ON THE SYLLABLE-TIMING OF CANTONESE AND BEIJING MANDARIN

MOK Pik Ki Peggy

Abstract: This study investigates the speech rhythm of Cantonese and Beijing Mandarin using some recently developed acoustic rhythmic measures. The two languages were compared with four languages in the BonnTempo corpus: German and English (stress-timed) and French and Italian (syllable-timed). Six Cantonese and six Beijing Mandarin native speakers were recorded reading the North Wind and the Sun story with a normal speech rate and telling the story semi-spontaneously. Both raw and normalised rhythmic measures were calculated using vocalic, consonantal and syllabic duration (ΔC, ΔV, ΔS, %V, VarcoC, VarcoV, VarcoS, rPVI_C, rPVI_S, nPVI_V, nPVI_S). Results confirm the syllable-timing impression of Cantonese and Beijing Mandarin, and suggest that Cantonese may have the most typical syllable-timed rhythm among the languages in this study, probably due to its lack of lexical stress. This study also shows that, in addition to consonantal and vocalic duration, syllable duration can potentially be useful in distinguishing speech rhythm.

Keywords: speech rhythm, Cantonese, Beijing Mandarin, rhythmic metrics

1. INTRODUCTION

Speech researchers have traditionally classified languages into different rhythmic groups: syllable-timed, stress-timed and mora-timed [1, 19]. English and German are typical stress-timed languages; French and Italian are typical syllable-timed languages, and Japanese is a typical mora-timed language. This rhythm class hypothesis was based on the notion of isochrony, i.e. there are units of equal or near-equal duration in the speech signal for such classification: syllables for syllable-timed languages, inter-stress intervals (feet) for stress-timed languages and mora for mora-timed languages. However, many experimental studies could not find concrete evidence for such isochronous units in the speech signal to support the rhythmic class hypothesis (see [9, 14, 18] for reviews). For example, the syllable duration of syllable-timed languages is equally variable as stress-timed languages [18], while duration of inter-stress intervals in stress-timed languages is not more variable than in syllable-timed languages [9]. Beckman [4] and Laver [15] concluded the early attempts to find acoustic correlates of speech rhythm by suggesting that speech rhythm is merely perceptual, since no reliable evidence could be found for isochrony.

Nevertheless, despite the lack of isochronous units, Dauer [9] and Roach [18] pointed out that stress-timed languages and syllable-timed languages differ in several important phonological aspects: syllable structure, vowel reduction and stress. Stress-timed languages have more variation in syllable length and structure, more reduced unstressed syllables, more variation in the phonetic realisation of stress and more stress-related rules than syllable-timed languages. These features, rather than any isochronous unit, combine with one another to give the impression of stress-timing versus syllable-timing. In addition, contrary to the early assumption of categorical distinction of speech rhythm, they suggested that languages can be more or less stress-timed or syllable-timed, with a continuum between the two.

The above insights are captured by several

recently developed acoustic measures of speech rhythm which could reflect the auditory impression of different rhythmic classes: %V (percentage of vocalic durations in speech), ΔC, ΔV (standard deviations of consonantal and vocalic duration respectively) by Ramus *et al.* [17] and Pairwise Variability Index (PVI) of vocalic and consonantal durations by Grabe & Low [14]. These measures depart from the search of isochronous phonological units; instead, they consider the durational variability in speech. They take only the duration of vowels and consonants as the basis for rhythmic classifications. Due to the various phonological differences mentioned above, stress-timed languages would have higher variability of consonant and vowel duration than syllable-timed languages. Their results show that %V and ΔC, the normalised vocalic PVI and the raw consonantal PVI can categorise different languages into distinct rhythmic clusters, while languages having less typical or unknown rhythm may fall between these clusters.

