International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:13, No:3, 2019

Duration Patterns of English by Native British Speakers and Mandarin ESL Speakers

Chen Bingru

***Abstract*—**This study is intended to describe and analyze the effects of polysyllabic shortening and word or phrase boundary on the duration patterns of spoken utterances by Mandarin learners of English in comparison with native speakers of English. To investigate the relative contribution of these effects, two production experiments were conducted. The study included 11 native British English speakers and 20 Mandarin learners of English who were asked to produce four 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 duration of words and segments was measured, and results from the data analysis suggest that the amount of polysyllabic shortening and the effect of word or phrase position are likely to affect a Chinese accent for Mandarin ESL speakers. This study sheds light on research on the duration patterns of language by demonstrating the effect of duration-related factors on the foreign accent of Mandarin ESL speakers. It can also benefit both L2 learners and language teachers by increasing their sensitivity to the duration differences and difficulties experienced by L2 learners of English. An understanding of the amount of polysyllabic shortening and the effect of position in words and phrase on syllable duration can also facilitate L2 teachers to establish priorities for teaching pronunciation to ESL learners.

***Keywords*—**Duration patterns, Chinese accent, Mandarin ESL speakers, polysyllabic shortening.

I. Introduction

ONE definition of foreign accent is "the overall impression concerning native speakers form whether or not and to what degree a person sounds native or non-native" [23]. There is a wide phenomenon that foreign accent is often perceived in the speech of Second Language (L2) learners, which has been investigated in a large number of experimental studies. It is also true of Chinese speakers who learn English as an L2 language [10]. Many scholars have tried to reveal the factors expected to affect overall degree of perceived accent for Chinese speakers who learn English as an L2 (ESL), however, most previous studies on the phonological features of Chinese-accent English focused on the segmental features (e.g. [4], [7]); the suprasegmental parts of English with Chinese accent received much less attention. Several research findings suggest that listeners' impression and perception on foreign-accented speech are potentially influenced by the duration patterns which include speaking rates, pausing, stress timing, intonation, etc. Reference [1] found that English speakers have less tolerance of suprasegmental deviance than of segmental errors when assessing L2 learners' accented speech. Reference [28] further confirmed that non-native timing patterns produced by Chinese ESL learners had a significantly negative impact on their intelligibility. The two languages, English and Chinese, are different in terms of duration patterns. It is generally assumed that all languages belong to three rhythm classes: stress-timed, syllable-timed, and mora-timed, and in each case, the named units are equal in duration [31], [12]. The empirical basis of the rhythm class hypothesis has been investigated extensively, however, acoustic evidence of such isochrony has been lacking. Even so, English is typically claimed to be stress-timed, and Chinese, on the other hand, tends to be syllable-timed [12].

According to [9], when the language patterns of L1 and L2 are identical, learning could take place easily through positive transfer of the L1 pattern, but when they are different, learning would be difficult and errors may arise as the result of negative transfer or interference. Based on the influence of L1 on L2 learning, and the above different timing patterns among English and Chinese, it is reasonable to assume that the accuracy of English pronounced by Chinese speakers depends, at least to some extent, on their L1 duration patterns.

Chen Bingru is with the Department of English, The Chinese University of Hong Kong, HK, China (e-mail: Vivian_Chan@link.cuhk.edu.hk).

II. Literature Review

On the basis of many previous studies, factors that determine the duration patterns are basically categorized from seven levels by [17]. The comparison of durational patterns of English by native English speakers and Chinese ESL speakers in the present study is mainly based on the following two factors.

A. Effect of the Number of Syllables

In the process of exploring temporal factors in speech, many researchers observed that the number of syllables of a word is an important factor in determining syllable duration [2], [18], [22]. When the number of segments increases, the duration of any given segments displays a tendency to become shorter, until a surmised minimum duration materializes; after this point segments can no longer be constricted [3], [11], [18], [15], [30]. While tracing the effect of speaking mode on vowel duration, it is discovered that the duration of a vowel in a monosyllabic word is longer than in the case of two, three or four syllable words; also, the vowel duration in a disyllabic word is longer than in a three or four-syllable word [13]. Such an effect is called polysyllabic shortening.

According to [33], polysyllabic shortening is equally relevant in Mandarin. They found that the duration of a syllable cluster is going to rise when the number of syllables in the group increases. However, these two uphill steps are not

commensurate– if the length of the phrase rises, the duration of the separate parts decreases.

Dissimilarities between English and Mandarin are also easy to be found. Compared to English, the shortening effect in Mandarin is greater. In Mandarin, the final syllable compresses considerably when changing from a monosyllabic word to disyllabic word pattern [33], which is different from English [25]. 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, whereas in English this considerable difference is much more subtle [25]. Additionally, [25] reported that there is a linear correlation in English, that is, as the duration of segments in a word goes up, there is also a corresponding increasing tendency for the duration of syllables, however, a less linear correlation exists in Mandarin.

B. Effect of Level of Stress on a Syllable

Reference [27] have pointed to the relationship between the duration of a syllable and its level of stress. The stress level of an utterance influences the segmental duration; also the stressed vowels are longer than their unstressed counterparts. When they compared words with opposing lexical characteristics - "insert" (a verb) and "insert" (a noun) – the result was that the duration of stressed vowels exceeds that of the non-stressed ones.

