Rhythmic Patterning in Malaysian and Singapore English
Rachel Siew Kung Tan and Ee-Ling Low
Language and Speech published online 31 July 2013
DOI: 10.1177/0023830913496058

The online version of this article can be found at:
http://las.sagepub.com/content/early/2013/07/30/0023830913496058

A more recent version of this article was published on - May 20, 2014

Published by:
SAGE
http://www.sagepublications.com

Additional services and information for Language and Speech can be found at:

Email Alerts: http://las.sagepub.com/cgi/alerts
Subscriptions: http://las.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav

Version of Record - May 20, 2014
>> OnlineFirst Version of Record - Jul 31, 2013

What is This?
Rhythmic Patterning in Malaysian and Singapore English

Rachel Siew Kuang Tan
University of Malaya, Malaysia

Ee-Ling Low
National Institute of Education, Nanyang Technological University, Singapore

Abstract
Previous work on the rhythm of Malaysian English has been based on impressionistic observations. This paper utilizes acoustic analysis to measure the rhythmic patterns of Malaysian English. Recordings of the read speech and spontaneous speech of 10 Malaysian English speakers were analyzed and compared with recordings of an equivalent sample of Singaporean English speakers. Analysis was done using two rhythmic indexes, the PVI and VarcoV. It was found that although the rhythm of read speech of the Singaporean speakers was syllable-based as described by previous studies, the rhythm of the Malaysian speakers was even more syllable-based. Analysis of the syllables in specific utterances showed that Malaysian speakers did not reduce vowels as much as Singaporean speakers in cases of syllables in utterances. Results of the spontaneous speech confirmed the findings for the read speech; that is, the same rhythmic patterning was found which normally triggers vowel reductions.

Keywords
Rhythm, syllable-timing, stress-timing, varieties of English

Introduction
1.1 Background
Traditionally, languages have been classified as stress-timed, syllable-timed and mora-timed. This categorization is based on the notion of isochrony, or the regular occurrence of some form of speech unit. Stress-timed languages were those that have stressed syllables occurring at regular intervals such as English and German, while syllable-timed languages have syllables with similar durations irrespective of whether they are stressed or unstressed such as Spanish and French, while “mora-timed” languages have morae of approximately the same duration (Ladefoged & Johnson, 2011, p. 252). The lack of experimental evidence from early research on rhythm of languages led to a view

Corresponding author:
Rachel Siew Kuang Tan, Faculty of Languages and Linguistics, University of Malaya, Wilayah Persekutuan, Kuala Lumpur, 50603, Malaysia.
Email: tansk@um.edu.my
of rhythm which suggested that rhythmic classifications of languages should be based on combinations of phonological, phonetic, lexical and syntactic qualities of the different languages (Dasher & Bolinger, 1982; Dauer, 1983, 1987; Low, 2006, p. 102). Dauer has proposed that the three main influences on the rhythm of a language are the complexity of syllable structure, the presence or absence of vowel reduction and the stress patterning of a language (1983). She also suggested that stress-based languages tend to have more complex syllable structures, while in syllable-based languages, absence of vowel reduction is common. Dasher and Bolinger (1982) offered the view that phonemic vowel length distinction tends to be absent in syllable-based languages. Laver adds that the perceived rhythm of a language stems from “the coincidence of segmental sonority, syllable weight and lexical stress in the lexicon of a language, and of the pragmatic use of the lexicon in the utterances of that language” (Laver, 1994, p. 527). The combination of lexical and pragmatic construction explains why some utterances are more regularly rhythmic than others of the same language. Linguists generally believe that languages fall along a continuum where they can be classified as being either more or less “stressed-based” or “syllable-based” (Dauer, 1983; Miller, 1984).

1.2 Rhythm indexes

In more recent years, rhythm indexes developed by Ramus, Nespor, and Mehler, (1999); Low, Grabe, and Nolan (2000); Deterding (2001); Grabe and Low (2002) and Dellwo and Wagner (2003) have produced empirical evidence to support the plausibility of rhythmically categorizing languages. Low (2006) provides a useful explanation and summary of the different rhythmic indexes. Indexes like Ramus et al. (1999) and Low et al. (2000) exploit syllable complexity, vowel reduction and stress-based lengthening to provide a measurement that is able to capture rhythmic differences between languages or varieties of a language. These acoustic measures are derived from the segmentation of speech segments into vocalic and consonantal intervals, enabling the measurement of successive variability in these intervals on the basis that stress-based languages tend to have greater contrast in vowel duration between stressed and unstressed syllables, while stress-timed languages tend to have greater variability in the complexity of the consonant clusters, which therefore influences the duration of consonantal intervals as well. One advantage that both the Low et al. pairwise variability index (PVI) and Ramus et al. interval measures (IM) have is that they enable the measurement of rhythm without reference to the phonological constituency of syllable intervals, the latter of which can be advantageous to those unfamiliar with the phonology of a particular language (Nolan & Asu, 2009).

