INTERSPEECH 2017
August 20–24, 2017, Stockholm, Sweden

A Preliminary Phonetic Investigation of Alphabetic Words in Mandarin Chinese

Hongwei Ding[1], *Yuanyuan Zhang*[1], *Hongchao Liu*[2], *Chu-Ren Huang*[2]

[1]Institute of Cross-Linguistic Processing and Cognition, School of Foreign Languages, Shanghai Jiao Tong University, China
[2]Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong
{hwding; yyuan.zhang}@sjtu.edu.cn, cliu.hongchao@connect.polyu.hk, churen.huang@polyu.edu.hk

Abstract

Chinese words written partly or fully in roman letters have gained popularity in Mandarin Chinese in the last few decades and an appendix of such Mandarin Alphabetical Words (MAWs) is included in the authoritative dictionary of Standard Mandarin. However, no transcription of MAWs has been provided because it is not clear whether we should keep the original English pronunciation or transcribe MAWs with Mandarin Pinyin system. This study aims to investigate the phonetic adaptation of several most frequent MAWs extracted from the corpus. We recruited eight students from Shanghai, 18 students from Shandong Province, and one student from the USA. All the subjects were asked to read both 24 Chinese sentences embedding the MAWs and all 26 letters of the English alphabet. The results showed that Letters *A O N T* were predominantly pronounced in Tone 1; *H* was often produced with vowel epenthesis after the final consonant; and *B* was usually produced in Tone 2 by Shanghai speakers and in Tone 4 by Shandong speakers. We conclude that the phonetic adaptation of MAWs is influenced by the dialects of the speakers, tones of other Chinese characters in the MAWs, as well as individual preferences.
Index Terms: alphabetic words, Mandarin Chinese, phonetic adaptation

1. Introduction

The adaptation of loan words provides important tests of the phonetic and phonological systems of a language, and it is well known that orthography plays a role in phonological adaptation [1]. However, what happens when orthography is also adapted has never been studied before. *Mandarin Alphabetic Words (MAWs)* [2] is exactly such a case where individual English letters are borrowed and incorporated into Chinese character orthography to form new words. These are not *lettered words* since by definition each word in Chinese can be turned into a lettered word when represented by Pinyin Romanisation. For instance 爱克斯射线 is the calque loan word of *X-ray*, and can be represented in Pinyin as a lettered word *àikèsīshèxiàn*. However, the conventionalized and preferred form in Chinese nowadays is *X*射线. We know for sure that *X* is not pronounced as *aikèsī*, but its exact phonological or phonetic value has never been established before. *MAWs* have gained popularity in the last few decades and now form an established category in Chinese. An appendix of 239 MAWs is included in the authoritative *Dictionary of Modern Chinese (* 现代汉语词典 *)* [3]. And even several MAWs dictionaries have been published in recent years [4]. However, they only provide glosses but not pronunciations for MAWs, which awaits standardized transcription.

There are different views on how Mandarin alphabetical words should be transcribed. Some linguists insist that MAWs should be transcribed with the Pinyin system so that these words could be included in the Mandarin Chinese lexical system [5]. But some letters such as *F H L M S W X* have to be transcribed into more than one syllable; moreover, the Chinese phonetic transcriptions may deviate greatly from the pronunciation Mandarin speakers really use. Other scholars claim that the English letters in MAWs should be read in the same way as they are pronounced in English ([6], [7]). But the English pronunciations of these MAWs do not conform to the Mandarin phonological system. Another problem is that Chinese speakers can hardly pronounce English letters as the natives do. Mandarin Chinese speakers usually pronounce English words with a certain Chinese accent [8]. When English letters are combined with Chinese characters, speakers even try to adjust the pronunciation of the letters to fit the Chinese phonological systems. For example, speakers may attach a Chinese lexical tone to the syllable when they read the English letters. It seems that a logical way to address the transcription issue of MAWs is to investigate the performance of Mandarin speakers, and to illustrate the pronunciation of people who use MAWs descriptively.