Based on another two studies, unstressed vowels have a shorter duration than stressed vowels. The phrase-final syllable shows the greatest difference, with the unstressed unreduced vowel being only two thirds of the length of the stressed vowel. Reducing an unstressed vowel to schwa further reduces the length of a vowel [20], [26].

Reference [16] found that the class of stressed vowels only has a subtle difference (5%) between the primary and secondary lexical stress carried by the vowels (e.g. the fourth and first syllables in the word 'applicability'). The length of a word may be further affected by varying the types of the phrasal stress patterns. For example, the durational pattern created by the stress pattern of the word "fifteen" is likely to vary in utterances of 'ten fifteen' and 'fifteen men' [31]. This example indicated that there may be phrase level stress alteration founded on the rhythmic constraints that, for example, encourage alternations between unstressed and stressed syllables.

It is often known that Chinese can be defined as a tone language with four lexical tones whereby each syllable has a tone of equal length. Within Chinese dialects, Mandarin contains neutral tones which occur at the ending of a word or phrase with pronunciation in a light and short manner [5]. Reference [21] has stated that neutralization of tone in weaker syllables such as in Mandarin is caused by shorter durations of phonetics or decreased intensity. The neutral tone takes the place of the final syllable causing the syllable to lose its own tone. The surface pitch pattern is then caused by the preceding full lexical tone. As a result of this characteristic, it is considered similar to an unstressed syllable. Such syllables are typically described as having a 'fifth tone' or 'neutral tone'. It is of vital importance to state that in Mandarin, except for the neutral tone, there are no equivalents to word stress in English as most words in Mandarin are not distinguished by stress [6], [8]; hence, shortenings are not a result of varying degrees in stress.

C. Presence of Word or Phrase Boundary Following a Syllable

Reference [17] stated that a word pronounced independently from others has a relatively identical duration as if it was at the final position of an utterance and probably at least twice as long as if it were at the initial point of a sentence. This phenomenon was described as a pre-pause elongation. The details can be seen as follows: word-final lengthening and clause- and phrase-final lengthening.

The duration of word-final syllables is longer than that of their counterparts found in different position in a word. The duration of a consonant is determined also by the position within a word, and it is greater when it is placed in the starting point and is minimal when placed in the medial position [14], [26], [29]. Consonants found at the end of words are greater in length if they also occupy the final position of the phrase.

Based on previous studies, reference [25] did additional experiments to confirm the significance of the above-mentioned findings and added some details to the framework of how they contribute to the syllable duration patterns. It is suggested that syllable duration is powerfully dominated by stress, final positions in words and phrases, yet barely influenced by non-final positions. The recorded duration of a word increases linearly with the size of the word and varies with stress pattern for words of constant proportion. For instance, duration of syllables in the final positions of a phrase and a word is consistently longer. Other positions show non-significant influence.

According to [33], the duration patterns in the Mandarin language are influenced by word-initial and word-final lengthening as well. They found that in 3-syllable and 4-syllable combinations of speech, the last syllable is the longest and the initial syllable is the second longest. On the other hand, in multi-syllabic words or phrase, the final syllable is the longest and the initial one is the second longest.

Reference [15] stated that eloquent pauses take place in sentences, for example between words which are syntactically unrelated. 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.

Reference [22] has inspected a great number of statistics for the Swedish language where elongation was found at the final points of linguistic combinations which include each sentence, phrase, and individual words. Comparable results are recognized in English where, according to [24], components tend to be elongated during spontaneous utterances right before significant grammatical component boundaries. Reference [16] analyzed the duration of all components in a connected

communication, as delivered by one speaker. He determined which components were longer than 1.4 times the medial duration for a particular component and found that, with the exception of one, all elongated components were located in a phrase-final position. Elongation was discovered at the final position of noun compounds, combined or fixed examples, also at the boundary of a noun and a verb phrase. The result showed a 30% rise in vowel duration compared with the phrase corpus.

Based on previous observations concerning duration patterns of English and Chinese, the present study thus tries to examine the duration patterns of English by Chinese ESL learners. Also, a comparative study is further conducted between Chinese ESL learners and native British English speakers, expecting to see whether the inaccurate pronunciations of Chinese ESL speakers, if any, are dependent on the different timing patterns of them. Two timing variables—syllable duration and word/ phrase boundary—will be acoustically measured and compared. This study seeks to address the following research questions:

(1) Is the duration of the base form in derived words shortened to a significant different degree when another syllable is added, for both native speakers and Mandarin ESL speakers?
(2) Is syllable duration strongly influenced by positions in phrases and words for both native speakers and Mandarin ESL speakers?

III. Method

A. Stimuli

Two types of stimuli were used: words and sentences. The test materials will be presented in the descriptions of experiment 1 and experiment 2.

B. Subjects

The overall experiments had a between-subjects design, with two subject groups divided by their L1. One group consists of 11 native British English speakers (7 females, 4 males) who are all from the Greater London area and are self-declared to speak Southern British English, Received Pronunciation, or BBC English according to their answers in a background questionnaire. Subjects in this group were 19-53 years old (average 33 years old). The other group includes 20 native Mandarin speakers (17 females, 4 males) who are all from mainland China and speak standard Mandarin. They were 20-27 years old (average 24 years old), and have been learning English for 10-20 years (average 14.5 years). Their English proficiency is 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 of the 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. None of them reported having any speech disorders.