The Ramus et al. (1999) rhythm index takes into account three different properties of the durational variation in languages. The %V, which is defined as the proportion of vocalic intervals in the sentence (that is, the section of speech between the vowel onset and offset); ∆V, which is the standard deviation of vocalic intervals; and ∆C, the standard deviation of consonantal intervals which is the section of speech between vowel offset and vowel onset. These metrics were applied to speakers reading sentences in stress-timed (English, Dutch and Polish), syllable-timed (French, Spanish, Italian, Catalan) and mora-timed (Japanese) languages. It was found that a combination of ∆V and either ∆C or %V appears to offer useful measures of rhythmic relatedness between languages. However, the limitation of especially the ∆C and ∆V is that while they take into consideration overall interval variations, the successive interval variations are ignored (Low et al., 2000). Other criticisms raised regarding the use of %V and ∆C are that the measure is not robust across speech rate and that it could be a measure of syllable complexity rather than rhythm (Wagner & Dellwo, 2004).

The Low et al. (2000) index, the PVI, measures the durations of vowels in successive syllables in an utterance. The underlying assumption is that there is a lack of vowel length distinction
between stressed and unstressed syllables in syllable-timed languages while this is not so for stress-timed languages. The main reason for this lack of vowel length distinction is due to the relative absence of reduced vowels in syllable-timed languages. Taylor (1981) claimed that vowel duration is the key to syllable-timing in languages and this prompted Low (1998) and subsequently Low et al. (2000) to consider vowel durations in the measurement of rhythm. The PVI index, which was found in Low et al. (2000), shows the mean absolute difference between successive pairs of vowels in an utterance. Calculations of the difference in duration between successive pairings of vowels in utterances are calculated and the absolute values (that is, discarding the negative sign where it occurs) are taken. The mean difference is calculated by dividing the difference between successive durations and dividing by the average of the two durations in order to control for speech rate variation, and this is expressed in an index,

\[
PVI = 100 \times \left( \frac{2}{m-1} \sum_{k=1}^{m-1} \left( \frac{d_k - d_{k+1}}{\left( \frac{d_k + d_{k+1}}{2} \right)} \right) \right)
\]

where \( m \) is the number of intervals and \( d \) is the duration of the \( k \)th interval. Thus, stress-timed languages would have a higher index as there is a greater variability in the duration between the successive vowels in the sentence, while syllable-timed languages would have a lower index. Low et al. (2000) compared the PVI scores for vocalic intervals for British English (BrE) and Singapore English (SgE) and found that the latter had a significantly higher PVI score than the former, reflecting the perceived rhythmic distinction where BrE is categorized as being stress-timed, and SgE as syllable-timed. Low et al. also used Ramus et al.’s (1999) IM and found that the \( \Delta V \) did show a pattern comparable to the PVI. However, they had reservations about the use of \%\( V \) as it did not reflect the differences between varieties of the same language.

Grabe and Low (2002) calculated the vocalic PVI (NPVI) and the raw consonantal PVI (rPVIC) scores for prototypical stress-timed languages (English, German and Dutch), prototypical syllable-timed languages (French and Spanish) and a prototypical mora-timed language (Japanese). In addition to these languages, they also calculated the vocalic PVI and raw consonantal PVI scores for Polish and Catalan, which are rhythmically indeterminate, and three previously rhythmically unclassified languages (Greek, Romanian and Estonian). They found that the prototypical stress-timed languages (English, Dutch and German) had higher PVI scores than the syllable-timed languages (French and Spanish) while many of the rhythmically indeterminate and unclassified languages had PVI scores which fell between the stress-timed and syllable-timed languages. In contrast to Ramus et al.’s (1999) clear separation of Japanese as a mora-timed language, Grabe and Low did not find that Japanese had a distinct PVI space. Grabe and Low did compare PVI and IM scores for the speech samples they measured, finding some similarities and some clear differences. However, conclusions were difficult to draw from the speech samples as they were based on only one speaker per language.

A number of studies have supported the use of normalization of metrics, as it has been found that speech rate can influence the values obtained in the use of metrics (Barry, Andreeva, Russo, Dimitrova, & Kostadinova, 2003; Dellwo & Wagner, 2003). Dellwo (2006) used a rate-normalized metric, VarcoC, which is the standard deviation of consonantal interval duration divided by the mean consonantal duration and then multiplied by 100. It was found that VarcoC was better than \( \Delta C \) at discriminating stress-timed English and German from syllable-timed French at all speech rates. It was found that VarcoC did vary according to the intended rate but there did not appear to (the PVI of consonantal durations) be a systematic correlation across languages. Dellwo and Wagner (2003) found little consistent correlation between \%\( V \) and speech rate, implying that normalization may not be essential for \%\( V \). White and Mattys (2007a) also used rate-normalized interval
measures, VarcoV (normalised standard deviation of vocalic interval durations divided by the mean vocalic duration) and VarcoC (normalised standard deviation of consonantal interval durations divided by the mean consonant duration). While VarcoV discriminated the influence of the rhythm of one’s first language on the second language spoken, VarcoC did not. Calculation for the VarcoV is done by calculating the standard deviation of vocalic interval duration divided by the mean vocalic interval duration and multiplied by 100. This is expressed in an index,

\[
\text{VarcoV} = \frac{\Delta V}{\bar{V}} \times 100
\]

where \( \Delta V \) refers to the standard deviation of the vocalic interval and \( \bar{V} \) refers to the mean duration of the vocalic intervals.

Justification for the use of a combination of different indexes comes from the work of Loukina, Kochanski, Shih, Keane, and Watson (2009). They found that the use of a combination of two metrics was more effective at classifying languages compared with the use of one matrix, although the most efficient pairs of measures differed in their success rate of identifying specific languages. They also found that the use of three metrics did not significantly improve the success rate of classifying languages compared to the use of two metrics. A number of studies have also used different combinations and variations of indexes to measure rhythm (Asu & Nolan, 2005; Benton, Dockenförf, Jin, Liu, & Edmondson, 2007; Dellwo & Wagner, 2003; Gibbon & Gut, 2001; Gut, Urua, Adouakou, & Gibbon, 2001; Lin & Wang, 2005).