Up to now, the studies of MAWs have been mainly focused on morphology; pronunciation research has been very sparse, and examination of the tones of MAWs has received little attention. It was remarked by Riha that "the tones for roman letter names have yet to be standardized in Mandarin" [9, p. 45]. The advancement of speech technology also requires a standardization of the pronunciation of MAWs, for a standard pronunciation should be provided for the speech synthesizer. An original English pronunciation of the letters in MAWs might sound non-Chinese, while a prescribed and deviated pronunciation with Mandarin Chinese Pinyin transcription might also be absurd. An intermediate approach which is used by most Chinese speakers might be accepted by listeners. A pilot investigation revealed that the pronunciation of the alphabetic words is dependent on numerous factors, such as speaker's level of education, English fluency, dialects, age, profession, personal interests, etc.

As a preliminary investigation in the phonetic adaptation of MAWs, we recruited the university students who belong to the population that use MAWs in their speech more frequently than other groups (e.g people who cannot speak English). We chose those MAWs which consist of one alphabetic letter followed by one Chinese character for the current investigation because the adaption of this kind of MAWs to Chinese phonological system might be more probable than other types of MAWs (e.g. those MAWs which consist of only English letters).

http://dx.doi.org/10.21437/Interspeech.2017-876

This study aims to answer these questions on the pronunciation (including tones) of English letters in MAWs:

1. Are there any regularities for the pronunciation?
2. Are the pronunciations influenced by speakers from different dialectal areas, by the way they produce English letters, by the combination of different tones of Chinese characters, or by the position of MAWs in the sentence?

2. Methods

Our method is to extract the basic empirical facts about MAWs from corpora, by which we will have a better idea of their distributional range and behaviours, and then can use the generalisations to design follow-up experiments.

2.1. Extraction from corpus

A study was carried out to extract nearly 60,000 types of MAWs from the Chinese Gigaword Corpus. It is a 14 billion word PoS (Part of Speech) tagged corpus containing data from China, Singapore, and Taiwan [10]. Based on this experience, we have established a taxonomy of distributional patterns of alphabetical letters in MAWs based on whether they co-occur with Chinese characters and on the number of letters and their positions. All the MAWs found are combinations of one or more Chinese characters and one or more English letters, which can be further divided into the following four categories [2]:

- LC (e.g. B股, T恤, H股, Ｂ型, B组, X光)
- CL (e.g. 甲A, 小S, 甲B, 大S, 卡拉ＯＫ)
- CLC (e.g. 朱F基, 程I青, 刘X福, 简X燕)
- the others (e.g. 奥迪A6, 阿Sa, 4A级, 5C班)

(Note: "C" and "L" stand for Chinese Characters (C) and English Letters (L) respectively.) Among the four categories, MAWs with LC combination rank the first in occurrence.

2.2. Reading materials

The most frequent LC MAWs were selected for the current investigation. The selected MAWs also have the same nominal PoS. This prevents the compounding effect of PoS and allows us to study the PoS effect in the future. The selected LC MAWs for the current investigation include six letters, which are grouped into three pairs: *A* and *B*, *H* and *N*, *T* and *O*. Among the most frequent MAWs in LC combinations, both *A* and *B* are found to be followed by four different lexical tones, *H* and *N* are mainly followed by Tone 3 (gǔ股), and *T* is mostly followed by Tone 4 (xù恤). In order to test the contrast with *T*, *O*恤 was constructed, which was not extracted from the corpus. Because *O* is usually pronounced at the initial and *T* at the final position of *OPQ RST* when the alphabet is read, we assumed that a tonal contrast might be formed. The investigated MAWs are listed in Table 1.

Table 1: *MAWs investigated*

	A	B	H	N	O	T
Tone 1	A区(qū)	B区				
Tone 1	A型(xíng)	B型				
Tone 3	A股(gǔ)	B股	H股	N股		
Tone 4	A队(dùi)	B队			O恤(xù)	T恤

These 12 MAWs were embedded in sentences both at the initial and final positions, which resulted in 24 sentences. The length of the sentences ranges from 9 to 18 characters. MAWs of the same pair were put in the same carrier sentences. These 24 sentences were presented to the subjects in a random order, so that no rhythm could be established in the elicitation process. Therefore, the reading materials consisted of two parts: 1) 24 sentences; 2) 26 letters in the English alphabet from a to z.