C. Recording Procedure

Subjects were given 2-3 minutes to get familiar with the test materials. During the recording, subjects were told to maintain constant speaking rate and loudness. The test words were presented on the screen of a notebook computer in a random order, and a different order was used for each subject. All words appeared three times, thus a total of 48 words were tested in experiment 1, and 18 sentences in experiment 2, respectively. If the subjects made a mistake, they were instructed to repeat the word.

D. Measurements

The acoustic sound of each word was analyzed by using ProsodyPro, a Pratt script for prosody analysis [32]. When the script was run, two windows were displayed on the screen, one with pulse markings and the other with TextGrid together with the waveform. Segmentation and labeling for each experiment was all 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. Consonant was considered terminated at the moment of release and the duration of a plosive consonant was equated with the duration of the closure. The overall data was analyzed by two-way ANOVAs, followed by a Tukey post-hoc analysis with a 95% confidence interval using IBM SPSS Statistics 19 (SPSS Inc., Illinois, USA).

IV. Experiments and Results

A. Experiment 1

This experiment represents a replication and extension of [19] with both British English-speaking subjects and Mandarin-speaking subjects. The correlation between the number of syllables and the duration of syllables for English, and whether Mandarin speakers produced the same duration patterns as the native speakers were designed to be tested.

The stimuli, as shown in Table I, consist of four sets of words, built around the base words stick, sleep, shade, and speed. Then three derivational suffixes were added to the base words separately, one of them (-y) monosyllabic and two (-ily, -iness) disyllabic. Thus, the words ranged from one to three syllables in length.

TABLE I
Test Materials Used in Experiment 1

base	stick	sleep	shade	speed
-y	sticky	sleepy	shady	speedy
-ily	stickily	sleepily	shadily	speedily
-iness	stickiness	sleepiness	shadiness	speediness

The data from Experiment 1 were presented in Tables II and III and graphically in Fig. 1 and Table II. Tables II and III displayed mean durations of the base in base form and in the derived word, mean duration of syllable nucleus (vowel) and B/D ratio (the ratio of the duration of the base word to the duration of the same segments occurring in the derived word) for British speakers and Mandarin ESL speakers, respectively.

Fig. 1 shows the comparisons of mean duration of base form, and Fig. 2 shows the comparison of mean duration of syllable nuclei produced by two groups of speakers.

TABLE II
MEAN DURATIONS (IN MILLISECONDS), STANDARD DEVIATIONS, AND B/D RATIOS FOR FOUR SETS OF WORDS AND CORRESPONDING SYLLABLE NUCLEI PRODUCED BY BRITISH SPEAKERS

Utterance	Mean duration of base	Mean duration of syllable nucleus	Standard deviation	B/D ratio
stick	607.18	144.89	52.13	
sticky	425.22	99.54	45.51	1.43
stickily	404.05	88.28	52.88	1.50
stickiness	384.98	86.45	44.38	1.58
sleep	676.55	231.73	45.74	
sleepy	479.03	171.70	57.03	1.41
sleepily	458.53	153.15	64.02	1.48
sleepiness	410.09	137.57	54.38	1.65
shade	660.55	314.44	57.50	
shady	416.36	188.39	37.87	1.59
shadily	377.94	152.37	50.90	1.75
shadiness	371.56	147.72	40.90	1.78
speed	731.49	300.00	94.40	
speedy	480.09	175.27	65.41	1.52
speedily	442.31	143.98	48.18	1.65
speediness	425.78	135.32	43.73	1.72

As can be seen from Fig. 1, there is a clear correlation between the duration of the base (or the duration of the same segments in the derived words) 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. For instance, when they read the words stick, sticky, stickily, and stickiness, as the number of syllables increases from one to two, then to three, the duration of the base form stick in the following three words decrease consistently. In addition, similar duration patterns are seen in the syllable nuclei (vowel), as displayed in Fig. 2, which is true of both British and Mandarin ESL speakers.

The present data are in accordance with Lehiste's statement that durational structure is conditioned by the number of syllables in a word. As the number of syllables increases, the duration of the base form in the derived words decreases and the shortening in the duration of base in the derived form is achieved more at the expense of vowels. Mandarin ESL speakers also produce a similar tendency when they read the same materials.