Some studies have examined the rhythm between different varieties within a language, such as Barry et al. (2003) for Italian (\( \Delta V, \Delta C, \% V, \text{PVI variants} \)) and Ferragne and Pellegrino (2004) for BrE (\( \Delta V, \Delta C, \% V, \text{VarcoV, VarcoC, PVI variants} \)). Yet other studies have examined the influence of the speakers’ first language on the second language rhythm. Gut (2003) found that German L2 is influenced by Chinese, English, French, Italian and Romanian L1 (\( \Delta C, \% V \) and PVI); Lin and Wang (2005) showed that Canadian English L2 is influenced by Chinese L1 (\( \Delta C, \% V \)); while Mok and Dellwo (2008) found that L2 English is influenced by Cantonese and Beijing Mandarin (\( \Delta V, \Delta C, \Delta S \) (the standard deviation of syllabic durations), \( \% V, \text{VarcoV, VarcoC, VarcoS} \) (which is the normalised standard deviation of syllabic durations), \( r\text{PVIC} \) (raw PVI of consonantal durations), \( r\text{PVIS} \) (raw PVI of syllabic durations), \( n\text{PVIV}, n\text{PVIS} \) (normalised PVI of vocalic and syllabic durations)). Carter (2005) showed that American English L2 is influenced by Mexican Spanish L1 (PVI), and Whitworth (2002) proved that English and German bilinguals in both languages are influenced by their parental languages (PVI). In Carter’s (2005) study, the speech of English–Spanish bilingual speakers who had moved from Mexico to North Carolina in their childhood was measured. He found that the PVI scores for L2 English were between the low PVI scores for L1 Spanish and high scores of L1 English and claimed that this was the result of lack of vowel reduction in L1 Spanish. This shows that the PVI is able to provide significant information regarding accommodation between rhythmically distinct languages. White and Mattys (2007a) measured the rhythm of L1 English, Dutch, Spanish and French as well as L2 English, Dutch, Spanish and French using \( \Delta V, \Delta I \text{Eng}, \text{VarcoV, VarcoC, NPVI and CPVI} \). Their hypothesis was that the rhythm of the L2 should be intermediate between the speaker’s native language and the native rhythm of the language learnt. For example, the rhythm of the Spanish of English speakers learning Spanish as an L2 should be intermediate between the rhythm of English and Spanish spoken by native speakers. Of all the indexes used in their study, VarcoV was found to be the best discriminator of the rhythm of languages spoken as L1 and L2.

The advent of indexes to measure rhythm has enabled a more objective means of categorizing languages based on rhythm. This is especially useful for the categorizing of languages within
rhythm classes as it is challenging to do it perceptually. Nazzi, Bertoncini, and Mehler (1998), as well as Ramus, Dupoux, & Mehler (2003) found that their subjects could not discriminate languages within rhythm classes like English and Dutch (which are both stress-based) even though they could discriminate languages between rhythm classes like English and Spanish.

From the extensive review of the literature above, it is evident that using indexes like the PVI (Low et al., 2000) and VarcoV (Dellwo, 2006) is useful in aiding the comparison of Malaysian English (MalE) and SgE since both varieties of English have traditionally been categorized as being syllable-based.

While indexes like Ramus et al. (1999) and Low et al. (2000) have been very influential in the area of rhythmic research over the past decade, they have focused mainly on the duration of vowels and consonants. Nolan and Asu (2009) notably have questioned this and cited Ferragne (2008), who found the PVI of mean syllable intensity as well as duration gave a better discrimination of accents or dialects of BrE as compared to the PVI of duration alone. Nolan and Asu also cited Deterding (2001, p. 221), who argued that too many syllables in the conversational data he measured were devoiced and thus supported the measurement of syllables instead of vowels. Nolan and Asu proposed the measurement of the PVI of the foot while questioning the notion of determining rhythm just based on a single parameter, namely duration. Justification for this was the fact that PVI measurements in their previous work showed that BrE and Estonian had similarities and differences in terms of rhythm since BrE tends towards stress-timing while Estonian has qualities of syllable-timing and stress-timing. Thus, the rhythmic classification of a language like Estonian is difficult in terms of a unidimensional measure. They argued that the notion of rhythm even in the constraints of duration should accommodate rhythmic patternings that coexist. Thus, stress-timing and syllable-timing should not be seen as extreme points of a continuum but as “orthogonal dimensions,” (p.70) enabling a language to be both stress and syllable-timed at the same time. Nolan and Asu’s proposal to measure the foot and the syllable, while interesting, does present some significant problems, namely decisions regarding segmentation of the syllable and foot. This might lead to difficulties in comparing the languages as decisions on segmentation would vary according to the language. More importantly, Nolan and Asu were attempting to find a way to categorize Estonian, which has both stress and syllable-timed qualities. A more serious problem that was faced was the difficulty in the identification of stressed syllables. As a foot begins with a stressed syllable, lack of reliability in its identification poses a major difficulty in the identification of the foot. As the results of many of the studies reviewed so far (Barry et al., 2003; Benton et al., 2007; Dellwo & Wagner, 2003; Gibbon & Gut, 2001; Gut, 2003; Gut et al., 2001; Lin & Wang, 2005; Low, 1998; Grabe & Low, 2002; Low et al., 2000; Rickard, 2006) in this section tend to show that it is possible to categorize the rhythmic patterning of most languages based on the durational domain alone, the decision is undertaken in this study to measure the PVI of vowel durations only as well as to use the VarcoV index which essentially measures the standard deviation of the vocalic intervals. Consonant durations were not considered as previous work on SgE has shown that it is the vowels that contribute to the overall perception of syllable-timed rhythm (Low, 1998). This is also supported by White and Mattys (2007a) as it was found that VarcoC was limited in its ability to differentiate the rhythm of different languages.