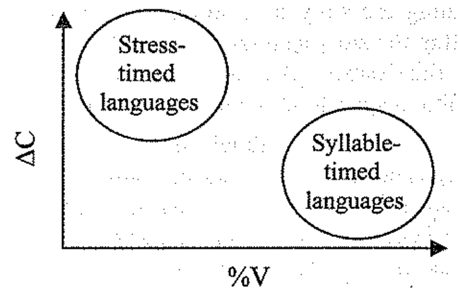

Figure 1: Patterns according to Ramus *et al.*

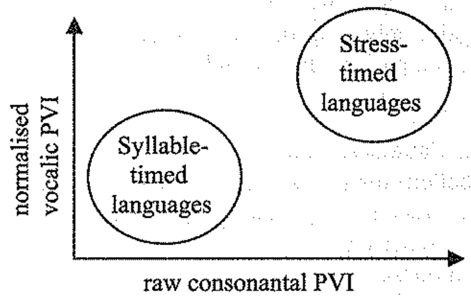

Figure 2: Patterns according to Grabe & Low.

Figures 1 and 2 show the hypothetical patterns of syllable-timed and stress-timed languages according to the measures proposed by Ramus *et al.* [17] and Grabe and Low [14]. Subsequent studies confirm that these acoustic measures can be used to distinguish languages with different speech rhythm, e.g. [20].

This study investigates the speech rhythm of Cantonese and Beijing Mandarin using the above acoustic measures. Cantonese has a very simple syllable structure with no lexical stress and no phonological vowel reduction. Every syllable carries a lexical tone. In emotionally neutral sentences, each syllable receives roughly equal emphasis [3]. Impressionistically, Cantonese is a typical syllable-timed language, but so far no experimental study has examined its rhythm using acoustic measures. Moreover, syllable-timed languages like Spanish and Italian have regular lexical stress while Cantonese does not. It is also unclear how Cantonese speech rhythm compares with other syllable-timed languages.

The speech rhythm of Beijing Mandarin is less clear. Mandarin is similar to Cantonese in that it also has lexical tones and a very simple syllable structure. Impressionistically, it also sounds quite syllable-timed. However, unstressed syllables (the so-called 'neutral tone') occur frequently in Mandarin. Duration of such toneless syllables is dramatically reduced and their vowel qualities are also reduced to schwa-like [6, 8]. The frequent occurrence of unstressed syllables is characteristic of stress-timing. Cao [7], by measuring syllables and feet duration, concluded that there is no evidence to regard Beijing Mandarin as syllable-timed because no isochronous units could be found. However, since the notion of isochrony was shown to be inadequate for speech rhythm classification, her conclusion should be considered tentative.

Other studies have used acoustic rhythm measures to investigate Mandarin speech rhythm. Grabe & Low [14] found that Mandarin has the lowest vocalic PVI values among all the languages in their study suggesting that it is a typical syllable-timed language. However, they looked at Singaporean Mandarin in which unstressed syllables occur much less frequently than in Beijing Mandarin because Singaporean Mandarin is heavily influenced by other southern Chinese languages. The two Mandarin accents sound quite different in various aspects.

It is likely that there may be subtle differences between the rhythmic patterns of these two Mandarin accents.

Benton *et al.* [5] compared Beijing Mandarin and American English using rhythmic measures with over 50 speakers in each language. They found that the rhythmic values for Mandarin and English are significantly different, but there was considerable diversity between individual speakers of both languages. Also, they did not state explicitly whether Mandarin is a syllable-timed language, although the text strongly implied that it is. In addition, given the continuum between stress- and syllable-timing, there can be statistically significant variation among languages belonging to the same rhythm class [10, 12]. Therefore, comparison with more languages using rhythmic measures is necessary in order to investigate the speech rhythm of Beijing Mandarin.

Besides investigating the speech rhythm of Cantonese and Beijing Mandarin using the above-mentioned acoustic rhythmic measures, this study also compares the two languages with four languages in the BonnTempo Corpus [12]: German and English (stress-timed), French and Italian (syllable-timed) for a clearer picture of Cantonese and Mandarin speech rhythm. In addition, in view of the salience of the syllable in speech production and perception, this study also calculates the rhythmic measures using syllable duration in order to test whether variability of syllable duration also correlates with different speech rhythm [2, 13].

2. METHOD

2.1 Speakers

Six native Hong Kong Cantonese speakers and six native Beijing Mandarin speakers (three male, three female) were recorded. They were either undergraduate or postgraduate students at the Chinese University of Hong Kong and were paid to participate in the experiment. None of them reported any speech or hearing problem.

