AN ACOUSTIC STUDY OF MANDARIN RHYTHM IN TAIWAN

Yi-hsiu Lai

National University of Kaohsiung, Taiwan
yhlai@nuk.edu.tw

ABSTRACT

This paper is to examine Mandarin rhythm spoken in Taiwan. Oral recordings of twenty speakers who took the National Exam of Qualified Mandarin Teachers held in Taiwan in 2010. Based on the oral scores, they were evenly divided into two different groups: high oral proficiency (HOP) and low oral proficiency (LOP). Rhythmic correlates in sentence reading were analyzed from two perspectives: (a) pitch contour: the highest F0, the lowest F0, the distance between the highest F0 and the lowest one within a prosodic word (PW), (b) timing pattern: syllable duration, vowel duration, consonant duration and vowel percentage (V%). Results indicated that speakers of HOP significantly outperformed those of LOP in most of the rhythmic correlates. Finally, suggestions for further rhythmic studies and pedagogical implications for Mandarin Chinese learners are offered.

Keywords: rhythm, pitch contour, timing pattern, Mandarin Chinese

1. INTRODUCTION

Rhythm is often perceptually defined as "the perceived regularity of prominent units in speech" [7] p.417. Rhythmic correlates include syllable weight (stressed vs. unstressed syllable), syllable length, and pitch range. In production, pitch changes are acoustically shown in the variation of fundamental frequency (F0). Bradlow, Torretta, & Pisoni [1] found that speakers of clear rhythmic patterns uttered with a wider range of F0 and a good control of segmental duration.

In the past decade, studies on Mandarin rhythm focus primarily on Pu-Tong-Hua ("spoken Mandarin") in Mainland China. Cao [2, 3, 4, 5], for example, made a series of reports on rhythmic patterns of Pu-Tong-Hua, analyzing syllable duration, rhythmic grouping, temporal structure of rhythmic unit, and distribution of pauses. Speech tempo adjustment in Pu-Tong-Hua was demonstrated in that the first-syllable duration is longer than the average duration, and that the final-syllable duration was shorter than the average one. Lin & Wang [12] studied the rhythmic patterns spoken by six Mandarin speakers in northern China. Four acoustic correlates of rhythm were examined: vowel percentage, consonant standard deviations, rPVI (raw Pairwise Variability Index) and nPVI (normalized Pairwise Variability Index). Speech samples were taken from conversations produced by six speakers of Beijing Mandarin. Findings in most of acoustic correlates, except for nPVI, proved that Mandarin was a syllable-timed language. Recently, Mok & Dellwo [13] compared native and non-native speech rhythm in Cantonese, Beijing Mandarin and English. Both raw and normalized acoustic measures, including vocalic, consonantal and syllabic durations were calculated. Results confirmed the syllable-timing impression of Cantonese and Beijing Mandarin. VarcoC and %V were the best classification of speech rhythm.

Mandarin in Taiwan, according to Yeh [16] is specified with certain supra-segmental properties, rhythm and intonation in particular. Yeh [16] observes that Mandarin in Taiwan is spoken with less prominence and with narrower ranges of pitch changes. This echoes the idea of "neural intonation" [6], in which utterances are expressed with flatter pitch contours and less emotion. "Expressive intonation", by contrast, is produced with clearer temporal patterns and wider pitch variations [6]. Different emotions can be conveyed with different expressive intonations, as frequently found in dramas. In Taiwan, few expressive intonations occur in daily conversations.

Starting from 1996, Ministry of Education in Taiwan annually holds one National Exam of Qualified Mandarin Teachers, which aims to promote the policy of teaching Mandarin for foreign learners and to enhance the oral proficiency as well as professional knowledge of Mandarin teachers. In the exam, there are two parts: one written test and one oral test. In the oral test, test-takers will be asked to read a list of words, sentences, a paragraph, and finally make spontaneous speech on an assigned topic. Huang & Yi [9] studied the oral test, indicating that it was

difficult for test-takers with daily neutral rhythm to pass the oral examination. Those who scored high in the oral test always possessed clear pronunciation and rhythmic patterns, which serve as a good model for foreign learners to imitate and follow.