1.3 Rhythm in Malaysian English and Singapore English

Rhythm in MalE has been described as syllable-based where all syllables are perceived to recur at nearly equal intervals of time, although MalE speakers do use a stress-timed rhythm in some
instances such as reading a formal passage (Baskaran, 2008). When informal speech is used, even educated MalE speakers use a syllable-timed rhythm (Baskaran, 2008). Instrumental research on the rhythm of MalE is scarce. Tan and Low (forthcoming) measured the rhythm of three female ethnically Malay speakers of MalE and three ethnically Malay speakers of SgE reading specially prepared sentences designed to include both full and reduced vowels, as used first in Low (1998). They found that while a significant difference was found between the PVI values of full and reduced vowel sentence sets for SgE speakers, no significant difference was found for the PVI values of full and reduced vowel sentence sets read by MalE speakers. This implies that there is greater vowel reduction in SgE compared to MalE.

According to Wee (2008, p. 273), the syllable-timed rhythm of colloquial SgE means that all syllables in an utterance have the same duration irrespective of whether the syllables are stressed or not. This kind of rhythm has been described as having a “staccato” effect (Platt & Weber, 1980, p. 57) and possessing a “machine-gun” like quality (Tay, 1993, p. 27). Deterding (2001) proposed that among the more important features that contribute to the staccato effect of SgE are the use of a glottal stop instead of a /t/ or /k/ for word-final plosives, final consonantal cluster simplification so fact may be realized as [fa:k] or even [fæʔ], neutralization contrasts of the long vowels and short vowels so pull and pool become homophones, and absence of reduced vowels in unstressed syllables so that the first syllable of concern has a full /ɒ/ vowel rather than a schwa.

Previous comparative studies on the rhythm of SgE have not always been consistent in their results. Low (1994) and Low and Grabe (1995), using read text, compared the vowel duration of one syllable with that of the following syllable and found greater variability in the BrE data, lending support to the notion that SgE is more syllable-timed. Deterding (1994, 2001), using the Variability Index with a normalizing procedure, measured syllables in the natural conversations of BrE and SgE speakers and confirmed that SgE is more syllable-timed compared to BrE. Low (1998), whose subjects were all ethnically Chinese, used the PVI and found a significant difference between the PVI values obtained for SgE and BrE. Lim (1996), whose subjects were ethnically Chinese, Indian and Malay, used the Variability Index with a normalizing procedure (Deterding, 1994) and found no significant difference between the rhythm of Chinese, Malay and Indian SgE speakers.

Based on the review of previous studies thus far, it can be seen that rhythm in MalE should be further investigated to substantiate impressionistic claims by earlier researchers as well as to enable a more comprehensive description of the rhythm of MalE. As researchers on the rhythm of SgE have used different materials to measure rhythm, it makes sense to measure the rhythm of MalE not only in carefully scripted sentences (Low, 1998), but to also measure the rhythm of MalE in a longer read text (Deterding, 2006) as well as in spontaneous speech.

2 Method

2.1 Subjects

The subjects of the study were carefully controlled for ethnicity, gender and educational background. They comprised 10 ethnically Malay Malaysian and 10 Malay Singaporean speakers of English. All the speakers were undergraduate students in their respective countries and had been educated in the local school systems. All the speakers resided in urban areas and spoke both English and Malay. The average age of the Malaysian speakers was 21.1 while the average age of the Singaporean speakers was 26 years.
2.2 Recordings and test materials

The recordings were done using a software called AvRack manufactured by “Realtek,” which is made in Taiwan, with the help of a mini clip-on microphone which was plugged in to a laptop computer. The recordings were done in quiet rooms with good acoustic support like carpeted flooring and cushioned walls where possible.

The Wolf passage (Deterding, 2006) was measured for rhythm. The ideal text for measuring rhythm is one that contains alternative strong–weak syllables with no initial approximants and no dark /l/. Words with initial approximants or dark /l/ should be avoided where possible as the measurement of the vowel duration would be problematic. The disadvantage of such a text is that it may sound artificial. Spontaneous speech is ideal, but for comparing rhythm it is crucial that the same material is used for different speakers, so a fixed text can be performed but at the same time it should not sound too artificial. Although Deterding (2006) has rightly pointed out that the North Wind and the Sun (NWS) passage poses some problems for measuring rhythm as it has a number of instances of consecutive strong syllables, the Wolf passage is no better because it also has similar problems. Similarly, Deterding (2006) noted that both the NWS and the Wolf passage pose problems for the measurement of vowel duration because of the presence of an initial /w/ or /r/ in a syllable resulting in problematic words like wind, were, which, was, when, warm, one, stronger; traveler; wrapped and around in NWS and words in the Wolf passage like was, once, watch, one, wolf; with, were, away, forest, raising, ran, rushed, tried, trick, threaten, racing, cried, trying and words like used and usual which had initial /j/. The instances of the occurrences of the dark /l/ make it difficult to measure the vowel duration before the dark /l/ as it is vocalized, even in BrE (Wells, 1982, p. 259). Both NWS and the Wolf passage have this problem, although with just one instance of a dark /l/, in fold, the problem is not as big in NWS. Nevertheless, the decision was made to use the Wolf passage while recognizing its limitations, and for purposes of triangulation, specially prepared sentences for measuring rhythm from a list of sentences designed by Low et al. (2000), which comprised only full vowels and those with a mixture of full and reduced vowels, were used.