These speakers were compared with previously published data (the BonnTempo corpus, see [12]). The number of languages and speakers in that corpus are as follows: German (15), British English (7), French (6) and Italian (3). German and English represent examples of stress-timed languages, French and Italian examples of syllable-timed languages.

2.2 Materials and procedures

The North Wind and the Sun stories in Cantonese [21] and Mandarin [16] were used as the experimental materials. The recording took place in a sound-treated room at The Chinese University of Hong Kong. Recordings were made directly to disk with a sampling rate of 22050 Hz. The speakers practised reading the story as many as times as they liked before the actual recording. They were recorded reading the story with three self-selected speech rates: normal, fast and slow. Then, they were recorded telling the story themselves without reading the script for semi-spontaneous speech. Finally, they were recorded doing free talk for about five minutes. Only data for reading the story in a normal speech rate and telling the story semi-spontaneously is reported in this paper. Analysis of the data with different speech rates and styles is underway.

The speech material in the BonnTempo Corpus consists of read speech based on a short passage from a novel in German, which was translated into the other languages by native speakers of the target language (English, French and Italian). Five speech rates were used: very slow, slow, normal, fast, very fast. Again, only data for normal speech rate is used for comparison in this study.

2.3 Speakers

All Cantonese and Mandarin sound files were labelled manually into syllabic, consonantal and vocalic intervals using Praat and were cross-checked by the author, a native Cantonese speaker who also speaks Mandarin. Syllable intervals were labelled as phonological syllables by reference to acoustic cues and careful listening, unless no acoustic cues of the syllable can be found as in the case of elision. Segmentation criteria followed those in [14] except that a 50 ms closure duration was added to all post-pausal initial stops for

consistency. The story was divided into several sentences. Any silent pause within a sentence was excluded from further analysis. Pre-pausal or utterance-final syllables were not excluded because they may be language-specific and may contribute to the perceived rhythmic pattern. The sound files in the BonnTempo Corpus were labelled in a similar way.

2.4 Calculation of rhythmic measures

Duration (ms) of syllabic, consonantal and vocalic intervals was extracted using a Praat script. Altogether eleven rhythmic measures were calculated for each sentence by each speaker, which were then averaged for each speaker. Details of the measures are as follows:

- ΔC: the standard deviation of consonantal duration
- ΔV: the standard deviation of vocalic duration
- ΔS: the standard deviation of syllabic duration
- %V: the proportion of vocalic duration within a sentence

Since ΔC and ΔV have repeatedly been demonstrated to interact with the average segment duration, a normalisation procedure was applied by calculating the coefficient of variation [10].

- VarcoC: (ΔC/mean consonantal duration)×100
- VarcoV: (ΔV/mean vocalic duration)×100
- VarcoS: (ΔS/mean syllabic duration)×100

In addition, two sets of PVI values, raw (1) and normalised (2), were calculated using the following two formulas from [14]. These indexes express the level of variability in successive intervals. Raw PVI, taking the absolute difference in duration between each pair of successive units, was calculated for consonantal (rPVI_C) and syllabic (rPVI_S) duration. Normalised PVI uses the mean duration of each pair of successive units to normalise for speech rate variations. Normalised PVI was calculated for vocalic (nPVI_V) and syllabic (nPVI_S) duration.

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

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

(where m = number of items; d = duration of the kth interval)

3. RESULTS

3.1 %V, ΔC, VarcoC, nPVI_V and rPVI_C

Following [17], [10] and [14], Figures 1 to 3 below show ΔC plotted against %V, VarcoC against %V and nPVI_V against rPVI_C respectively of the languages used in this study. In all the figures and tables below, Can = Cantonese, Man = Mandarin, _n = reading with a normal speech rate, _t = telling the story semi-spontaneously.

