

On the Syllable Timing in Taiwan English

Hua-Li Jian

Department of Foreign Languages and Literature
National Cheng Kung University, Tainan, Taiwan
hljian@mail.ncku.edu.tw

Abstract

In this paper the syllable timing of Taiwan English is compared to that of American English. A variability index reflecting vowel length over utterances is computed based on acoustic measurements of Taiwan and American English speech. The results show that successive vowel durations are more equal in Taiwan English than in American English. Further, vowel durations are generally larger in Taiwan English than in American English. These observations suggest that unlike American English, which is stress-timed, Taiwan English is not stress-timed.

1. Introduction

A vast body of literature exist addressing the rhythm and timing in British English (BE) and American English (AE), which are both known to be stressed-timed. However, very little has been written about the rhythm and timing of Taiwan variety of English (TE), although some studies have addressed other South Asian varieties of English – especially Singapore (SE) and Malaysian English (ME) since English is one of the official languages in Singapore (see for example the studies by Low, Grabe and Nolan [7], Brown [4], Platt and Weber [10], Tay [15] and Torque [16]). All these studies describe Singaporean English as being syllable-timed and thus staccato sounding. In this study we investigate Taiwan English from the same viewpoint. English is not an official language in Taiwan though everyone is now taught English from an early age. Taiwan English is a distinct variety of English, although American pronunciation is generally taught throughout the island. Factors influencing the Taiwan variety of English include Japanese-style English brought over during the second world war occupation, mainland Chinese English mainly brought over just before 1945, British textbooks used before the introduction of the Kenyon and Knott general American pronunciation system in 1969, persistent defective pronunciation of English in classrooms, and effects from local languages – especially Mandarin and Southern Min (see Chung [5]). Our hypothesis is that the Taiwan English variety shares some characteristics with the Singaporean English due to similar influences such as the Chinese language. The study attempts to uncover whether Taiwan English is syllable-timed. This is achieved by the means of an acoustical analysis of speech samples. A pair-wise variability index, proposed by Low, Grabe and Nolan [7], is used to quantify the timing characteristics of the vowels.

2. Background

It has long been assumed that there is near equal interval between stresses in stress-timed languages and that there are near equal intervals between successive syllables in syllable-

timed languages (Pike [9]). Most of the early studies were not based on acoustic studies, and Couper-Kuhlen [6] claims that this isochrony is a perceptual phenomenon. Studies also show that isochrony cannot be related to constant inter-stress intervals in stress-timed languages (Strangert [14]). Further, studies show that successive syllable durations vary greatly in syllable-timed languages (Pointon [11]). Other studies show that stress-timed languages are no more regular in inter-stress intervals than syllable-timed languages (Roach, [13]).

Another problem is that most studies consider English as a stress-timed language in general, avoiding the complexity of all the varieties of English such as Singaporean and Malaysian English (Brown [4]), and studies into timing in Taiwan English are next to non-existent.

One of the reasons for the difficulties of contrasting the different languages lies in the choice of timing syllable durations. Studies by Bertinetto [1] and others have shown that vowel duration is probably more important than syllable duration, and this is one of the reasons for considering vowel duration in this paper. Further, syllable boundaries are hard to identify in English.

3. Method

3.1. Subjects

Five Taiwanese English (TE) and five American English (AE) speakers took part in the recording task. All subjects live in Taiwan. The native Taiwanese TE speakers were students at the National Cheng Kung University in Tainan. The native American AE speakers were mostly teachers in the Tainan and Kaohsiung areas in the south of Taiwan. The TE speakers were mostly in their early twenties. Two of the AE are in their forties and the rest in their mid twenties. None of the subjects reported having any articulation disability.^f

3.2. Materials

It was decided to adopt material with a similar structure to that described in Low, Grabe and Nolan [7]. The TE and AE subjects were asked to read a list of 10 sentences. Five sentences consist of a mixture of full and potentially reduced vowels (i.e., the reduced vowel set) and five sentences comprise only full vowels (i.e., the full vowel set). See Table 1 as an example (taken from Low, Grabe and Nolan [7]).

The sentences were shuffled into pseudorandom order and presented to the speakers without context and fillers.

3.3. Recording procedure

The subjects were recorded in an office in the National Cheng Kung University using a Shure KSM32 studio microphone and a portable mixer with a built-in microphone amplifier.

Table 1: *Example of the full vowel set and reduced vowel set.*

Full vowel set	John came back through France last Sunday
Reduced vowel set	John was sick of Jack and Sandy

The microphone was placed approximately 20 cm in front of the speakers. The subjects were given time to read the material prior to the recording and the subjects were manually informed to repeat utterances when mistakes were made. The material was digitally recorded onto a Minidisk using a stationary Sony deck. The digitally recorded material was subsequently transferred to audio files stored on a personal computer.