2.3. Subjects and recording

We recruited undergraduate students from Shanghai Jiao Tong University between 18-21 years old, including 16 students (8 females, 8 males) from Shandong Province and 8 students (4 females, 4 males) from Shanghai, which represent dialects of northern and central zones in China respectively ([11], [12]). One American male student who can speak Mandarin Chinese was also recruited as a reference. And these subjects were of average English level, who had passed the lower level of the national English test in the People's Republic of China–*College English Test (CET4)*, but not the higher level (*CET6*).

Every speaker was recorded individually in the professional recording studio at the School of Foreign Languages of Shanghai Jiao Tong University. The speech materials were presented to the speakers through individual PowerPoint slides. Recordings were made with a headset microphone and Mbox mini at a 44,100 Hz sampling rate with a 16-bit resolution.

3. Results

The results were presented in perceptual evaluation and acoustic analysis.

3.1. Perceptual evaluation

The tones and phonemes were assessed by two native Chinese speakers, and they agreed on all the evaluation results.

3.1.1. Phonemes

Letters *A B O T* were pronounced as /eɪ/, /biː/, /əʊ/, /tiː/ respectively in both the Chinese sentences and the English alphabet by all the US, Shanghai and Shandong speakers.

Letter *N* was pronounced by all speakers as /ən/ (嗯) in the Chinese text, and by two Shanghai speakers and 12 Shandong speakers as /ən/ in the English alphabet, the ratio of which is shown in Table 2.

Table 2: *Overview of the pronunciation of Letter N*

Context	Chinese sentences		English alphabet	
IPA	/ɛn/	/ən/	/ɛn/	/ən/
US	0	**100%**	**100%**	0
Shanghai	0	**100%**	**75%**	25%
Shandong	0	**100%**	25%	**75%**

Letter *H* was pronounced in one syllable (/eɪtʃ/) or two syllables (/eɪtʃɤ/). The percentages of the pronunciations by the speakers are listed in Table 3.

3.1.2. Tones

The classification of tones of Letter *H* with consonant finals or two syllables was not straightforward; therefore, *H* was not included in the statistics of tones. The identification of the mono-

Table 3: *Overview of the pronunciation of Letter H*

Context	Chinese sentences		English alphabet	
IPA	/eitʃ/	/eitʃɤ/	/eitʃ/	/eitʃɤ/
US	**100%**	0	**100%**	0
Shanghai	43.75%	**56.25%**	**100%**	0
Shandong	28.1%	**71.9%**	**87.5%**	12.5%

syllables of other letters was clear. The evaluation results of Shanghai and Shandong speakers are presented in the percentage of each lexical tone in Table 4 and Table 5 respectively.

Table 4: *Tones of students from Shanghai*

Letters	A	B	N	O	T
Context	Chinese sentences				
Tone 1	**100%**	4.7%	**100%**	**100%**	**100%**
Tone 2	0	**57.8%**	0	0	0
Tone 3	0	0	0	0	0
Tone 4	0	37.5%	0	0	0
Context	English alphabet				
Tone 1	**87.5%**	12.5%	**75%**	12.5%	25%
Tone 2	12.5%	**62.5%**	6.2%	**75%**	0
Tone 3	0	0	0	0	0
Tone 4	0	25%	18.8%	12.5%	**75%**

Table 5: *Tones of students from Shandong*

Letters	A	B	N	O	T
Context	Chinese sentences				
Tone 1	**100%**	0.8%	**100%**	**100%**	**84.4%**
Tone 2	0	0	0	0	0
Tone 3	0	0	0	0	0
Tone 4	0	**99.2%**	0	0	15.6%
Context	English alphabet				
Tone 1	**93.75%**	0	**75%**	**87.5%**	0
Tone 2	6.25%	12.5%	6.25%	6.25%	0
Tone 3	0	0	0	0	0
Tone 4	0	**87.5%**	18.75%	6.25%	**100%**

The tones of the US speaker in Chinese texts were usually level tones, but not so high as Tone 1. They could hardly be classified as any tone category in Mandarin Chinese.

