Mandarin Rhythm: An Acoustic Study*

Hua Lin Qian Wang
Department of Linguistics, University of Victoria, Victoria, BC, V8P, 5C2, Canada
Luahin@uvic.ca wangqian@uvic.ca

Abstract

Is Mandarin Chinese a syllable-timed language? Based on auditory impression, traditional analyses say it is. However, this question has rarely been investigated from the perspective of acoustic phonetics. Following Ramus et al. (1999) and Grabe and Low (2002), we measured the four rhythmic correlates: vowel percentage, consonant standard deviation, rPVI and nPVI in passage readings and conversations of native Mandarin speakers. Except those for nPVI, the results confirmed the impression that Mandarin is a syllable-timed language.

Keywords
Rhythm, syllable-timed, stress-timed, acoustic phonetics, consonant duration, vowel duration, Mandarin Chinese, suprasegmental

1 Introduction

It is common practice for linguists to classify languages into two rhythmic groups: stress-timed (StrT) and syllable-timed (SylT). Typical stress-timed languages include English and Dutch while Spanish and Italian are typical syllable-timed languages. Since Pike first named the dichotomy (1945) based on Lloyd-James' (1940) observation of a distinction between 'machine-gun' (SylT) and 'Morse-code' (StrT) languages, studies have

* This work has been supported in part by the project 'The Formal Syntactic Theory and A Study of Some Essential Chinese Syntactic Issues' (05JJD740184) from the Key Project Schema of the Advanced Research Bases in Humanities and Social Sciences, the Ministry of Education, China. We also wish to thank the reviewers of this paper and the audience at the 13th International Conference on Chinese Linguistics, Leiden, the Netherlands, June 2005, where an earlier draft of this paper was presented.

been carried out to pinpoint exactly what this dichotomy means phonologically, acoustically, and perceptually (see Bolinger 1965, Borzone de Manrique and Signorini 1983, Deterding 2001, Roach 1982, 2004, and Wenk and Wiolland 1982). Many earlier studies (e.g., Roach 1982, and Dauer 1983, 1987) were quick to dismiss the distinction after not being able to find evidence for the syllable isochrony in so-called syllable-timed languages or the lack of it in the stress-timed language. However, recent psycho-linguistic studies (Bahrick and Pickens 1988, Bertoncini and Mehler 1981, Christopher and Dupoux 1996, Mehler *et al*. 1988, Nazzi *et al*. 1998, Ramus 2000, and Ramus *et al*. 2003) have been able to confirm that both adults and infants are able to perceive rhythmic differences when presented with speech signals of these two types of languages, even if segmental information is reduced or removed. On the acoustic-phonetic front, two recent experimental studies, one by Ramus *et al*. (1999) and the other by Grabe and Low (2002), have established methodologies that contribute to the understanding of the dichotomy in concern.

Previous studies have investigated the rhythm patterns of more than two dozen of the world's languages including the StrT languages such as English, Russian, and Arabic, the SylT languages such as Spanish, French, Telugu and Yoruba, and the unclassified languages such as Greek, Thai, Italian, Japanese, and Tamil. However, there is an interesting gap: little attention has been paid to Mandarin Chinese. Our study attempts to help fill this gap by focusing on the rhythmic properties of Mandarin Chinese. Adopting both Ramus *et al*. and Grabe and Low's methodologies, we measured four rhythmic qualities of Mandarin: vowel quantity, consonant variance, *rPVI* and *nPVI*.

Traditionally, Mandarin Chinese is considered a SylT language. If it is indeed SylT, Mandarin Chinese should yield measurements that are comparable to those of typically SylT languages such as French and Spanish. The purpose of the present study is thus three-fold: (1) to measure Mandarin Chinese in terms of the four acoustic properties identified as rhythmic property indicators by Ramus *et. al* (1999) and Grabe and Low (2002); (2) to decide how Mandarin Chinese rhythm should be classified in terms of the dichotomy (Ramus *et al*) or the continuum (Grabe and Low); (3) if and how style affects rhythm. To achieve our research goals, we collected and analyzed data from both passage reading and conversation. The latter was never investigated in previous studies on rhythm.