Despite the growing research interest in Mandarin rhythm, there are still open questions. Most of the current reports address the temporal patterns of Mandarin in Mainland China, the so-called Pu-Tong-Hua [2, 3, 4, 5, 12]. Little attention is directed to the pitch ranges, being one of vertical acoustic cues, in Mandarin rhythm. Also, few studies have discussed the rhythmic patterns of Mandarin in Taiwan. In what way and to what degree Mandarin rhythm in Taiwan differs from that in Mainland China are worthy of investigation. Though Torgerson [14] acoustically analyzed the pitch difference between Beijing Mandarin and Taiwan Mandarin, the main focus was on tone register and no discussion was made from the perspectives of timing patterns. Moreover, no acoustic studies have been conducted on the test-takers' oral performance in the annual National Exam of Qualified Mandarin Teachers. It remains unknown in what way and to what degree the testees who passed the oral test acoustically differentiated from those who failed. To concretely and objectively describe the rhythmic patterns of those who passed the oral test will also be of great help in suggesting advice to future test-takers and to learners of Mandarin Chinese. Hence, the objectives of the present study aims (a) to examine possible vertical rhythmic patterns of Mandarin in Taiwan, as revealed by test-takers of National Exam of Qualified Mandarin Teachers; (b) to investigate possible horizontal rhythmic patterns of Mandarin in Taiwan, as revealed by test-takers of National Exam of Qualified Mandarin Teachers.

2. METHOD

An acoustic experiment on Mandarin rhythm in Taiwan is carried out. Description of subjects, data collection, and data analysis is presented in this section.

2.1. Participants

Participants in the current study include twenty female native speakers of Mandarin Chinese. They grew up in Taiwan and were college graduates who planned to be teachers of Mandarin Chinese for foreign students in Taiwan and abroad. In 2010, they took the National Exam of Qualified Mandarin Teachers held in Taiwan. In the exam, there were two parts: one written test and one oral test. In the oral test, these participants were asked to read a list of words, sentences, a paragraph, and finally they made spontaneous speech on an assigned topic. The whole oral session were recorded in a quiet room.

Based on scores in the oral tests, these twenty participants were evenly divided into two different groups: ten in the group of "high oral proficiency" (HOP) and ten in the group of "low oral proficiency" (LOP). The former group passed the oral rest, while the latter one did not. Also, speakers in HOP scored significantly higher than those in LOP. Oral recordings of these participants were analyzed from the perspectives of Mandarin rhythmic patterns.

2.2. Stimulus material & analysis

The current study investigates on the Mandarin rhythm produced by twenty female test-takers of National Exam of Qualified Mandarin Teachers 2010. One declarative sentence of their recorded readings in a read speech "氣象局預測明天的氣溫將下降三度" "The Weather Bureau predicts that temperature tomorrow will drop by three Celsius degrees" was transformed into the wave files in Praat (Boersma & Weenink, 1999-2000). Segmentation of each rhythmic unit in this sentence was first conducted, inclusive of segment (vowel, consonant), syllable, prosodic word (PW), prosodic phrase (PPh), and prosodic phrase group (PG) [15]. The example of rhythmic grouping is shown below: (|: boundary of PPh; %: boundary of PG)

PG

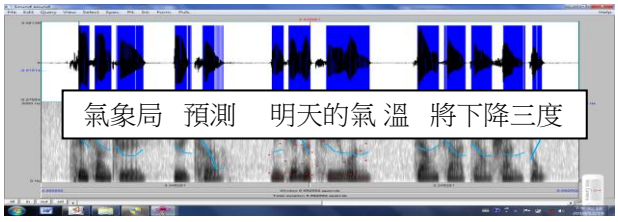
PPh | PPh | PPh

(PW) (PW) | (PW) (PW) | (PW) (PW)

%(氣象局)(預測)|(明天的)(氣溫)|(將下降)(三度)%

"The Weather Bureau predicts that temperature tomorrow will drop by three Celsius degrees"

Figure 1: Spectrogram of one declarative sentence.



Rhythmic correlates in the present analysis are examined from two perspectives: vertical acoustic cues and horizontal acoustic cues. Regarding the vertical acoustic cues, pitch changes are acoustically shown in the variation of fundamental frequency (F0), usually characterized by visible blue pitch lines in spectrograms (Figure 1). In each prosodic word (PW), the highest F0, the lowest F0, the distance between the highest F0 and the lowest one are respectively measured. As for horizontal acoustic cues, the temporal information is easily fetched in the spectrograms by defining the vowel region and the consonant region. The vowel region is usually characterized by visible periodic waves and vertical pulses in spectrograms [8, 10, 11]. Syllable duration, vowel duration, consonant duration and vowel percentage (V%) in one syllable are calculated.