3.2. Acoustic analysis

The acoustic analysis was conducted in Praat [13].

3.2.1. Phonemes

As we have mentioned that many speakers pronounced *N* as /ɛn/ in English but /ən/ in Chinese, which are illustrated in Figure 1 on the left and right respectively.

The left figure shows an English /ɛn/ segmented from the English alphabet, while the right figure displays an /ən/ from the Chinese text of the same speaker. The boundary between the vowel and the nasal is much clearer in /ɛn/ than in /ən/, and F0 is mid-level in /ɛn/, but high level in /ən/, which was clearly perceived as Tone 1.

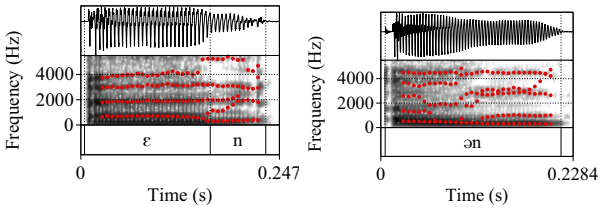

Figure 1: *Waveform, spectrogram, formant and annotation of /ɛn/ and /ən/ produced by a Shanghai female speaker.*

Similarly, Letter *H* was pronounced by more speakers as /eɪtʃ/ in English but /eɪtʃɤ/ in Chinese, which are illustrated in Figure 2 on the left and right respectively. A vowel epenthesis

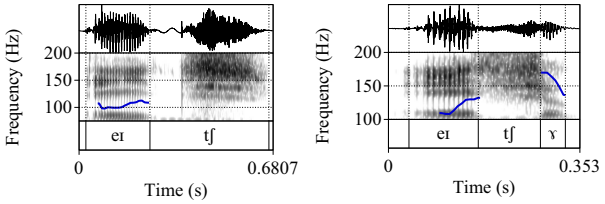

Figure 2: *Waveform, spectrogram, f0 and annotation of /eɪtʃ/ and /eɪtʃɤ/ produced by a Shandong male speaker.*

is discernable in the right figure. F0 goes down in the vowel epenthesis, like a lengthening of the rising tone of the first syllable, and was perceived as a neutral tone.

3.2.2. Tones

No difference in gender of speakers or position of MAWs in the sentence were found in the production of tones and phonemes of MAWs. Males and females only differed in F0 measures in Hertz. To accommodate the pitch range differences across male and female speakers, F0 was normalized in which F0 measures in Hertz obtained from each speaker were converted to their logarithms, using a formula commonly adopted for such purposes [14]:

$$T(X) = 5\frac{\lg X - \lg L}{\lg H - \lg L} \quad (1)$$

where H and L are the highest and lowest F0 for a given speaker, and X is any given point of a pitch contour. The output (T) is a value between 0 to 5, which is similar to the 5-point pitch scale for Mandarin tones proposed by Chao [15].

The problem was to define the highest and lowest F0 values. If the values are taken on the sentence level, the tone values at the sentence initial are much higher than those at the sentence final because of sentence declination effect. If they are taken from the word, no lowest values can be defined in a word with two high tones. In order to find both highest and lowest F0 values in a word, we selected *A B* followed by Tone 4 at the sentence initial position as an example. To counterbalance the differences in speaking rate, duration was normalized per segmental annotation across speakers. F0 of each utterance was estimated at 10 equal points of each annotation using the script developed by Xu [16]. The tone values can be observed in Figure 3, Figure 4, Figure 5, and Figure 6.

The pitch values of Letter *A* produced by Shandong speakers are slightly higher than those by Shanghai speakers, thus they sound more like Tone 1.