As some might argue that the rhythm of read speech is not necessarily the same as that of spontaneous speech, an attempt is made to measure the rhythm of spontaneous speech. The speakers were asked to speak for approximately five minutes on the topic “My most memorable holiday” and the rhythm of this data was analyzed. Measuring the rhythm of spontaneous speech is problematic as speakers can be inconsistent in their rhythm. Another problem with measuring spontaneous speech is that it is very difficult to compare the rhythm of different speakers as the phonetic environments of the utterances measured are usually very different. In order to have some form of similarity for purposes of comparison, six utterances from each speaker, each utterance consisting of eight syllables, were selected for measurement. In the selection of the utterances, attempts were made to avoid words with /j/, /r/ and /w/ as well as dark /l/ to facilitate measurement of the interval durations, as the boundaries between vowels and approximants can be difficult to identify reliably from the visual analysis of waveforms and spectrograms (White & Mattys, 2007b), although this was not always possible. Following Deterding (2001), utterances with pauses in between words were avoided. Similarly, utterances where the speakers lengthened some words as they searched for words to continue their speech were not chosen. The words that were lengthened were often function words like “to” and “and.” The unnaturally lengthened words will influence the PVI value obtained. Figure 1 demonstrates an example of this when the speaker lengthened his realization of the preposition “to” before he is able to name the place he visited, namely Universal studios. The unnaturally lengthened /tu/ resulted in a vowel duration of 302 ms, which would influence the PVI value and was thus discarded.
2.3 Data analysis

The Pairwise Variability Index (PVI) suggested by Low et al. (2000) was used to estimate the degree of rhythmic variation in the utterances measured. Two modifications to the PVI were made following the method adopted by Deterding (2006). Firstly, the PVI was modified so that the minimum duration of all syllables measured was 30 msec. to overcome problems that can arise when there are no vowels in a syllable when elision occurs or where clear vowel length cannot be determined. A second reason is that when two extremely short syllables occur consecutively, especially when one is measured as twice the duration of the other, a large value of the PVI will result even though the difference may not be very meaningful. A second modification was to exclude the sentence-final syllable from the measurement, as was done by Deterding (2001, 2006). This is because inclusion of the sentence final syllable will influence the measurement of the PVI since sentence final syllables tend to be lengthened (Low et al., 2000).

Another measure, VarcoV (White & Mattys, 2007a), which is the standard deviation of vowel interval duration divided by the mean vocalic interval duration and multiplied by 100, was employed. White and Mattys (2007a) showed that this metric is robust for speech rate variation and is able to discriminate between languages perceived to be different rhythmically as well as between first and second language speakers. Thus, using two different measures would perhaps provide a better insight of the rhythm of MalE and SgE. Measurements of sentence final syllables were also excluded for this measure.

The data was analyzed using PRAAT software (version 4.4.22) by Boersma and Weenink (2006). Two rounds of measurements for the full and reduced vowel sentence sets were performed and a two-tailed Pearson’s correlation revealed a significant result ($r = .91$, $df = 199$, $p < 0.001$), showing a high inter-measurement reliability.

![Figure 1. I went to Universal studios.](image-url)
3 Results

3.1 Full and reduced vowel sentence sets

A total of 350 vowel durations were measured each for the full vowels sentence sets and full and reduced vowel sentence sets for each variety of English, making a total of 700 vowel durations measured for each variety of English. The average PVI values obtained for the full vowel set sentences as well as the reduced vowel set sentences for both MalE and SgE are shown in Table 1.

As can be seen from Table 1, the PVI values obtained for the full vowel and reduced vowel sentence sets were higher for SgE compared to MalE suggesting that MalE is more syllable-timed than SgE. Table 2 reports the results of the paired t-tests conducted.

A Bonferroni-adjusted alpha value of .0063 shows that there was a difference for comparisons of SgE full and reduced vowels for PVI and the value for VarcoV was very close to the level of
significance. The reduced vowel sentence sets comparison between MalE and SgE utterances yielded no significant difference. Bearing in mind that the full vowel sentence set contains words that tend not to be reduced whereas the reduced vowel sentence set contains words with a high possibility of reduction, what was seen was that the SgE speakers tended to reduce their vowels in the reduced vowel sentence set whereas the MalE speakers tended not to. This suggests that there is greater vowel-to-vowel variability in SgE compared to MalE. In comparison, the Low et al. (2000) study that measured the rhythm of Chinese speakers of SgE and speakers of BrE found that the difference between the PVI of the SgE full vowel and reduced vowel set was 6 while the difference for BrE was 34.78. Low’s results show that the BrE speakers tended to have a lot more vowel reduction than the SgE speakers, showing that the rhythm of the BrE speakers tended to be stress-based while the rhythm of the SgE speakers tended to be syllable-based.