It can be seen in Figures 1 and 2 that ΔC and %V seem to give a clearer separation of stress-timed languages and syllable-timed languages than VarcoC and %V, although both sets of measures show distinct clusters of languages. In both cases, Cantonese and Mandarin pattern with syllable-timed Italian and French more than stress-timed German and English. The %V values of Cantonese and Mandarin are higher than Italian and French, suggesting that the two languages may show even stronger syllable-timing than Italian and French (comparing Italian and French with the normal version of Cantonese and Mandarin [$F(3,17) = 5.758$, $p = 0.007$]). Post hoc comparisons with Bonferroni adjustment shows that French is significantly different from both Cantonese ($p = 0.020$) and Mandarin ($p = 0.047$), while Cantonese and Mandarin are not significantly different. The %V values of semi-spontaneous speech of Cantonese and Mandarin (_t) are higher than read speech with a normal speech rate (_n). Paired-samples t-tests show that this stylistic difference is significant for both Cantonese [$t(5) = -3.591$, $p = 0.016$] and Mandarin [$t(5) = -3.754$, $p = 0.013$]. In addition, Mandarin has a higher VarcoC value than Cantonese for both versions, but independent-samples t-tests show that this difference is insignificant (_n: [$t(10) = -0.913$, $p = 0.383$]; _t: [$t(10) = -1.175$, $p = 0.267$]).

The nPVI_V and rPVI_C parameters differentiate syllable- and stress-timing less clearly than ΔC and %V. In Figure 3, the stress-timed German is indistinguishable from syllable-timed languages in the nPVI_V parameter. Cantonese and Mandarin again pattern with syllable-timed Italian and French. In addition, the stylistic difference between read speech and semi-spontaneously speech in Cantonese and Mandarin observed in %V disappear in both nPVI_V and rPVI_C. The two languages also do not have more extreme values than Italian and French.

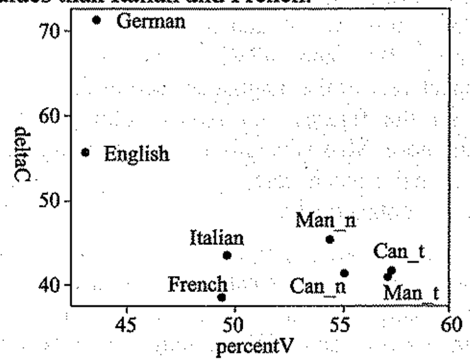

Figure 3: ΔC and %V of all the languages.

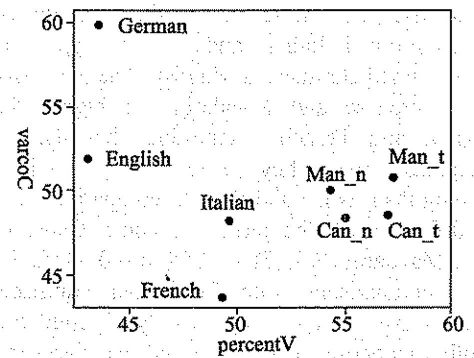

Figure 4: VarcoC and %V of all the languages.

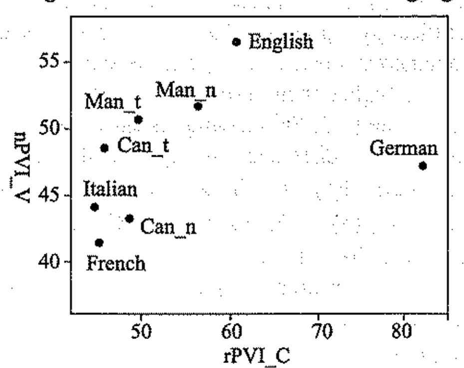

Figure 5: nPVI_V and rPVI_C of all the languages.

3.2 Indexes of syllable duration

In addition to calculating various indexes for consonantal and vocalic duration, this study also calculates such indexes for syllable duration. Tables 1 and 2 show the values of ΔS, VarcoS, rPVI_S and nPVI_S of all the languages in this study in descending order.

Table 1: ΔS and VarcoS of all the languages.

Language	ΔS		Language	VarcoS
English	88.74		English	51.87
German	80.78		Italian	46.73
Man_n	75.80		German	43.53
Man_t	68.33		Man_t	39.27
Italian	67.61		Man_n	38.17
Can_t	62.90		French	36.15
Can_n	57.48		Can_t	34.70
French	55.30		Can_n	30.71

Table 2: rPVI_S and nPVI_S of all the languages.