2 Methodology and Hypotheses

As mentioned above, we adopted methodologies developed in two previous studies. Below we will provide a brief review of both starting with Ramus *et. al.*'s. (1999).

2.1 *Ramus et al.'s Method*

In Ramus *et al.*'s (1999) study, three groups of languages were investigated: the StrT group of English, Polish and Dutch, the SylT group of French, Spanish, Italian and Catalan, and the so-called mora-timed group with Japanese as the only member. Four speakers from each of the eight languages read five short news-like declarative statement sentences. The sentences were segmented into either vocalic or intervocalic intervals. Vocalic intervals were defined as existing between the onset and the offset of one or a sequence of vowels. Intervocalic intervals constituted the part between the onset and the offset of one consonant, or a sequence of consonants. Two crucial variables were identified (Ramus *et al*, 1999: 272) and calculated:[1]

(1) Vowel Quantity: the proportion of vocalic intervals within the sentence; the sum of the vowel intervals divided by the total duration of the sentence.
(2) Consonant Variance: the standard deviation of the duration of the consonant intervals within each sentence.

It was found that the rhythmic group factor had a significant effect on the two variables that they identified. StrT languages shared higher consonant variance values and lower vowel percentages while SylT languages shared the reverse, lower consonant variance values but higher vowel percentages. The so-called mora-timed languages had the highest value of vowel quantity and lowest consonant variance. In other words, Ramus *et al*.'s measurements of vowel percentage and consonant standard deviation appear to 'support the idea that the standard rhythm classes are meaningful categories, that not only appeal to intuitions about rhythm, but also reflect actual properties of the speech signal in different languages' (Ramus *et al*., 1999: 387).

2.2 *Grabe and Low's Method*

Grabe and Low studied 18 languages including StrT languages such as British English, German, Dutch and Thai, SylT languages such as Tamil, French, Spanish, and Singapore

---

[1] The study also calculated 'vowel variance', the standard deviation of the duration of the vowel intervals within each sentence, but found it not a reliable correlate for rhythmic distinction. We will disregard it in this study as well.

English, one mora-timed language, Japanese, and two ‘mixed’ languages, Polish and Catalan. The study also examined seven previously unclassified languages: Estonian, Greek, Luxembourg, Malay, Rumanian, Welsh, and, of particular interest to us, Mandarin. One speaker from each language was recorded reading the passage “The North Wind and the Sun” in his/her respective native language.

Like Ramus *et al.*, Grabe and Low (2002) also measured the duration of vowel and consonant intervals. However, they differ from Ramus *et al.* in that rather than vowel proportion and consonant standard deviation, Grabe and Low calculated differences between successive pairs of the intervals using a variance measuring formula called the *Pairwise Variability Index* (*PVI*; see Figures 1 and 2 below). According to the authors, the formula compares “the level of variability in successive measurements” (Grabe and Low, 2002: 519). The formula consists of two parts, the ‘raw PVI’ and the ‘normalized PVI’. The former is used for calculation of the consonant intervals while the latter the vowel intervals. Figure 1 below is the formula for the consonantal raw PVI (*r*PVI) where *m* is the number of vocalic or intervocalic intervals in a passage of speech and *d* is the duration of the k^{th} interval.

$$rPVI = \left[\sum_{k=1}^{m-1} \left| d_k - d_{k+1} \right| \ / (m-1) \right]$$

Figure 1 *r*PVI formula

Grabe and Low found that vocalic duration is directly correlated with speaking rate. In order to account for speaking rate effect that can be present, a normalized version of the PVI (*n*PVI) was also adopted for vocalic PVI. Figure 2 provides the normalized PVI:

$$nPVI = 100 \times \left[\sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{(d_k + d_{k+1}) / 2} \right| \Big/ (m-1) \right]$$

Figure 2 *n*PVI formula

The *n*PVI is achieved by first calculating the duration difference between each pair of successive intervals, then dividing it by the mean duration of the pair, and taking the absolute value. The results for each pair are then summed and divided by the number of differences. The final output is multiplied by 100 to avoid fractional values.