3.4. Post-processing and analysis

The speech material comprising 100 utterances of which 50 utterances were produced by the TE speakers, and the remaining 50 were produced by the AE speakers. The digital recording was partitioned so that each audio-file consisted of one sentence. The open source analysis package PRAAT by Boersma and Weenink [2], [3] was used to segment and label the utterances. The spectrogram feature and formant analysis tool in PRAAT was used to select the vowel boundary, which is not a trivial task. See Peterson and Lehiste [8] for a general discussion on segmentation of vowels by the means of spectrograms.

PRAAT stores segmentation information in separate text files. A small script was written to extract the relevant vowel information from the segmentation file so that it could be exported into a Microsoft Excel spreadsheet. Subsequent calculations were performed in Excel and statistical significance tests were conducted by the means of the data analysis plug-in for Excel.

In order to compare consecutive vowels amongst the TE and AE speakers the pair-wise variability index p was computed (Low, Grabe and Nolan [7]) which is given by

$$p = \frac{2}{m-1} \sum_{k=1}^{m-1} \frac{|d_k - d_{k+1}|}{d_k + d_{k+1}} \quad (1)$$

where d_i is the duration of vowel i and m is the total number of vowels. For each neighbouring vowel-pair the absolute difference in vowel duration is taken and divided by the average duration of the two vowels in order to get a normalised difference. These normalised duration differences are summed and divided by the total number of differences.

4. Results and discussion

The mean pair-wise variability indices (PVI) across the subjects for the four groups are presented in Figure 1 - namely TE full vowels, reduced vowels, AE full vowels and reduced vowels.

Table 2 lists the PVI values for the individual subjects and the overall mean PVI and the overall vowel duration. A repeated measures two-way analysis of variance (Anova) was used in the analysis of the data, where variety (TE and AE) and set (full, reduced) are independent variables and the PVI

is the dependent variable. The PVI values for each of the five speakers in each category are used to compute the statistics.

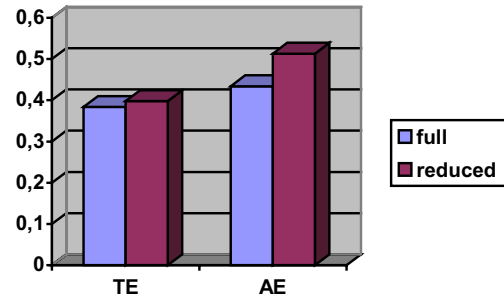


Figure 1: *The mean pair-wise variability indices for TE full vowels, TE reduced vowels, AE full vowels and AE reduced vowels.*

Table 2: *The mean pair-wise variability indices for the subjects including TE full vowels, TE reduced vowels, AE full vowels and AE reduced vowels.*

Variety Set	TE		AE	
	Full	Reduced	Full	Reduced
subject 1	0.428	0.398	0.529	0.527
subject 2	0.394	0.432	0.417	0.471
subject 3	0.353	0.445	0.509	0.579
subject 4	0.414	0.409	0.532	0.651
subject 5	0.334	0.426	0.394	0.547
Mean	0.384	0.398	0.434	0.513

Table 3: *The mean vowel duration (in milliseconds ms) for the subjects organized into TE full vowels, TE reduced vowels, AE full vowels and AE reduced vowels.*

Variety Set	TE		AE	
	Full	Reduced	Full	Reduced
subject 1	144	131	184	152
subject 2	171	181	154	110
subject 3	131	120	124	96
subject 4	140	140	130	116
subject 5	137	150	131	102
Mean	145	145	138	114

Note that the values are in the range 0 and 1 instead of the range 0 and 100 (percentage) as used by Low, Grabe and Nolan [7]. By inspecting the data it is clear that there is a significant difference in PVI across the TE vowels and the AE vowels ($F(16,1)=23.6$; $p=0.0$), as the mean PVI for TE vowels is below 0.4 and the mean PVI for the AE is above 0.4. Thus, there is a clear difference in vowel timing between Taiwan English and American English. Further, for TE there is very little difference in mean PVI when comparing full and

reduced vowels, while there is a larger difference in PVI for AE when comparing full and reduced vowels. However, the PVI across full and reduced vowels are significantly different for both TE and AE ($F(16,4)=6.29; p=0.02$). These observations adhere to the theory, consistent with the study into syllable timing in Singapore English speech (Low, Grabe and Nolan [7]), and confirm our expectations, namely that successive vowels in TE remain more or less constant while it varies hugely in AE. TE is therefore not a stress-timed variety of English, but more likely to be syllable-timed.