Language	rPVI_S		Language	nPVI_S
English	115.50		English	69.67
German	99.62		German	56.42
Man_n	86.08		Italian	54.78
Italian	82.68		French	49.47
Man_t	79.37		Man_t	45.95
French	75.89		Man_n	45.02
Can_t	65.97		Can_t	36.77
Can_n	63.62		Can_n	34.32

It is interesting to note that except VarcoS, the other three measures all rank stress-timed English and German at the top, followed by other syllable-timed languages. The rPVI_S parameter seems to give the best separation between stress-timed and syllable-timed languages, followed by ΔS. Although nPVI_S gives the same order, there is only a small difference between German and Italian suggesting that there may not be a clear-cut separation. Finally, all four measures rank Mandarin higher than Cantonese meaning that there is more variation of syllable duration in Mandarin than Cantonese, in line with expectation because of the frequent occurrence of unstressed syllables in Beijing Mandarin. Cantonese is ranked the lowest by three out of the four measures, indicating that Cantonese may sound even more syllable-timed than Italian and French, echoing the results of %V.

4. DISCUSSION

The main focus of this study is to investigate the speech rhythm of Cantonese and Beijing Mandarin. All rhythmic measures confirm the syllable-timing impression of the two languages. The results also suggest that Cantonese may have an even stronger syllable-timed rhythm than Mandarin, French and Italian, which presumably is contributed by the absence of lexical stress in Cantonese. A similar situation is also found in Singaporean Mandarin which has far fewer unstressed syllables than Beijing Mandarin. The data in [14] shows that Singaporean Mandarin has the lowest nPVI_V value and the highest %V value among all the languages in their study, suggesting that Singaporean Mandarin is the most typical syllable-timed language. It will be of interest to compare more syllable-timed languages with and without lexical stress to assess the effect of lexical stress on the perception of syllable-timing. Results from the present study and [14] suggest that the presence of lexical stress not only contributes to the distinction between stress-timing and syllable-timing, but can also affect the degree of syllable-timing.

Although both Cantonese and Beijing Mandarin have a syllable-timed rhythm, the rhythm of natural Mandarin speech sounds more variable than that of Cantonese impressionistically. The results of syllable durations in both read speech and semi-spontaneous speech confirm this impression of the two languages. Further analysis using naturally occurring speech materials in both languages with acoustic rhythmic measures is underway for a more thorough investigation of their speech rhythm. In addition, since Ramus *et al.* [17] showed that listeners can distinguish speech rhythm by listening to highly reduced synthesized speech (flat 'sasasa' speech), it will be interesting to investigate whether listeners can distinguish Cantonese and Mandarin speech rhythm perceptually using highly reduced synthesized speech based on naturally occurring speech data.

The significant difference in %V values of the two styles in Cantonese and Mandarin (read speech vs semi-spontaneous speech) implies that speakers may slightly change their rhythmic patterns according to speaking styles. This seems quite possible because read speech and spontaneous speech can differ in many aspects, including prosody. The stylistic difference in %V can also be partly explained by segmentation issues. Initial /j/ and /w/ were considered consonantal if there were acoustic cues for segmentation. However, in semi-spontaneous speech, many of these initial glides could not be separated from the following vowels so they could only be considered vocalic. This contributed to a higher percentage of vocalic portions in semi-spontaneous speech. On the other hand, Benton *et al.* [5] showed that in Mandarin, genre (news broadcast vs interview) indeed gave significantly different values for various rhythmic measures, which parallels the stylistic difference found in this study. So far, most studies on speech rhythm use only one speaking style, either read speech or spontaneous speech. Analysis of the data on free talk versus read speech by both Cantonese and Mandarin speakers is currently underway. More studies comparing speaking styles are needed in order to further explore the relationship between speech styles and rhythm.