Statistical analysis is conducted following the acoustic measurements. ANOVA is performed to examine group differences in their oral performance. Rhythmic patterns of different groups could be concretely and objectively described and compared from different perspectives: (a) vertical acoustic cues (i.e. pitch contour): the highest F0, the lowest F0, the distance between the highest F0 and the lowest one within a prosodic word (PW), (b) horizontal acoustic cues (i.e. timing pattern): syllable duration, vowel duration, consonant duration and vowel percentage (V%).

3. RESULTS & DISCUSSION

Several acoustic factors from both vertical and horizontal perspectives are compared to demonstrate the way how the HOP speakers of Mandarin Chinese in Taiwan differentiate from the LOP ones in rhythmic patterns. Statistical results are presented in terms of vertical acoustic cues and horizontal acoustic cues.

3.1. Vertical acoustic cues

Vertical acoustic features, including the highest F0, the lowest F0, the distance between the highest F0 and the lowest one within a prosodic word (PW), between HOP and LOP are examined by two-way repeated measures ANOVA (i.e. 2 groups x 3 measures). Statistical results are indicated in Table 1.

Table 1 reveals significantly different pitch patterns for two groups: HOP and LOP. First of all, HOP significantly differentiated from LOP in all two vertical features: the highest F0, the distance between the highest F0 and the lowest one within a prosodic word (PW). Concerning the highest F0 in one PW, speakers of HOP (M=354.77) were significantly higher than those of LOP (M=308.13) [$F(1, 18)=7.72$, $p < 0.01$]. Also, significant larger pitch distance in one PW was found in HOP (M=176.20) than in LOP (M=135.67) [$F(1, 18)=7.61$, $p < 0.01$]. No significant distinction, however, was identified in the lowest F0 between these two groups.

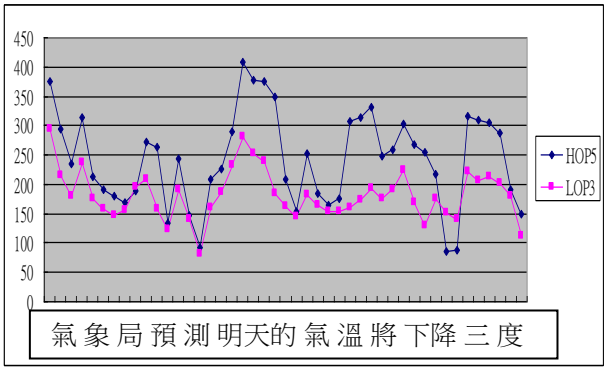
Table 1: Pitch patterns in PW produced by LOP and HOP.

Items	Group	N	M	SD	F
Highest F0	LOP	10	308.13	67.52	7.72**
	HOP	10	354.77	62.35	
Lowest F0	LOP	10	172.47	49.95	0.273
	HOP	10	178.57	39.92	
Pitch Distance	LOP	10	135.67	57.97	7.61**
	HOP	10	176.20	55.79	

Note: N: number; M: mean; SD: standard deviations; **p< 0.01

Findings in vertical acoustic cues implied that speakers of HOP produced significantly wider pitch contour changes within one prosodic word, which made their utterances more fluctuating. As demonstrated in Figure 2, one speaker in HOP (HOP5) read the statement "The Weather Bureau predicts that temperature tomorrow will drop by three Celsius degrees" with a dramatically-changing pitch contour. By contrast, one speaker in LOP (LOP3) uttered the sentence with a flat pitch pattern.

Figure 2: Pitch contour produced by HOP5 and LOP3.



3.2. Horizontal acoustic cues

Horizontal acoustic cues, inclusive of syllable duration, vowel duration, consonant duration and vowel percentage (V%), are investigated by two-way repeated measures ANOVA (i.e. 2 groups x 4

measures). Table 2 summarizes statistical results of temporal patterns.