3.2 “The boy who cried Wolf” text

Data of 10 speakers each of MalE and SgE were measured reading the *Wolf* passage. For each speaker, 25 utterances (251 vowel durations) were measured, making it a total of 2510 vowel durations measured each for MalE and SgE for this part of the work.

Post hoc comparisons with Bonferroni adjustment show that the rhythm of MalE is significantly different from SgE for both PVI and VarcoV, implying greater variability in the rhythm of SgE compared to MalE for the *Wolf* passage. As can be seen from Table 4, the rhythm of the SgE speakers showed more variability in the successive vowels compared to the MalE speakers as the PVI values for the Singaporean speakers were higher for all the sentences except for ws12 (Wolf sentence) (This gave the boy so much pleasure that a few days later)) and ws17 (overcoming its fear of being (shot)). A closer examination of these two sentences reveals that in these two sentences, there are more content words and thus there is less potential for vowel reduction. In ws12, there are 13 vowels durations that were measured. Of these, only three vowels are likely to be reduced (thE, thAt, A). The vowel in the word “so” is not likely to be reduced as it does not generally have a weak
form. Wells (2008, p. 754) states that there is only occasionally a weak form for “so”; and Jones, Roach, Hartman, and Setter (2003, p. 495) state that “The weak form is used only rarely, and only in casual speech before adjectives and adverbs.” As the reading of the Wolf passage cannot be regarded as casual speech, one must assume that no weak form is possible for “so” in this instance.

Of the nine other vowel durations, four of them are diphthongs, one a long monophthong and four are short vowels. For ws17, nine vowel durations were measured and only one of those vowels has the potential of being reduced (O). Of the remaining vowel durations, two are diphthongs, one a long vowel and five short vowels. It is likely, therefore, that in these two sentences, the SgE speakers tended to have less variation in the successive vowel lengths resulting in a lower PVI value as compared to the MalE speakers, while for the other sentences, as there was greater variation in the length of the successive vowels, the PVI values were higher than those of MalE speakers. The lower VarcoV score for SgE speakers for sentence 17 also shows that there was less vowel length variation among the SgE speakers for this sentence.

A close examination of ws2 in Table 4 shows that three (whO, tO, thE) out of the eight vowel durations measured had potential for vowel reduction. Of the rest of the vowel durations, two were

<table>
<thead>
<tr>
<th>Variety sentence</th>
<th>MalE PVI</th>
<th>MalE VarcoV</th>
<th>SgE PVI</th>
<th>SgE VarcoV</th>
</tr>
</thead>
<tbody>
<tr>
<td>ws1</td>
<td>45.20</td>
<td>34.91</td>
<td>50.19</td>
<td>42.58</td>
</tr>
<tr>
<td>ws2</td>
<td>38.54</td>
<td>27.04</td>
<td>49.18</td>
<td>38.76</td>
</tr>
<tr>
<td>ws3</td>
<td>42.27</td>
<td>31.06</td>
<td>46.00</td>
<td>41.81</td>
</tr>
<tr>
<td>ws4</td>
<td>43.51</td>
<td>35.51</td>
<td>49.84</td>
<td>39.67</td>
</tr>
<tr>
<td>ws5</td>
<td>31.49</td>
<td>28.33</td>
<td>45.80</td>
<td>37.03</td>
</tr>
<tr>
<td>ws6</td>
<td>31.22</td>
<td>49.10</td>
<td>44.46</td>
<td>57.02</td>
</tr>
<tr>
<td>ws7</td>
<td>41.47</td>
<td>42.26</td>
<td>45.49</td>
<td>45.60</td>
</tr>
<tr>
<td>ws8</td>
<td>23.60</td>
<td>68.44</td>
<td>31.43</td>
<td>82.17</td>
</tr>
<tr>
<td>ws9</td>
<td>34.34</td>
<td>30.46</td>
<td>42.31</td>
<td>41.74</td>
</tr>
<tr>
<td>ws10</td>
<td>40.66</td>
<td>32.29</td>
<td>45.73</td>
<td>36.30</td>
</tr>
<tr>
<td>ws11</td>
<td>37.29</td>
<td>22.39</td>
<td>41.39</td>
<td>31.88</td>
</tr>
<tr>
<td>ws12</td>
<td>53.07</td>
<td>40.10</td>
<td>50.07</td>
<td>45.19</td>
</tr>
<tr>
<td>ws13</td>
<td>43.67</td>
<td>31.59</td>
<td>45.86</td>
<td>36.97</td>
</tr>
<tr>
<td>ws14</td>
<td>39.80</td>
<td>39.17</td>
<td>52.13</td>
<td>49.27</td>
</tr>
<tr>
<td>ws15</td>
<td>36.10</td>
<td>26.45</td>
<td>40.70</td>
<td>32.49</td>
</tr>
<tr>
<td>ws16</td>
<td>49.36</td>
<td>39.61</td>
<td>51.64</td>
<td>38.46</td>
</tr>
<tr>
<td>ws17</td>
<td>46.75</td>
<td>41.59</td>
<td>44.90</td>
<td>40.88</td>
</tr>
<tr>
<td>ws18</td>
<td>39.89</td>
<td>28.38</td>
<td>45.21</td>
<td>31.90</td>
</tr>
<tr>
<td>ws19</td>
<td>39.41</td>
<td>32.36</td>
<td>46.90</td>
<td>35.65</td>
</tr>
<tr>
<td>ws20</td>
<td>44.38</td>
<td>55.49</td>
<td>60.32</td>
<td>68.37</td>
</tr>
<tr>
<td>ws21</td>
<td>51.50</td>
<td>44.84</td>
<td>54.78</td>
<td>43.89</td>
</tr>
<tr>
<td>ws22</td>
<td>39.08</td>
<td>31.57</td>
<td>41.79</td>
<td>35.50</td>
</tr>
<tr>
<td>ws23</td>
<td>42.21</td>
<td>30.11</td>
<td>47.09</td>
<td>38.36</td>
</tr>
<tr>
<td>ws24</td>
<td>45.12</td>
<td>33.64</td>
<td>56.51</td>
<td>47.48</td>
</tr>
<tr>
<td>ws25</td>
<td>50.36</td>
<td>59.75</td>
<td>52.69</td>
<td>64.78</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>41.21</strong></td>
<td><strong>37.46</strong></td>
<td><strong>47.30</strong></td>
<td><strong>44.15</strong></td>
</tr>
</tbody>
</table>