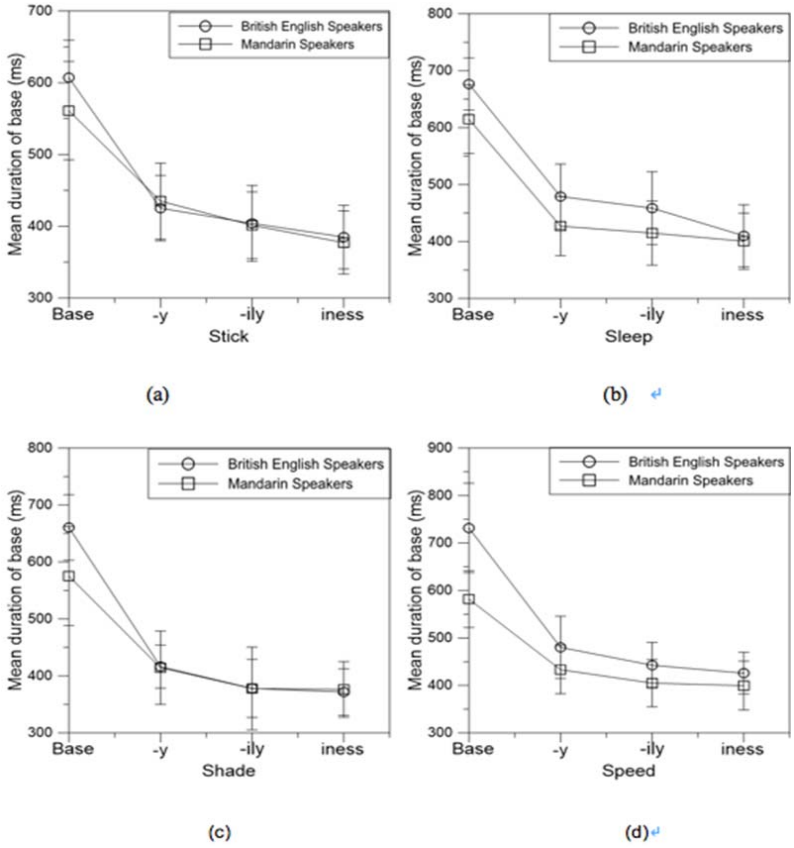

Fig. 1 Mean durations with standard deviations of base and of the same segments in the derived words for the words stick, sleep, shade and speed by British English speakers and Mandarin ESL speakers. The base word and the derivative forms are indicated on the horizontal axis; the vertical axis displays mean durations of the base form and of the same segments in derived words

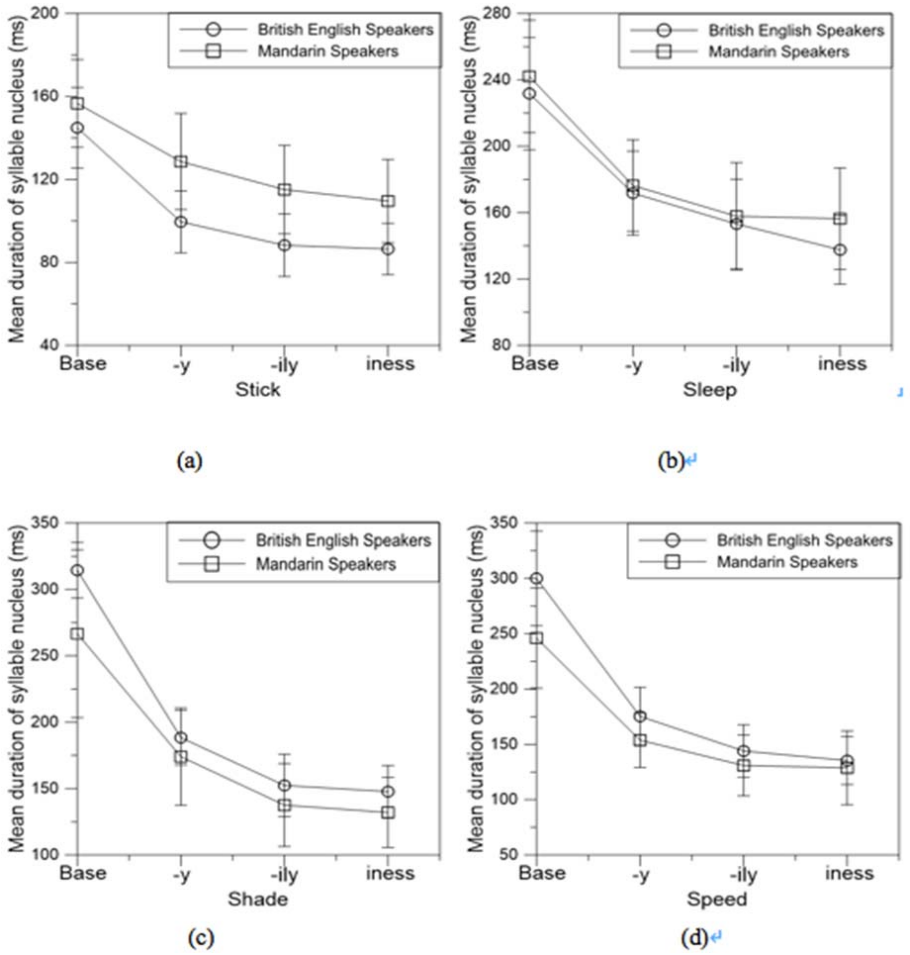

Fig. 2 Mean durations with standard deviations of syllable nuclei for the words stick, sleep, shade and speed and their derivative words by British English speakers and Mandarin ESL speakers. The base word and the derivative forms are indicated on the horizontal axis; the vertical axis displays mean durations of the syllable nuclei in the base word and derived words

TABLE III
MEAN DURATIONS (IN MILLISECONDS), STANDARD DEVIATIONS, AND B/D RATIOS FOR FOUR SETS OF WORDS AND CORRESPONDING SYLLABLE NUCLEI PRODUCED BY MANDARIN SPEAKERS