The acoustic rhythmic measures were developed to capture the variability of consonantal and vocalic intervals in speech. These two kinds of intervals were used partly because of the failure of early attempts to find isochronous phonological units in the speech signal, and also partly because infants are able to distinguish languages based on their speech rhythm, see [17] for discussion. If infants who have no prior knowledge of a language's phonological structure can distinguish languages with different speech rhythm, they must be attending to some basic acoustic properties in the speech signal. Dellwo *et al.* [11] also showed that stress- and syllable-timed languages can be distinguished on the basis of voiced and voiceless intervals alone. Nevertheless, the present study shows that the variability of phonological syllable duration can also potentially distinguish stress- versus syllable-time languages, echoing [2, 13]. It is quite reasonable in view of the salient role played by the syllable in speech production and perception. Of course, further studies with more languages are needed in order to verify this claim, as well as to investigate which rhythmic measures work best with syllable duration.

5. CONCLUSION

This study confirms the syllable-timing impression of Cantonese and Beijing Mandarin with acoustic rhythmic measures. Results show that Cantonese may have an even stronger syllable-timed rhythm than Mandarin, French and Italian, probably due to its lack of lexical stress. The results also show that different speech styles exhibit variations in speech rhythm. In addition to consonantal and vocalic durations, this study also demonstrates that variations of syllable duration can potentially be useful in distinguishing languages with different rhythmic patterns.

6. ACKNOWLEDGEMENTS

The author would like to thank Volker Dellwo from UCL for helpful discussion and the generous sharing of his data. Thanks also go to Thomas Tsoi for segmenting the sound files.

7. REFERENCES

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

[2] Asu, E.L.; Nolan, F. 2006. Estonian and English rhythm: a two-dimensional quantification based on syllables and feet. *Speech Prosody 2006*, Dresden, 249-252.

[3] Bauer, R. S.; Benedict, P.K. 1997. *Modern Cantonese Phonology*. New York: Mouton de Gruyter.

[4] Beckman, M. E. 1992. Evidence for speech rhythms across languages. In *Speech Perception, Production and Linguistic Structure*, Y. Tohura, E. Vatikiotis-Bateson & Y. Sagisaka (eds.). Tokyo: IOS Press, 457-463.

[5] Benton, M.; Dockendorf, L.; Jin, W.; Liu, Y.; Edmondson, J. 2007. The continuum of speech rhythm: computational testing of speech rhythm of large corpora from natural Chinese and English speech. *The 16th ICPhS*. Saarbrücken, 1269-1272.

[6] Cao, J. F. 1986. An analysis of Mandarin neutral tone syllables [in Chinese]. *Applied Acoustics* [in Chinese], 4, 3-8.

[7] Cao, J. F. 2000. Rhythm of spoken Chinese — linguistic and paralinguistic evidences. *The 6th ICSLP*. Beijing, 2000, 357-360.

[8] Chao, Y. R. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

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

[10] Dellwo, V. 2006. Rhythm and Speech Rate: A Variation Coefficient for ΔC. In *Language and Language-Processing*, Karnowski, P.; Szigeti, I. (eds.). Frankfurt am Main: Peter Lang, 231-241.

[11] Dellwo, V.; Fourcin, A.; Abberton, E. 2007. Rhythmical classification of languages based on voice parameters. *The 16th International Congress of Phonetic Sciences (ICPhS)*, Saarbrücken, Germany, 1129-1132.

[12] Dellwo, V.; Aschenberner, B.; Dancovicova, J.; Wagner, P. 2004 The BonnTempo-Copus and Tools: A database for the combined study of speech rhythm and rate. *INTERSPEECH-2004 (ICSLP)*. Jeju Island, Korea, 777-780.

[13] Deterding, D. 2001. The measurement of rhythm: a comparison of Singapore and British English. *Journal of Phonetics*, 29, 217-230.

[14] Grabe, E. & Low, E. L. 2002. Durational variability in speech and the rhythm class hypothesis. In *Laboratory Phonology VII*, Gussenhoven C.; Warner, N. (eds.). Berlin: Mouton de Gruyter, 515-546.

[15] Laver, J. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.

[16] Lee, W.S.; Zee, E. 2003. Illustrations of the IPA: Standard Chinese (Beijing). *Journal of the International Phonetic Association* 33, 109-112.

[17] Ramus, F.; Nespor, M.; Mehler, J. 1999. Correlates of linguistic rhythm. *Cognition* 73, 265-292.

[18] Roach, P. 1983. On the distinction between stress-timed languages and syllable-timed languages. In *Linguistic Controversies: Essays in Honour of F.R. Palmer*, D. Crystal (ed.). London: Arnold.