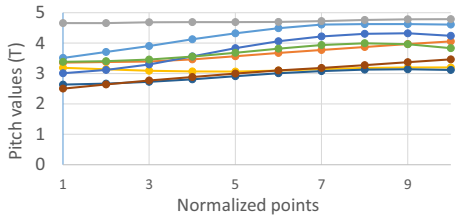

Figure 3: *Tone contours of Letter "A" by Shanghai speakers.*

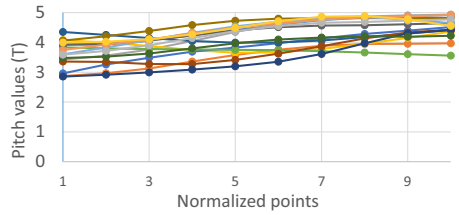

Figure 4: *Tone contours of Letter "A" by Shandong speakers.*

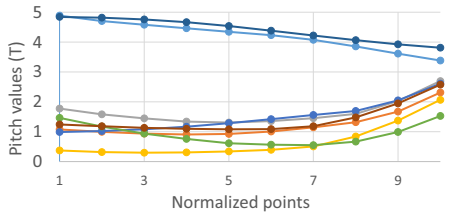

Figure 5: *Tone contours of Letter "B" by Shanghai speakers.*

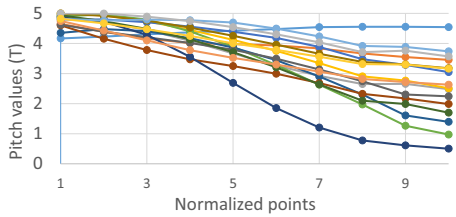

Figure 6: *Tone contours of Letter "B" by Shandong speakers.*

Because the syllable shown in Figure 3, 4, 5, and 6 is followed by a falling tone, the first syllable usually does not fall to the lowest level, nor does it rise to the highest, as shown in Figure 5 and Figure 6.

4. Discussion

On the basis of the results, several conclusions can be drawn on the phonetic adaption of Letters *A B H N O T*.

- The pronunciation of /eɪ/, /biː/, /əʊ/, /tiː/ exist in Mandarin Chinese; thus *A B O T* can keep their English pronunciation in Mandarin Chinese.
- All English /ɛn/ of *N* are changed to Chinese /ən/.
- In most cases, English /eɪtʃ/ has been changed to /eɪtʃɤ/, so that two syllables are resulted with CV structure, which conforms to the Chinese phonological system.
- Some letters tend to be pronounced in certain tones in Mandarin Chinese. For example, *A N O* are generally pronounced in Tone 1.
- Some letters have variations in tones in Mandarin Chinese, e.g. *B*. The variation is influenced by 1) dialects, 2) coarticulation, and 3) pronunciation in the English alphabet. We have found that speakers from northern China prefer falling tones, while speakers from Shanghai prefer rising tones. But still two Shanghai speakers tend to produce *B* in a rising tone before Tone 1 and 4, and in a falling tone before Tone 2 and 3. The reason seems to be that a higher beginning pitch of Tone 1 and 4 requires a previous rising tone for tonal coarticulation, while a lower beginning of Tone 2 and 3 demands a previous falling tone. Still another two Shanghai speakers pronounce all *B*s in Tone 4 in Chinese sentences. A further investigation reveals that they also pronounce *B* in a falling tone when they read the English alphabet.
- Mandarin speakers do not always transfer their pitch change patterns in English into Mandarin Chinese. Most speakers read the English alphabet with such intonation groups *ABCD, EFG, HIJK, LMN, OPQ, RST, UVW, XYZ*. Letters at the beginning of a phrase (for example, *B*) are pronounced in a high or a rising tone; while others at the end of a phrase (for example, *T*) are usually pronounced in a falling tone. However, in Chinese text, *B* is generally pronounced in a falling tone by Shandong speakers and *T* is predominantly pronounced in a high level tone by both Shanghai and Shandong speakers, which is different from the way they are read in the English alphabet.
- No investigated letters were read in Tone 3 in Mandarin.
- Tones are not influenced by the position of MAWs in the sentence. The reason may be that as a tone language, tone coarticulation plays an important role in tone contour modification; sentence intonation does not change the category of the syllable, but only lowers or raises the tone register and pitch range.