Table 3 shows the mean vowel duration for each of the five subjects and the overall vowel duration for all the subjects organized into TE full vowels, TE reduced vowels, AE full vowels and AE reduced vowels. The table reveals that there is a large variability in durations across different subjects. This variability is probably the result of individual and different speaking rates and speaking style, as the long and short durations are consistent with the different speakers. For example, TE subject 2 has both the largest mean full vowel duration and mean reduced vowel duration, while TE subject 3 has both the shortest mean full and reduced vowel duration. The same pattern can be observed for the AE subjects, where AE subject 1 has both the largest mean full and reduced vowel duration, while subject 3 has both the shortest mean full and reduced vowel duration.

For the Taiwan English subjects the mean of the full vowels varies the least in relative terms, with means from 131 to 171 ms (a difference of 40 ms, or 30.5 %), while the mean reduced vowel duration varies from 120 to 181 ms (a difference of 61 ms, or 50.8 %). The relative variability is even larger for the American English speakers. Mean full vowel duration ranges from 124 to 184 ms (a difference of 60 ms, or 48.4 %), and mean reduced vowel duration ranges from 96 to 152 ms (a difference of 56 ms, or 58.3 %).

The overall mean vowel duration for TE is larger than the mean vowel duration for AE. For TE the mean vowel duration is 145 milliseconds (ms) for both full and reduced vowel set (equal), which is a relatively long time interval. For AE subjects the mean vowel duration varies. Full AE vowels have a mean duration of 138 ms and reduced AE vowels have a mean duration of only 114 ms. AE reduced vowels which have the largest PVI also have the smallest mean vowel duration, and the TE vowels which have the smallest PVI have the longest duration. It is difficult to control the subjects' speaking rate in this type of phonetic study, and such problems can be solved by increasing the number of subjects and compensated by employing statistical techniques.

5. Conclusion

In this study the acoustic correlates of rhythm in Taiwan English and American English have been compared. Pair-wise variability indices for speech samples taken from both varieties of English were computed and showed that there was significantly smaller variation in successive vowel lengths in Taiwan English compared to that of American English. American English is known to be stress-timed, while the acoustic measurements described in this paper suggest that Taiwan English is syllable-timed.

Acknowledgements

The author would like to thank the assistants **Hoi Wei-Yei, Tan Te-Kong, and Chen Chia-Sian** for assisting in the overwhelming job of corpus labeling and segmentation. I am

also grateful for the subjects that took the time to conduct the recording. This work is supported by the National Science Council of Taiwan grant **NSC 91-2411-H006-020**.

6. References

- [1] Bertinetto, P. M., 1977. Syllabic blood, ovvero l'italiano come lingua ad isocronismo sillabico *Studi di Grammatica Italiana*, 6, 69-96.
- [2] Boersma, P.; Weenink, D., 2003. *The PRAAT homepage*, <http://www.fon.hum.uva.nl/praat/>
- [3] Boersma, P.; Weenink, D., 2001. PRAAT, a system for doing phonetics by computer, *Glott International* 5(9/10): 341-345
- [4] Brown, A., 1988. The staccato effect in the pronunciation of English in Malaysia and Singapore. In J. Foley (Ed.). *New Englishes, The case of Singapore*. Singapore, Singapore University Press, 115-128.
- [5] Chung, K., 2002, *The sounds and allophones of Taiwan English*, <http://ccms.ntu.edu.tw/~karchung/intro%20page%202027.htm>
- [6] Couper-Kuhlen, E., 1990. Discovering rhythm in conversational English: Perceptual and acoustic approaches to the analysis of isochrony. *KonTRI working paper no. 13*, University of Constance, Fachgruppe Sprachwissenschaft.
- [7] Low, L. E.; Grabe, E.; Nolan, F., 2000. Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English, *Language and Speech*, 43(4), pp.377-401.
- [8] Peterson, G. E.; Lehiste, I., 1960. Duration of syllable nuclei in English, *Journal of the Acoustical Society of America*, 32, pp.693-703.
- [9] Pike, K., 1945. *The intonation of American English*, Ann Arbor, University of Michigan Press.
- [10] Platt, J. T.; Weber, H., 1980, *English in Singapore and Malaysia : studies features and functions*, Kuala Lumpur, Oxford University Press.
- [11] Pointon, G. E., 1980. Is Spanish really syllable-timed? *Journal of Phonetics*, 8, 293-304.
- [12] Ramus, F.; Nespor, M.; Mehler, I., 1999. Correlates of linguistic rhythm in the speech signal, *Cognition*, 72, pp.1-28.
- [13] Roach, P., 1982. On the distinction between "stress-timed" and "syllable-timed" languages, In D. Crystal (Ed), *Linguistic Controversies*, London, Arnold.
- [14] Strangert, E., 1985. Swedish speech rhythm in a cross-language perspective. Umeå. *Umeå Studies in Humanities*, 69.
- [15] Tay, M. W. J., 1982, The phonology of educated Singapore English, *English worldwide* 3(2), pp.135-145.
- [16] Tongue, R. K., 1974, *The English of Singapore and Malaysia*, Singapore Eastern Universities Press.