*ws* = Wolf sentence; PVI: pairwise variability index; VarcoV: normalised standard deviation of vocalic interval durations divided by the mean vocalic duration
long vowels and three were short ones. In ws5, out of the seven vowel durations measured, three (And, hAve, A) had potential for vowel reduction. Of the remaining four vowel durations, there was a diphthong, one long vowel and two short vowels. A close scrutiny of ws6 shows that out of six vowel durations, one had potential for vowel reduction (thE). Of the remaining five vowel durations, one was a diphthong and four were short front vowels. In ws14, three of the vowel durations measured had the potential for vowel reduction (And, hE, wAz). There were also three short vowels and one back vowel in ws14. In ws20, two of the vowel durations measured had the potential for vowel reduction (tO, thE). In ws20, there were also two diphthongs and two short front vowels.

Table 5 lists the sentences where SgE speakers had a lower PVI value compared to MalE speakers, while Table 6 lists the five sentences from the Wolf passage where the PVI values of the MalE and SgE speakers were the greatest.

Dasher and Bolinger (1982) suggested that the degree of vowel reduction in unstressed syllables make the stressed syllables relatively more prominent in stress-timed languages. Instrumental
work by Low et al. (2000) lends support to the notion that the dichotomy between stress and syllable timing can be attributed to vowel reduction.

If the sentences from Tables 5 and 6 are compared, the pattern that can be seen is that, comparatively, the sentences in Table 6 had a higher proportion of vowel durations with potential for vowel reduction. In these sentences, too, there were also a few long vowel durations (either diphthongs or long, back vowels). As such, it is likely that in these sentences, the SgE speakers had greater variation in their vowel lengths as compared to the MalE speakers who probably had less vowel reduction and thus had less vowel to vowel variation. As a result of this, the average PVI value obtained for the MalE sentences was lower, implying that the rhythm of the MalE speakers was more syllable-timed compared to the SgE speakers.

Pearson’s correlation was done to check if there was a correlation between the PVI and the VarcoV results. PVI values of the five greatest differences in PVI values were compared with the difference in the VarcoV results of the same utterances. A 1-tailed Pearson’s correlation test did not show a significant relationship ($r = .1401, df = 4, p = 0.41$). When the PVI and VarcoV values of the two utterances where SgE had a lower PVI value were added, a significant correlation was observed ($r = .8089, df = 6, p = 0.014$). The weak correlation between the PVI and VarcoV scores of the MalE and SgE speakers for sentences with the greatest difference in PVI scores implies that when there was a lot of difference in the pairwise vowel durations, there was little variation in the standard deviation of vowel durations between the two varieties of English. Table 7 lists the difference between the SgE and MalE PVI and VarcoV values.

**3.3 Results of the measurement of “My most memorable holiday”**

The speech of 10 speakers each of Malaysian and Singapore English was measured. There were equal numbers of males and females. For each speaker, six utterances containing eight syllables each were measured, making it a total of 480 vowel durations measured for each variety of English. Although Deterding (2001) measured utterances with a minimum of six syllables, it was decided to go with eight syllables, following the number of syllables for the full and reduced vowel sentence sets (Low, 1998). A total of 480 tokens for each variety of English were measured for this part of the study. Table 6 lists the average PVI value from the six utterances obtained for each speaker.

The lower average PVI value obtained for MalE in the measurement of “My most memorable holiday” shows that indeed, in spontaneous speech, there is less variation in the successive vowels of Malaysians speakers compared to the Singaporeans. This suggests that the rhythm of the
Malaysian speakers was more syllable-timed compared to the Singaporeans. This is confirmed by the VarcoV scores where the MalE speakers had a lower score, showing that there was less variability in the duration of vowels compared to the duration of the SgE vowel in “My most memorable holiday.”

Post hoc Bonferroni adjustment showed that the results were significantly different for PVI but not for VarcoV.