Grabe and Low hypothesized that StrT languages should exhibit higher vocalic and intervocalic variability. SylT languages, on the other hand, should have more equal vowel and consonant intervals. Their results confirmed their hypotheses. The calculation of *r*PVI and *n*PVI allowed for identification of the difference between languages in terms of rhythmic properties. It is important to note, however, that Grabe and Low's study does not support a strict categorical distinction between StrT and SylT languages, rather, that languages are either more or less StrT or SylT. That is, there is a continuum from clearly StrT languages to clearly SylT languages.

A notable feature of Grabe and Low's study is that it is the only acoustic-phonetic study on rhythm that has addressed Mandarin. The study found that Mandarin has the lowest vocalic *n*PVI values of all the 18 languages and its *r*PVI is also quite low. According to Grabe and Low, Mandarin clearly patterns with French and Spanish in both vocalic and intervocalic PVI. This supports the traditional perception that Mandarin is a SylT language.

2.3 *Hypothesis*

As mentioned earlier, we adopted Ramus *et al.*'s method in our study, focusing on vowels and consonants. Specifically, we measured proportions of vowel intervals and standard deviation of consonant intervals. We differed however from Ramus *et al*. in three ways. First, we had three men and three women as our subjects while Ramus *et al*. had four *women* speakers of each language. Next, we had *passage reading* as well as *conversation* in our data collection and analysis. We believe that both are more natural than Ramus *et al*.'s individual declarative sentences. Finally, we focused on Mandarin Chinese which was not among the languages studied in Ramus *et al.*

We also adopted Grabe and Low's method in our study. One major problem with Grabe and Low is that it has only one speaker from each language (see Asu and Nolan 2005). Our study differs from Grabe and Low's in three important ways, one of which is that we had six speakers, an equal number of each gender. Another difference between Grabe and Low and us is that our Mandarin speakers are from Northern China, speaking Standard Mandarin, while Grabe and Low's speaker spoke Singapore Mandarin[2]. The fhrid

[2] There have been studies that point out crucial differences between Singapore Mandarin and Standard Mandarin due to influences from other Chinese dialects. See more discussion on this in Section 5.

difference is that we collected and analyzed both passage reading and conversation while Grabe and Low had only passage reading.

Based on these methods, we want to investigate, instrumentally, the syllable-timed characteristic of Mandarin. Specifically we hypothesize that (1) Mandarin Chinese has a vowel duration and consonant standard deviation comparable to those found in typical SylT languages such as Spanish and French, as reported in Ramus et al's study; (2) Mandarin has a vocalic and an intervocalic PVI comparable to those of typical SylT languages, such as French and Spanish in Grabe and Low's study.

3 The Experiment

3.1 Subjects

We recruited six native Mandarin speakers, three men and three women, who were from Northern China, speaking Standard Mandarin. At the time of the data collection, they were either an undergraduate or graduate student at the University of Victoria (UVic). This means that their English levels were above the intermediate levels as they had previously been able to achieve the UVic admission requirement of 575 or higher on the TOEFL tests.

3.2 Recording

The recording took place in the sound-treated booth in the Phonetic Laboratory in UVic's Department of Linguistics. It consisted of two parts: passage and conversation recording. For passage recording, the subjects read the Chinese version of the passage the *North Wind and the Sun* from Lee and Zee (2003). Before the recording, the subjects were given as much time as they needed to read the passage to become familiarized with it. Then each subject was individually recorded six inches in front of a condenser microphone (AKG C1000S). Recordings were digitized on a computer running Windows XP, which was located in an adjacent room. The recordings were saved as uncompressed wav files.

Following the passage reading, a conversation between the subject and one of the authors was recorded in the same environment. Before the recording, the author had a casual warm-up chat with the subject to ensure that the subject became comfortable speaking with her. During the chat, several topics of interests to the subject were identified and used in the interview-like conversation. For instance, one of the subjects was a serious golf player. In this case, the conversation was oriented around the topic of golf. Using this technique, a reasonably large quantity of quite natural, connected speech was produced by the subject. A total of approximately eight minutes of conversation was obtained for each subject.

3.3 *Segmentation and Analysis*

To ensure objectivity of segmentation, a trained phonetician, who was not otherwise associated with the research, was hired to do the segmentation of the recorded speech. All the passage files and approximately 30 seconds of every conversation extracted from the original recording were segmented. Relying on both visual and audio cues, the phonetician carried out the segmentation on computer using the acoustic-analysis software Praat (Version 4.2.31).