Utterance	Mean duration of base	Mean duration of syllable nucleus	Standard deviation	B/D ratio
sticky	435.08	128.59	52.96	1.29
stickily	401.38	115.10	46.40	1.40
stickiness	377.25	109.58	43.89	1.49
sleep	614.75	242.05	60.16	
sleepy	427.03	176.33	51.86	1.44
sleepily	414.86	157.81	56.34	1.48
sleepiness	400.61	156.36	49.27	1.53
shade	575.21	266.63	86.98	
shady	414.31	174.00	64.42	1.39
shadily	377.65	137.56	72.67	1.52
shadiness	376.15	132.00	48.96	1.53
speed	581.78	245.99	60.09	
speedy	433.08	153.78	50.80	1.34
speedily	404.72	130.87	49.75	1.44
speediness	399.44	128.77	51.56	1.46
Mean	449.65	163.25	57.16	1.44

In order to statistically test whether there is a significant difference between these two groups in this tendency, a two-way ANOVA was conducted to examine the effect of subject groups and varying suffixes (-y, -ily, -iness) on the mean B/D ratio. The result shows that there is no significant interaction between the effect of subject groups and varying suffixes on the mean B/D ratio for each set of words with $p>0.05$ ($p=0.90$ for the set of stick, $p=0.12$ for the set of sleep, $p=0.87$ for the set of shade, $p=0.73$ for the set of speed). As can be seen from Table IV, there are significant differences in the mean B/D ratio of each set of words between British English speakers and Mandarin ESL speakers ($p<0.05$), except for the set of sleep. 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. 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 ($p<0.05$). Duration of the base form in derived words has the largest polysyllabic shortening when suffix -ily and -iness are added to a base word, and second largest polysyllabic shortening when suffix -y is added to a base word.

TABLE IV
MEAN VALUES OF B/D RATIO OF FOUR SETS OF WORDS UNDER THE EFFECTS OF GROUPS OF SUBJECTS AND VARYING SUFFIXES. ALSO DISPLAYED ARE THE F- AND P-VALUES OF THE MAIN EFFECT OF A TWO-FACTOR ANOVA

Set of word		Subject group		Suffix		
		British	Mandarin	-y	-ily	-iness
Mean B/D Ratio	Stick	1.52	1.40	1.35	1.45	1.53
		F(1,30)=9.92		F(2,91)=7.25		
		P<0.01		P<0.01		
	Sleep	1.53	1.50	1.44	1.49	1.59
		F(1,30)=1.09		F(2,91)=9.86		
		P=0.30		P<0.01		
	Shade	1.72	1.50	1.47	1.62	1.63
		F(1,30)=20.17		F(2,91)=4.93		
		P<0.01		P<0.01		
	Speed	1.64	1.42	1.42	1.52	1.56
		F(1,30)=31.84		F(2,91)=5.92		
		P<0.01		P<0.01		

B. Experiment 2

This experiment used three sets of phrases in the corresponding carrier sentences to test the view that syllable duration was strongly influenced by stress and by positions in words, phrases, and sentences and to check whether it is also true of Mandarin ESL speakers when they read the same materials. The stimuli for this experiment consist of two-word adjective-noun phrases which were all three syllables in size, as displayed in Table V.

It is known that a particular combination of phrase size, stress pattern, and boundary location specified a phrase paradigm. Some instances of phrase paradigms include 1+01#, 12+10#, etc. The digits 0, 1 and 2 denote no stress, primary stress and secondary stress respectively. The + symbol demotes a word boundary, and the # symbol denotes a phrase boundary (as shown in [25]). Altogether, 113 different phrase paradigms were reported by [25], and this experiment only used stress level 01+1# and "The + adjective + noun+ verb phrase" phrase type. The three sets of phrases were put in a standard syntactic context in which the target phrase was put in the initial position of the sentences and was followed immediately by a verb phrase (1a. 2a. 3a in Table V), and then in a nonstandard syntactic context (1b, 2b, 3b in Table V).

TABLE V
TEST MATERIALS IN EXPERIMENT 2

(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.

Figs. 3 and 4 show how the duration of syllable is dependent on the position of the syllable in a word and phrase. For both British English speakers and Mandarin ESL speakers, 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. Furthermore, considering the effect of stress, word-final syllables with primary stress were consistently elongated. These are true of both standard context and nonstandard context. Specific data are displayed in Tables VI and VII.