The conclusion is only based on the combination of LC MAWs and speakers from northern and central dialect zones. More data will be collected to cover wider categories of MAWs and larger areas of dialects for further investigation in the future.

5. Conclusion

In this paper, we report a preliminary study of phonological adaption when foreign orthography is incorporated. Our results show that the pronunciation of MAWs has violated the autonomy of Chinese phonology. This experiment on Chinese MAWs shows that there are complex interactions among orthography, perception, and phonological adaptation. It suggests that phonological integrity is less rigid than assumed, and that *dialects* as mother tongues can have strong influence on tonal adaptation. Both facts may have significant implications for linguistic theories beyond the seemingly idiosyncratic phenomena of MAWs and merit in-depth future studies.

6. Acknowledgements

Work by the first author was sponsored by the Interdisciplinary Program of Shanghai Jiao Tong University (14JCZ03), and work by the last author was supported by a RGC GRF grant (543512H) and the VariAMU international collaboration project by the Faculty of the Humanities of the Hong Kong Polytechnic University.

7. References

[1] I. Vendelin and S. Peperkamp, "The influence of orthography on loanword adaptations," *Lingua*, vol. 116, no. 7, pp. 996–1007, 2006.

[2] C.-R. Huang and H. Liu, "Corpus-based automatic extraction and analysis of Mandarin alphabetic words," *Journal of Yunnan University of Nationalities (Social Sciences, in Chinese)*, vol. (To Appear), 2017.

[3] P. R. o. C. The Institute of Linguistics of Chinese Academy of Social Science, Ed., *Modern Chinese Dictionary (in Chinese)*, 7th ed. The Commercial Press, 2012.

[4] Y. Liu, Ed., *The Dictionary of Lettered-words (in Chinese)*. Shanghai Lexicographic Publishing House, 2001.

[5] B. Jia, "Some considerations about and trial of transcribing words composed of Latin letters," *Applied Linguistics (in Chinese)*, vol. 3, pp. 77–80, 2000.

[6] X. Li, "The issue of pronunciation of lettered words in Chinese revisited," *Applied Linguistics (in Chinese)*, vol. 3, pp. 93–99, 2002.

[7] Y. Liu, "The issue of lettered words in Chinese," *Applied Linguistics (in Chinese)*, vol. 1, pp. 85–90, 2002.

[8] H. Ding, R. Hoffmann, and D. Hirst, "Prosodic transfer: A comparison study of f0 patterns in L2 English by Chinese speakers," in *Speech Prosody*, 2016, pp. 756–760.

[9] H. Riha, "Lettered words in Chinese: Roman letters as morpheme-syllables," in *OSU WPL*, 2010, vol. 59, pp. 44–51.

[10] C.-R. Huang, "Tagged Chinese gigaword version 2.0," LDC2009T14. Linguistic Data Consortium, 2009.

[11] J. Norman, *The Chinese dialects: phonology*, ser. Routledge Language Family Series. Routledge, London and New York, 2003, vol. The Sino-Tibetan languages, ch. 5, pp. 72–83.

[12] A. O. Yue, *Chinese dialects: grammar*. Routledge, London and New York, 2003, vol. The Sino-Tibetan languages, ch. 6, pp. 84–145.

[13] P. Boersma and D. Weenink. (2016) Praat: doing phonetics by computer [computer program]. Version 6.0.23, retrieved 24 December, 2016. [Online]. Available: http://www.praat.org

[14] M. Peabody, S. Seneff, and C. Wang, "Mandarin tone acquisition through typed interactions," in *ICALL*, 2004, pp. 173–176.

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

[16] Y. Xu. (2005-2017) Prosodypro.praat. [Online]. Available: http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/