Segmentation was conducted under the assumption that anything on the spectrogram with glottal striations (vocal pulses) and well-defined formants were labeled as a vowel, indicated as V in the CV annotation tier. Postvocalic glides /w, y/ typically appeared to be a part of adjacent vowels and no clear boundary could be determined. Thus, we chose to segment them as part of the preceding V. However, reduced amplitude is usually a cue to prevocalic glides /w, y/ and as a result these were segmented as Cs. Nasals were segmented as consonants (C) when there was a discernible amplitude difference between them and vowels. However, intervocalic nasals were sometimes labeled as part of the V interval when they were unidentifiable from the neighboring vowels. Consonants such as stops shown as gaps and bursts, fricatives shown as high frequency aperiodic energy, and affricates were clearly identifiable from the spectrogram and were labeled as Cs. If there was a gap that bore no indication that it was one sound or another, it was marked as B for breath. Any two consonants split by a B were combined into the same consonant interval in calculation by subtracting the duration of the B. The same approach was used for vowels split by a B.

In summary, there were a total of three labels in the segmentation: V for a vowel or a sequence of vowels, C for a consonant or a sequence of consonants and B for breath and other non-linguistic sounds. See Figure 3 below.

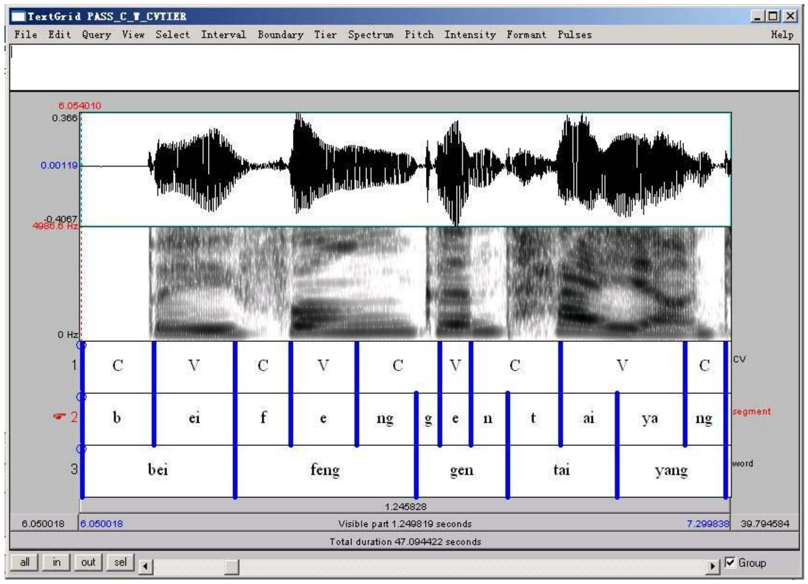

Figure 3. Segmentation of consonant and vowel intervals

After the segmentation, a script was written in Praat to measure the duration of each C or V interval. The results of the measurements and calculations were saved into Excel files. Statistical analysis was then conducted using the Statistic Package for Social Sciences (SPSS) software.

4 Results

Table 1 gives the overall mean vowel percentage, consonant standard deviation, *n*PVI and *r*PVI values.

Table 1 Overall mean results

	%V (Std.)	ΔC (Std.)	*n*PVI (Std.)	*r*PVI (Std.)
Passage	56.23 (2.46)	4.50 (.39)	48.72 (6.30)	52.64 (4.51)
Conversation	56.07 (2.11)	5.59 (.87)	50.51 (4.68)	55.50 (7.98)
Overall	56.15 (2.19)	5.04 (.86)	49.61 (5.38)	54.07 (6.356)

And Table 2 gives individual results for the four measures.