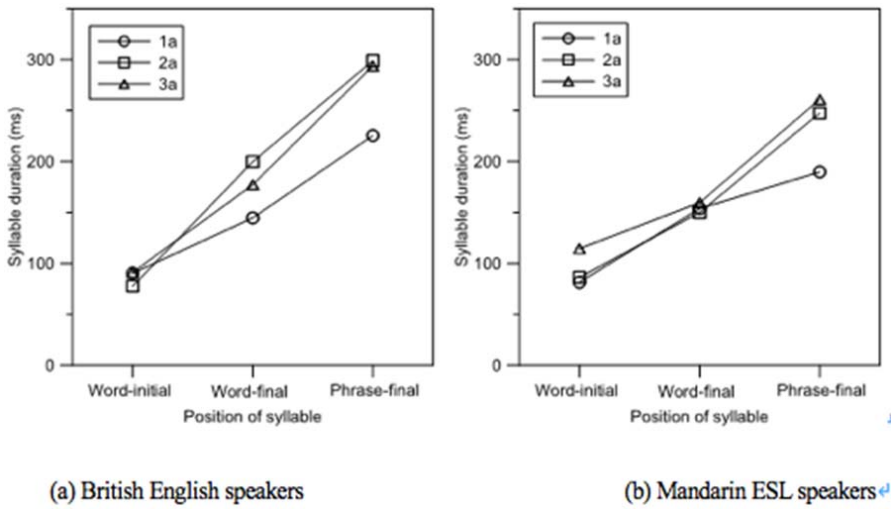

Fig. 3 Duration of syllables as a function of their positions within words and phrases (1a, 2a, 3a are displayed in Table V) in standard syntactic context for British English speakers and Mandarin ESL speakers

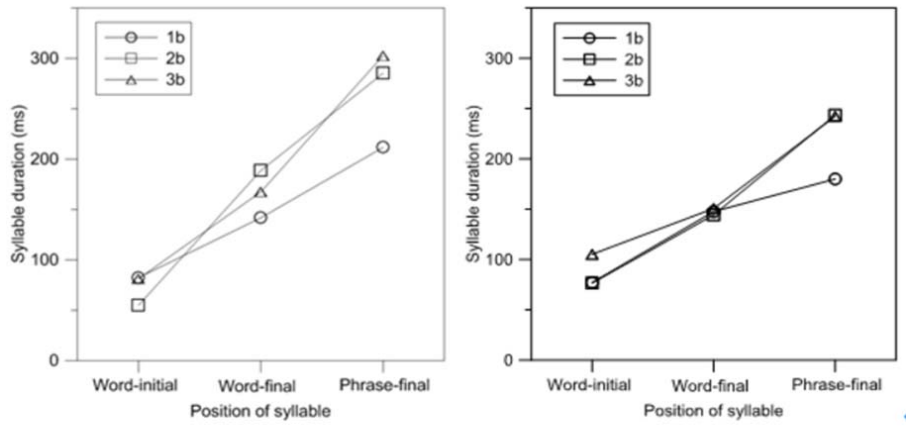

Fig. 4 Duration of syllables as a function of their positions within words and phrases (1b, 2b, 3b are displayed in Table V) in nonstandard syntactic context for British English speakers and Mandarin ESL speakers

TABLE VI
MEAN DURATIONS OF WORD-INITIAL SYLLABLE, WORD-FINAL SYLLABLE AND PHRASE-FINAL SYLLABLE IN SENTENCES FOR BRITISH SPEAKERS

Utterance	Word-initial syllable	Word-final syllable	Phrase-final syllable
1a	90.35	144.71	225.54
1b	82.36	141.75	211.61
2a	77.67	199.90	298.99
2b	54.95	188.64	285.29
3a	91.19	177.30	293.94
3b	81.23	167.44	302.61
Mean	79.63	169.96	269.66

TABLE VII
MEAN DURATIONS OF WORD-INITIAL SYLLABLE, WORD-FINAL SYLLABLE AND PHRASE-FINAL SYLLABLE IN SENTENCES FOR MANDARIN ESL SPEAKERS

Utterance	Word-initial syllable	Word-final syllable	Phrase-final syllable
1a	81.54	153.88	189.83
1b	77.15	147.65	180.06
2a	86.37	150.06	247.75
2b	76.55	144.33	243.63
3a	114.54	159.66	261.23
3b	105.16	150.76	242.41
Mean	90.22	151.06	227.48

Apart from this, we also investigate the effect of syntax for syllable durations by varying the syntactic contexts into standard syntactic context (early position in the sentence) and nonstandard syntactic context (later position in the sentence). For example, for the phrase "remote stream", a standard context would be "The remote stream was perfect for fishing" with the phrase always preceded by the sentence-initial "The" and followed by a verb phrase. A corresponding nonstandard context would be "Fishing in the remote stream was perfect for fishing". Fig. 3 and Fig. 4 show syllable durations in standard context and nonstandard context for the two subject groups. Syllables with the same positions in words and phrases have the same stress level (word-initial words are unstressed; word-final words are stressed; phrase-final words are monosyllabic). If equivalent syllables had the same duration in both standard and nonstandard contexts, then the data points are likely to fall on the diagonal line. The data in Fig. 5 display 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 British English speakers and Mandarin ESL speakers.

