*.

[17] White, L., Mattys, S. L. 2007a. Calibrating rhythm: first language and second language studies. *Journal of Phonetics*, 35(4), 501-522.

[18] White, L., Turk, A. E. 2010. English words on the procrustean bed: polysyllabic shortening reconsidered. *Journal of Phonetics*, 38(3), 459-471.

[19] Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, 27, 55-105.

[20] Xu, Y. 2013. ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis, 7–10.

[21] Xu, Y., Wang, M. 2009. Organizing syllables into groups — evidence from f0 and duration patterns in mandarin. *Journal of Phonetics*, 37(4), 502-520.