Table 2 Results for the six individuals performing two tasks

Subj	Pass		Conv		Pass		Conv	
	%V	ΔC	%V	ΔC	rPVI(ms)	nPVI	rPVI(ms)	nPVI
1	54.0	4.5	55.0	7.0	50.06	57.78	68.56	51.88
2	56.4	4.2	54.6	4.8	53.07	47.30	49.31	48.66
3	56.9	5.0	56.8	5.5	60.16	46.28	49.97	47.48
4	52.7	4.7	53.3	5.4	50.76	53.84	56.05	44.21
5	58.5	4.7	57.8	6.1	54.70	39.75	60.76	53.47
6	58.9	3.9	58.9	4.7	47.11	47.35	48.34	57.35

As mentioned earlier, we made four sets of calculations. The first two, the mean vowel quantity (%V) and the mean consonant standard deviation (ΔC) were based on Ramus *et al*. The other two, the *n*PVI and the *r*PVI were based on Grabe and Low. Note that as in Ramus *et al*., ΔC figures are given in the x*100 format. That is, if the ΔC is 0.04500, it is reinterpreted as 4.5000 (=0.04500*100).

Recall that the main purpose of this study is to see where Mandarin Chinese, as it is spoken in China, stands in the rhythmic classification of the world's languages. To that end, we need to compare our figures with those of Ramus *et al.*'s and Grabe and Low's. In terms of mean vowel percentage, our figure decisively indicates that Mandarin is not to be grouped with a StrT language like English, which had the smallest vowel percentage of all the languages in Ramus *et al.* In fact, our average %V figure for Mandarin, 56.15%, is even larger than the largest value found in Ramus *et al.,* 53.1% for Japanese. We will discuss this finding in the next section. In terms of consonant standard deviation (ΔC), our figure in Table 1 definitely places Mandarin Chinese among the SylT language. It ranks between the StrT languages of Catalan (4.52) and French (4.39) in Ramus *et al*. Based on the two indices of %V and ΔC, our first hypothesis that Mandarin Chinese would pattern with other SylT language, in respects to %V and ΔC, is supported.

Now let's compare our two PVIs with Grabe and Low's. Our *r*PVI value at 54.07 is quite comparable to Grabe and Low's for Singapore Mandarin at 52.0, and it places Mandarin Chinese close to and slightly above the typical SylT French (50.4), suggesting once again that Standard Mandarin is a SylT language. However, when it comes to *n*PVI, we find a more complex situation. On the one hand, our figure 49.6 can still be arguably said to fall within the range of the SylT languages since it is closer to Grabe and Low's *n*PVI for French (43.5) than to that for English (57.2). On the other hand, however, our figure is considerably different from Grabe and Low's—a mere 27.0—for Singapore Mandarin. We will discuss the possible reasons for such discrepancy in the next section.

Another immediate observation we can make from Table 1 concerns the difference (or similarity) between passage and conversation figures. Just by visual inspection, it is apparent that style produced no significant difference in vowel proportion (at 56.23 for passages and 56.07 for conversations). However, all the other three variance measures, the ΔC and the two PVIs, are found to be consistently larger for conversations than for passages. When we put all the pairs of values from passages and conversations under statistical analysis of paired-sample T-Test, the results revealed a significant difference between one set of values—the consonant standard deviation (ΔC) values ($df = 5$, $p < 0.05$), and no significant difference between any other set of values. More on the difference caused by style is found in the next section.

5 Discussion

5.1 Vowel Percentage (%V)

As we have presented in the result section, our %V is not only higher than the typical StrT languages such as English, but also higher than the so-called mora- timed language, Japanese. One of the possible reasons for this is that our data source was different from *Ramus et al.* Specifically, our conversation and passage reading, being more natural speech styles than Ramus *et al.*'s simple declarative sentences, may have generated more consonantal erosion at the coda position, thereby yielding larger vowel percentage. Coincidentally, our %V at 56.15 is quite comparable to Grabe and Low's %V for Singapore Mandarin at 55.8.[3] In that study, Singapore Mandarin has the largest %V value, larger than that for Japanese at 45.5. It appears that Grabe and Low's use of passage reading (as opposed to declarative sentence reading) may have resulted in their %V value being comparable to ours.