A two-way ANOVA was conducted to examine the effect of subject groups and location in words and phrases on the mean duration of syllables under standard and nonstandard context, as is displayed in Table VIII. 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 are no statistically significant differences between British English speakers and Mandarin ESL speakers 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 differences can be found. Duration of word-final syllables produced by British English speakers (mean duration 173.97 ms) is longer than that produced by Mandarin ESL speakers (mean duration 154.53 ms); and duration of phrase-final syllables produced by British English speakers (mean duration 272.82 ms) is longer than that produced by Mandarin ESL speakers (mean duration 232.94 ms). There are also significant differences among syllables in word-initial position, word-final position and phrase-final position ($p<0.05$). Duration of syllables in phrase-final position (mean duration 248.07 ms) is longer than duration of syllables in word-final position (mean duration 161.91 ms); and duration of syllables in word-final position (mean duration 161.91 ms) is longer than duration of syllables in word-initial position (mean duration 91.21 ms). 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 ($p<0.05$) can also be found under nonstandard context. Simple main effect analysis

shows that no significant difference exists between British speakers and Mandarin ESL speakers 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 165.94 ms) than Mandarin ESL speakers (mean duration 147.58 ms); and British English speakers also produce longer duration for syllables in phrase-final position (mean duration 266.50 ms versus mean duration 222.03 ms). Significant differences can also be tested within the locations in words and phrases (p<0.05). Duration of syllables in phrase-final position (mean duration 238.57 ms) is longer than duration of syllables in word-final position (mean duration 154.46 ms); and duration of syllables in word-final position (mean duration 154.46 ms) is then longer than duration of syllables in word-initial position (mean duration 81.19 ms).

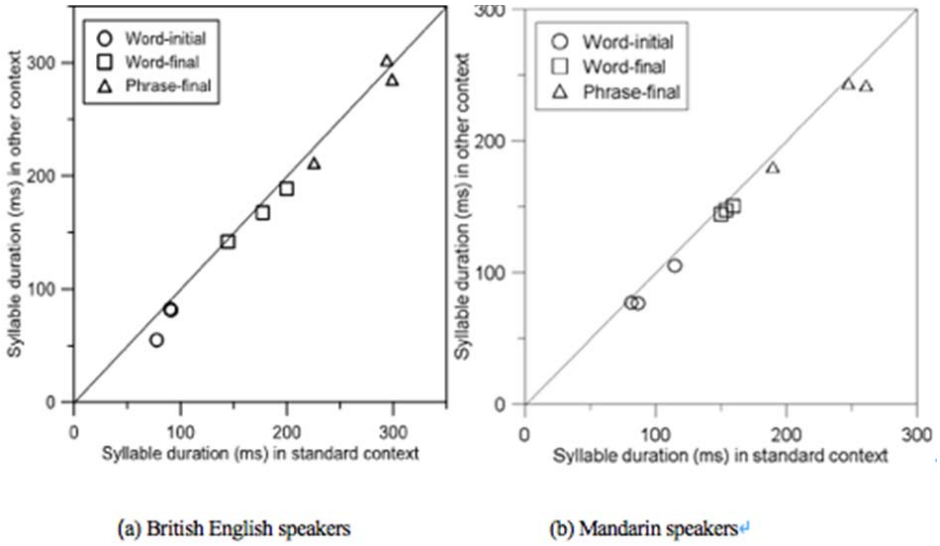

Fig. 5 Duration of syllables in standard syntactic context versus nonstandard syntactic contexts (labeled "other" on the ordinate) for British English speakers and Mandarin ESL speakers. Each point represents two syllables equated for stress-level, position within word, and position within phrase in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

TABLE VIII
MEAN DURATION OF SYLLABLES IN TWO SENTENCE CONTEXTS UNDER THE EFFECTS OF SUBJECT GROUPS AND LOCATION IN WORDS AND PHRASES. ALSO DISPLAYED ARE THE F- AND P-VALUES OF THE MAIN EFFECT OF A TWO-FACTOR ANOVA

Sentence context		Subject group		Location in word and phrase		
		British	Chinese	word-initial	word-final	phrase-final
Mean Duration	Standard	177.73	160.54	91.21	161.91	248.07
		F(1,30)=11.17			F(2,259)=334.03	
		P=0.01			P<0.01	
	Non-standard	168.43	151.75	81.19	154.46	238.57
		F(1,30)=13.18			F(2,259)=427.07	
		P<0.01			P<0.01	

The result shows that in both the standard and nonstandard context, British English speakers and Mandarin ESL speakers are significantly different in the mean duration of word-final and phrase-final syllables. British English speakers produce longer duration for syllables in word-final position and phrase-final position. And also, syllables in different positions in words and phrases are significantly different from each other. Duration of syllables in phrase-final position is longer than duration of syllables in word-final position; and duration of syllables in word-final position is longer than duration of syllables in word-initial position.

V. GENERAL 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. From the data collected and analyzed in the two experiments, we found:

First, durational pattern is conditioned by the number of syllables. As the number of syllables increases, the duration of a base form in the derived words is considerably shortened for British English. This is also true of Mandarin speakers when they read the same words. However, although the similar tendency occurs to both British English and Mandarin ESL speakers, there is a significant difference between the native

English speakers and L2 learners in the degree of that reduction. British English speakers show greater reduction in the base form of derived words as they have greater B/D ratio than Mandarin ESL speakers. 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.

Second, both British and Mandarin ESL speakers have a consistently increasing tendency for duration of syllables from word-initial position to phrase-final position. That is, durations of word-final words are longer than durations of word-initial words; durations of phrase-final words are longer than durations of word-final words. In both standard and nonstandard context, British English speakers and Mandarin ESL speakers are significantly different in the mean duration of word-final and phrase-final syllables. British English speakers produce longer duration for syllables in word-final position and phrase-final position.