[19] Warner, N.; Arai, T. 2001. Japanese mora-timing: a review. *Phonetica, 58*, 1-25

[20] White, L.; Mattys, S. L. 2007. Calibrating rhythm: first language and second language studies. *Journal of Phonetics* 35, 501-522.

[21] Zee, E. 1999. Illustrations of the IPA: Chinese (Hong Kong Cantonese). *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, 58-60.

MOK, Pik Ki Peggy, Department of Linguistics and Modern Languages, Fung King Hey Building, The Chinese University of Hong Kong, Shatin, Hong Kong.

普通话语流中的声调音高特征分析

熊子瑜

摘要：本文以篇章语音材料为研究对象，通过提取和分析音节调形段的音高特征数据，并在此基础上采用决策树工具对普通话语流中的声调类型进行了判别分析，以考察声调在语流中的音高特征。研究结果表明，即便在不考虑前后语音环境等因素的条件下，仅依据音节调形段的音高特征数据就可以在很大程度上有效识别普通话语流中的声调类型。基于数据分析的结果，本研究认为，尽管在语流中由于受到各类因素的影响，普通话音节的声调音高曲线可以有纷繁复杂的表现，声调调值也可以有很大幅度的变化，但除了轻声或轻音之外，绝大多数音节内部的音高对比关系（主要体现为声调调形）在语流中仍然具有较强的稳定性。

关键词：普通话 声调识别 音高特征

1. 引言

汉语是声调语言，其普通话有四个基本调类：阴平、阳平、上声和去声。从声调调形上看，阴平字属平调，阳平字属升调，上声字属先降后升的曲折调，有时会变读为阳平，有时会读为呈低降调的半上，去声字属降调。从调值上看，标准阴平字为 55 调，阳平字为 35 调，全上字为 214 调，半上字为 21 调，去声字为 51 调[11]。除了这四个基本调类，普通话中还有一些必读轻声的字词，由于轻声没有固定的调形和调值，所以通常不被看做一个独立的调类。另外，有些带调音节在语流中由于被弱化，也可能失去或改变其原有的声调，而读成“语流轻声”或“轻音”[6]。

单念时，普通话的字调调形和调值都比较稳定，主要受其调类控制。但在语流中，由于受到前后声调环境[1]、前后韵律边界类型[10]、语句焦点位置[2]、语气类型[8]、说话人的情绪态度，甚至说话人的生理状况等各类因素的影响，声调调值在语流中可以产生很大变化，有时甚至其调形也可能会发生一些显著的变化。

例如，图 1 显示了一个语句的音高曲线，其纵坐标单位为半音，对比其中阴平字“单、加、说”的调值，阳平字“极”和“明”的调值，上声字“简”和“(加）以”的调值，去声字“用、事、例”的调值，不难看出在诸多因素的影响下，声调调值在语流中已发生了显著变化，同一类声调的调值在不同环境下可以存在显著差异。

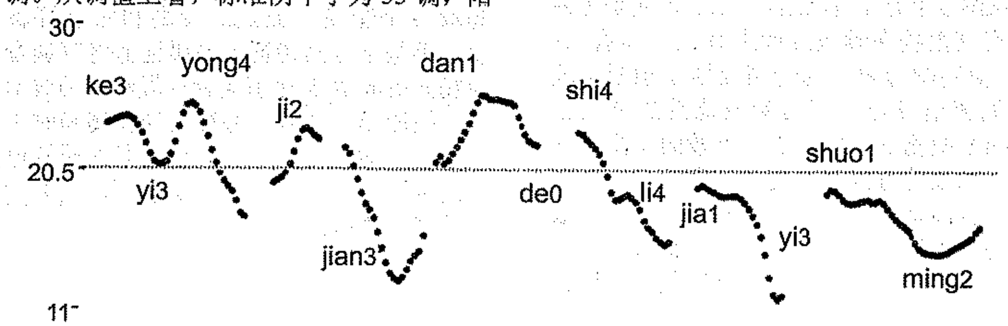

图 1：“可以用极简单的事例加以说明”的语句音高曲线