5.2 Vocalic nPVI

Another issue that deserves further discussion is our *n*PVI, which is much higher than Grabe and Low's value for Singapore Mandarin. In other words, our data vary much more in the vowel intervals when adjacent V pairs are compared. One plausible explanation may come from the fact that we studied Standard Mandarin spoken in China while Grabe and Low studied Singapore Mandarin. Many studies have made observations on the differences

---

[3] Grabe and Low also calculated %V in order to compare their study with Ramus *et al.*'s.

between these two Mandarins (e.g., Goh *et al*. 2006, Zhou 2006, Zhou and Zhou 2007, Zhao *et al*. 2007). It is generally observed that Singapore Mandarin has a distinctive accent from that of Standard Mandarin. Zhou (2006), for instance, found that Singapore Mandarin does not normally have the [r] coda, or the neutral tone. Zhou and Zhou (2007) found that Singapore Mandarin shares a number of important segmental features found in Southern Chinese dialects such as Min, Yue and Kejia (Hakka). For instance, it has a distinctive fifth tone, the *ru* tone, found in all major Southern Chinese dialects but not Northern Mandarin. As the Min dialect, Singapore Mandarin lacks the retroflexed alveopalatals <zhi, chi, shi>[4] found in Standard Mandarin (see Lin 2001 for a description of Standard Mandarin consonants) and in their place one normally finds <z, c, s> respectively. And like the Min dialect, Singaporean Mandarin does not have the umlaut high-front-rounded vowel [ü] and in its place one normally finds its unrounded counterpart [i]. What do all the tonal and segmental differences mean when it comes to the suprasegmental qualities such as rhythm? We unfortunately have no answer to this question other than speculating that these differences must somehow have had an impact on the rhythmic patterns. Further studies need to be done to shed light on it.

Another factor that may have caused the difference in our and Grabe and Low's *n*PVI values is that we had more subjects than Grabe and Low. That is, our *n*PVI result is averaged over six people and 12 figures from two tasks while Grabe and Low's comes from a single individual and one figure from doing one task (i.e., reading a passage). A detailed look at our figures in Table 2 shows a range from 39.75 to 57.78, which is a difference of 18.03. From our standard deviations of these mean values in Table 1, one can also see that the standard deviation value is much higher for PVI values ($\sigma = 5.38$ on $\mu = 49.61$ for *n*PVI; $\sigma = 6.36$ on $\mu = 54.07$ for *r*PVI) than for %V values ($\sigma = 2.19$ on $\mu = 56.15$), suggesting that PVI values are much less homogenous than %V values. The relatively wider range of and higher variance among the PVI scores suggest that larger samples are necessary to achieve valid PVI results.

To sum up, our results of vowel proportion and consonantal variation, as measured by consonant standard deviation and *r*PVI, have supported the auditory impression that Mandarin Chinese is a syllable-timed language similar to French on all the three measures. However, more research is needed to understand the *n*PVI measure better.

5.3 Comparison between passage and conversation

It was mentioned in the last section that on average, all the three variance measures, the ΔC

---

[4] These sounds are represented using the *pinyin* alphabets.

and the two PVIs, are found to be consistently larger for conversations than for passages. It thus appears that conversation generates more consonantal and vocalic durational variations than passage reading. Detailed examination of the results for the six individuals in Table 2 also shows that every subject's conversation ΔC value is consistently higher than his/her individual passage ΔC value. All clearly indicates that style does make a difference in rhythmic measure of variance, especially by ΔC values.

6 Conclusion

Our study followed the well-known studies of Ramus *et al*. (1999) and Grabe and Low (2002) and measured vowel percentage, consonant standard deviation, variation of the pairs of two adjacent vowel intervals (*n*PVI), and variation of the pairs of two adjacent consonant intervals (*r*PVI) of Mandarin Chinese. Except on the measure of *n*PVI, our results have confirmed the traditional auditory impression that Mandarin Chinese has a syllable-timed rhythm.

References

Asu, E. L. & F. Nolan 2005. Estonian rhythm and the Pairwise Variability Index. *Proceedings of Fonetik 2005,* Gothenburg, pp. 29-32.

Bahrick, L. E., and Pickens, J. N. 1988. Classification of bimodal English and Spanish language passages by infants. *Infant Behavior and Development*, 11, 277-296.

Bertoncini, J. & Mehler, J. 1981. Syllables as units in infant perception. *Infant Behavior and Development*, 4, 247-260.