Table 2: Temporal patterns produced by LOP and HOP

Item	Group	N	M	SD	F
Syllable duration	LOP	10	204.98	71.57	26.52***
	HOP	10	264.53	70.05	
Vowel duration	LOP	10	155.20	60.35	15.21***
	HOP	10	194.93	64.38	
Consonant duration	LOP	10	49.78	41.20	7.17**
	HOP	10	69.59	49.08	
Vowel Percentage (%)	LOP	10	76.65	15.81	0.77
	HOP	10	74.36	16.04	

Note: N: number; M: mean (syllable, vowel, consonant duration in msec.); SD: standard deviations; ***p< 0.00; **p< 0.01

Several findings can be made in Table 2. To begin with, these two groups (HOP vs. LOP) expressed significantly different temporal patterns, especially in syllable duration, vowel duration and consonant duration. Speakers of HOP uttered with longer syllable duration (M=264.53) than those of LOP (M=204.98) [F(1, 18)=26.52, p < 0.00]. Likewise, longer vowel duration was identified in HOP (M=194.93) than in LOP (M=155.20) [F(1, 18)=15.21, p < 0.00]. A similar pattern occurs in the consonant duration, in which HOP (M=69.59) produced significantly longer than LOP did (M=49.78) [F(1, 18)=7.17, p < 0.01]. No significant differentiation, however, was observed in the vowel percentage between these two groups.

4. CONCLUSIONS

The present study aims to examine possible vertical rhythmic patterns and horizontal rhythmic patterns of Mandarin in Taiwan, as revealed by test-takers of National Exam of Qualified Mandarin Teachers. Results indicated that speakers of high oral proficiency (HOP) significantly outperformed those of low oral proficiency (LOP) in most of the vertical and horizontal rhythmic correlates (i.e. the highest F0, the distance between the highest F0 and the lowest one within a prosodic word (PW), syllable duration, vowel duration, consonant duration). The current findings offer concrete and objective description of the rhythmic patterns, produced by speakers of better oral proficiency in Mandarin rhythm. These findings echo that in Bradlow, Torretta, & Pisoni [1], in which speakers of clear rhythmic patterns uttered with a wider range of F0 and a good control of segmental duration. For future test-takers of National Exam of Qualified Mandarin Teachers, it will be of great help for them to pass the test if they try to enhance these rhythmic features. As for learners of Mandarin Chinese, it will also be helpful if they work more on these rhythmic correlates to make their oral expressions with better Mandarin rhythm.

5. REFERENCES

[1] Bradlow, A.R., Torretta, G.A., Pisoni, D.B. 1996. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication* 20, 255-272.

[2] Cao, J.F. 1991. Durational patterns of syllables in Standard Chinese. *Proc. 12th ICPhS* Aix-en-Province, France.

[3] Cao, J.F. 1994. The effects of accentual focus and lexical stress upon temporal distribution in a sentence. *The Proceeding of ICSLP 4*, Yokohama, Japan.

[4] Cao, J.F. 1995. Basic temporal structure of a sentence in Standard Chinese. *Journal of Chinese Linguistics* 7.

[5] Cao, J.F. 1999. Acoustic-phonetic characteristics on the rhythm of Standard Chinese. *Proceeding of 4th National Conference on Modern Phonetics*, Beijing.

[6] Chao, Y.Z. 2002. *Linguistic Papers of Chao Y.Z.* Beijing: Shang-wu.

[7] Crystal, D. 2008. *A Dictionary of Linguistics and Phonetics* (6th ed.). UK: Blackwell

[8] Fry, D.B. 2001. *The Physics of Speech.* Cambridge: Cambridge University Press.

[9] Huang, S.I., Yi, Y.P. 2008. *Mandarin Oral Expression.* Taipei: New Sue-lin.

[10] Ladefoged, P. 1993. *A Course in Phonetics* (3rd ed.). N.Y.: Harcourt Brace Jovanvich Inc.

[11] Ladefoged, P. 2001. *Vowels and Consonants.* Mass: Blackwell.

[12] Lin, H., Wang, Q. 2005. Mandarin rhythm: an acoustic study. *Journal of Chinese Language and Computing* 17, 127-140.

[13] Mok, P.K., Dellwo, V. 2008. Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English. *www.cuhk.hk/lin/people/peggey/Mok&Dellwo_SP2008.pdf*

[14] Torgerson, R.C. 2005. *A Comparison of Beijing and Taiwan Mandarin Tone Register: An Acoustic Analysis of Three Native Speech Styles.* Unpublished MA thesis, Brigham Young University.

[15] Tseng, C.Y., Pin, S.H., Lin, Y.L., Wang, H.M., Chen, Y.C. 2005. Fluent speech prosody. *Speech Communication 46*, 284-309.

[16] Yeh, D.M. 2007. Prosody of Mandarin in Taiwan. *The third International Conference on Teaching Mandarin*.