Based on these findings above, 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 accuracy of English for Mandarin ESL speakers. This study not only offers some clues for later research on duration patterns of languages but also benefits L2 learners by improving their sensitivity to the duration differences between their L1 and L2. Mastering a good understanding of the effect of polysyllabic shortening and position in words or phrases can be also helpful for L2 teachers to set priorities when they teach pronunciation to L2 learners.

ACKNOWLEDGMENT

The author would like to express her gratitude to Professor. Xu Yi for his supervision and offering so many helpful suggestions. A special thank you would also go to Miya Mai for her help on processing data and providing so much encouragement.

REFERENCES

[1] Anderson-Hsieh, J., & Venkatagiri, H. (1994). Syllable duration and pausing in the speech of Chinese ESL speakers. *TESOL Quarterly, 28*, 807-812.
[2] Barnwell III, T. P. (1971). *An Algorithm for Segment Durations in a Reading Machine Context.* DTIC Document.
[3] Bolinger, D. L. (1963). Length, vowel, juncture. *Linguistics*, *1*(1), 5–29.
[4] Chang, J. (1987). Chinese speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teachers' guide to interference and other problems* (pp. 224–237). Cambridge, MA: Cambridge University Press.
[5] Chao, Y. R. (1968). *A grammar of spoken Chinese*. Univ of California Press.
[6] Chen, M. Y., Anderson, S. R., Bresnan, J., Comrie, B., Dressler, W., Ewen, C. J., & Huddleston, R. (2000). *Tone sandhi: Patterns across Chinese dialects*. Cambridge University Press Cambridge, UK:
[7] Deterding, D. (2006). The pronunciation of English by speakers from China. *English World-Wide*, *27*(2), 175–198.
[8] Duanmu, S. (2007). *The phonology of standard Chinese*. Oxford University Press.
[9] Ellis, Rod. 1994. *The study o f second language acquisition.* Oxford: Oxford University Press.
[10] Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *The Journal of the Acoustical Society of America*, *84*(1), 70–79.
[11] Gaitenby, J. H. (1965). The elastic word. *Haskins Laboratories Status Report on Speech Research*, *2*, 1–3.
[12] Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology*, *7*(515-546).
[13] Harris, M. S., & Umeda, N. (1974a). Effect of speaking mode on temporal factors in speech: vowel duration. *The Journal of the Acoustical Society of America*, *56*(3), 1016–1018.
[14] Klatt, D. (1974). The duration of (s) in English words. *Journal of Speech, Language, and Hearing Research*, *17*(1), 51–63.
[15] Klatt, D. H. (1972). Generative theory of segmental duration in English. *The Journal of the Acoustical Society of America*, *51*(1A), 101.
[16] Klatt, D. H. (1975). Vowel Lengthening is Syntactically Determined in a Connected Discourse. *Journal of Phonetics*, *3*(3), 129–140.
[17] 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.
[18] Lehiste, I. (1971). Temporal organization of spoken language. *Form and Substance*, 159–169.
[19] Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *The Journal of the Acoustical Society of America*, *51*(6B), 2018. Retrieved from http://scitation.aip.org/content/asa/journal/jasa/51/6B/10.1121/1.1913062.
[20] Lehiste, I. (1975). Some factors affecting the duration of syllabic nuclei in English. In *Proceedings of the First Salzburg Conference on Linguistics* (pp. 81–104).
[21] Lin, T. (1985). Preliminary experiments on the nature of Mandarin neutral tone. *Working Papers in Experimental Phonetics. Beijing University Press, Beijing*, 1–26.
[22] Lindblom, B., & Rapp, K. (1973). Some temporal regularities of spoken Swedish. *Auditory Analysis and Speech Perception (London, 1975)*, 387–396.
[23] Major, R. C. (2001). *Foreign accent.* Oxford: Blackwell Publishing Ltd.
[24] Martin, J. G. (1970). On judging pauses in spontaneous speech. *Journal of Verbal Learning and Verbal Behavior*, *9*(1), 75–78. doi:10.1016/S0022-5371(70)80010-X.
[25] Nakatani, L. H., Connor, K. D. O., Laboratories, B., & Hill, M. (1981). Prosodic Aspects of American English Speech Rhythm, *106*.
[26] Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *The Journal of the Acoustical Society of America*, *54*(5), 1235–1247.
[27] Parmenter, C. E., & Trevino, S. N. (1935). The length of the sounds of a Middle Westerner. *American Speech*, 129–133.
[28] Tajima, K., Port, R.., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign accented English. *Journal of Phonetics, 25*, 1-24.
[29] Umeda, N. (1975). Vowel duration in American English. *The Journal of the Acoustical Society of America*, *58*(2), 434–45. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2766870&tool=pmcentrez&rendertype=abstract.
[30] Van Santen, J. P., & Shih, C. (2000). Suprasegmental and segmental timing models in Mandarin Chinese and American English. *The Journal of the Acoustical Society of America*, *107*(2), 1012–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10687710.
[31] Vanderslice, R., & Ladefoged, P. (1972). Binary suprasegmental features and transformational word-accentuation rules. *Language*, 819–838.
[32] Xu, Y. (2013). ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis, 7–10.
[33] Xu, Y. and Wang, M. (2009). Organizing syllables into groups -- Evidence from F0 and duration patterns in Mandarin. *Journal of Phonetics* 37: 502-520.