Bolinger, D. 1965. *Forms of English: Accent, Morpheme, Order*, ed. by I. Abe & T. Kanekiyo. Cambridge, Mass: Harvard University Press, 279–315.

Borzone de Manrique, A. M. & Signorini, A. 1983. Segmental durations and the rhythm in Spanish. *Journal of Phonetics*, 11, 117-128.

Christopher, A., & Dupoux, E. 1996. Bootstrapping lexical acquisition: the role of prosodic structure. *The Linguistic Review,* 13, 383-412.

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

Dauer, R. M. 1987. Phonetic and phonological components of language rhythm, Proceedings of the XIth International Congress of Phonetic Sciences, Tallinn, Estonia, 447-450.

Deterding, D. H. 2001. Letter to the Editor: The measurement of rhythm: a comparison of Singapore and British English. *Journal of Phonetics*, 29, 217-230.

Goh, Hock Huan, Chunsheng Zhao & Yongbing Liu (2006). Singaporean Mandarin and Identity: A preliminary analysis on Mandarin of Chinese Preschoolers. The 5th International Conference on Chinese Sociolinguistics: Language and Identity, Research Institute of Linguistics and Applied Linguistics Conference, School of Foreign Languages of Peking University, Beijing, China, December 2006.

Grabe, E. & Low, E. L. 2002. Durational variability in speech and the rhythm class hypothesis. In Carlos Gussenhoven & Natasha Warner (eds.) *Laboratory Phonology 7*. New York: Moutonde Gruyter. 515-546.

Lee, Wai-Sum & Zee, Eric 2003. Standard Chinese (Beijing). *Journal of the International Phonetic Association* 33(1), 109-112.

Lin, Hua 2001. *A Grammar of Mandarin Chinese*. Munchen: Lincom Europa.

Lloyd-James, A. 1940. *Speech Signals in Telephony.* London**.**

Mehler, J., Jusczyk, E W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. 1988. A precursor of language acquisition in young infants. *Cognition*, 29, 143-178.

Nazzi, T., Bertoncini, J., and Mehler, J. 1998. Language Discrimination by Newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24(3), 756-766.

Pike, K. L. 1945. *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.

Ramus, F. 2000. Perception of linguistic rhythm by newborn infants. Manuscript.

Ramus, F., Dupoux, E. & Mehler, J. 2003. The psychological reality of rhythm classes: Perceptual studies. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 337-342.

Ramus, F., Nespor, M. & Mehler, J. 1999. Correlates of linguistic rhythm in the speech signal, *Cognition,* 72, 1-28.

Roach, P. 1982. On the distinction between `stress-timed' and `syllable-timed' languages. In D. Crystal, *Linguistic controversies*, London: Edward Arnold. 73-79.

Roach, P. 2004. British English: Received Pronunciation *Journal of the International Phonetic Association* 34(2), 239-245.

Wenk, B. & Wioland, F. 1982. Is French really syllable-timed? *Journal of Phonetics,* 10, 193-216.

Zhao, Chunsheng, Yongbing Liu & Hock Huan Goh 2007. Singaporean Chinese Children's Special Chinese Vocabulary:Their Identity. Proceedings of the Redesigning Pedagogy: Culture, Knowledge and Understanding. Singapore, May 2007.

Zhou, Qinghai 周清海 2006. The differences between Singapore Mandarin and the Standard Mandarin and how to deal with the differences. 新加坡华语和普通话的差异与处理差异的对策. Published in series on March 21-23 in Singapore United Morning Newspaper. 连载于2006年3月21日、23日新加坡《联合早报》。

Zhou, Qinghai & Changji Zhou 周清海 / 周长楫 1998. Singapore Southern Min dialect and the Standard Mandarin. 新加坡闽南话与华语。Singapore United Morning Newspaper 载于1998年10月20日新加坡《联合早报》。

Zhou, Qinghai & Guozheng Xiao 周清海/ 萧国政 1998. Choices of lexical forms, meanings and usage in Singaporean Mandarin. 新加坡华语词的词形选择、词义选择和词用选择。Presented at the International Conference on the Teaching of Putonghua (the Standard Mandarin). Macao. 1998年3月11－13日(澳门)普通话(国语)教学国际研讨会。