4 Discussion

4.1 Summary of findings

Table 10 summarizes the main findings obtained for all the measurements done for rhythm. As can be seen from the PVI values obtained, the rhythm of the SgE speakers generally exhibited more qualities of stress timing compared to MalE speakers in the read sentences and Wolf text, especially when there is greater potential for vowel reduction in the utterances. This difference in rhythm is also seen in their natural conversation. The VarcoV results for the reduced vowel sentence set and “My most memorable holiday” show that there is no significant difference in the distribution of the vowel duration.
A significant difference was found between the SgE full and reduced vowel sentence sets for both PVI and VarcoV, whereas no significant difference was found between the MalE full and reduced vowel sentence sets for either of these indices. The PVI and VarcoV values for SgE suggest that SgE exhibits more properties of stress-timing compared to MalE. This is so despite the similarity of the languages that influence both MalE and SgE. Morais (2000) for MalE and Low and Brown (2005) for SgE listed the languages that influence both the varieties of English as Malay, Chinese dialects like Mandarin, Hokkien, Cantonese and Hakka, and Indian languages including Tamil, Malayalam, Punjabi, Telegu and Gujarati. The difference is that for MalE, there are the languages spoken in Sarawak and Sabah like Iban and Kadazan which affect mainly those who originate from East Malaysia. However, as all the MalE subjects selected for this study are from West Malaysia, the languages from East Malaysia would not be a strong influence.

Comparing the PVI values obtained for MalE and SgE for both the full vowel and reduced vowel sentence sets reveals interesting trends. There is no significant difference between MalE and SgE for the full vowel sentence set. Similarly, no significant difference was found between MalE and SgE for the reduced vowel sentence set. In contrast, a significant difference was found between the full and reduced vowels sentence sets in SgE but not in MalE. This differs from Low’s (1998) finding where no significant difference between the full and reduced vowels sentence sets for SgE was seen. It should be noted, however, that Low’s subjects were all ethnically Chinese and in the present study the PVI measurement has been modified from Low’s original method, namely having a minimum value of 30 msec. and not including the sentence final syllables.

The fact that a significant difference was found in the PVI values obtained for the measurement of the Wolf passage as well as “My most memorable holiday” implies that MalE is more syllable-timed compared to SgE. It should be noted that as past research has found that BrE is stress-timed and SgE is syllable-timed (Low, 1998), this present study shows that if the difference between stress-timed languages and syllable-timed languages is seen as part of a continuum, then MalE is even more syllable timed than SgE. However, it should be noted that whereas the difference in the PVI value between BrE and SgE was 30 (Low, 1998, p. 40), in this study, the difference between MalE and SgE is less than 10. Even though there was a very significant difference found when the VarcoV results for the Wolf text for MalE and SgE were compared, there was no significant difference found between SgE and MalE in the VarcoV results for “My most memorable holiday.” This shows that for the natural conversation,
although there was a difference in the pairwise vowel-to-vowel variation between MalE and SgE, there was no significant difference in the variance of the distributions of the vowel durations. While acknowledging the limitation of the findings due to the fact that much of the past research has not shown that speech that is perceptually isochronous is acoustically isochronous (Dilley, Wallace, & Hefner, 2012), the difference in the PVI results between SgE and MalE could perhaps prove useful in the classification of the language and possibly could be accounted for by the difference in the educational and social environments of Malaysia and Singapore.

4.2 Implications for research on rhythmic patterning

Two rhythmic indexes were used, PVI and VarcoV. While VarcoV returned fewer significant differences between the two varieties, PVI was able to tease out significant differences. This shows that PVI is a more sensitive and robust measurement of inter-varietal differences in rhythmic patterning. This study is limited in that %V, which has also been found previously to differentiate rhythmic classes, was not used. Future research should consider this.

VarcoV essentially measures standard deviation of vowel durations in the entire utterance whereas PVI measures successive vowel durations. It was found that when PVI values differed most greatly between MalE and SgE in the read passage, no correlation was found between PVI and VarcoV. This suggests that the rhythmic differences between two varieties of a language are more effectively captured by considering successive vowel durations rather than overall standard deviations in vowel durations.

This study was based on instrumental work only. Perceptual data could also be used in future studies to complement acoustic findings for purposes of triangulation as well as to test if listeners are able to distinguish these two varieties of English based only on speech rhythm.

Funding

This work was supported by the University of Malaya Research Fund [F0226-2007B].

References


Appendix

Full vowel set

1. John came back through France last Sunday.
2. Don seemed quite cross with John last week.
3. Paul drives past huge towns by highway.
4. Jane gets four by post each Thursday.
5. Grace works through huge mounds each Friday.

Reduced vowel set

1. John was sick of Fred and Sandy.
2. Don was across at Jonathan’s.
3. Paula passed her trial of courage.
4. Jane has four to last the winter.
5. Grace was tired of Matthew Freeman.

The boy who cried wolf

There was once a poor shepherd boy who used to watch his flocks in the fields next to a dark forest near the foot of a mountain. One hot afternoon, he thought up a good plan to get some company for himself and also have a little fun. Raising his fist in the air, he ran down to the village shouting “Wolf, Wolf.” As soon as they heard him, the villagers all rushed from their homes, full of concern for his safety, and two of his cousins even stayed with him for a short while. This gave the boy so much pleasure that a few days later he tried exactly the same trick again, and once more he was successful. However, not long after, a wolf that had just escaped from the zoo was looking for a change from its usual diet of chicken and duck. So, overcoming its fear of being shot, it actually did come out from the forest and began to threaten the sheep. Racing down to the village, the boy of course cried out even louder than before. Unfortunately, as all the villagers were convinced that he was trying to fool them a third time, they told him, “Go away and don’t bother us again.” And so the wolf had a feast.