ISCA Archive
http://www.isca-speech.org/archive

4th International Symposium on
Tonal Aspects of Languages (TAL-2014)
Nijmegen, The Netherlands
May 13-16 2014

The Effect of Orthography on L2 Perception

Robert Bo Xu, Jingwen Li, Peggy Mok

Department of Linguistics and Modern Languaegs, The Chinese University of Hong Kong
xuborobert@gmail.com, joanneljw@gmail.com, peggymok@cuhk.edu.hk

Abstract

Mandarin Chinese has two orthographic systems: Chinese characters and Pinyin. While Pinyin is transparent to Mandarin pronunciation, characters are opaque and seldom relate to sound. This study aims to find out the effect of these two systems on Cantonese listeners who are L2 learners of Mandarin. Native Hong Kong Cantonese speakers participated in word recognition experiments which included a monosyllabic task and a disyllabic task. The results show that Cantonese listeners most often confused T1 and T4, and T2 and T3 in Mandarin. Pinyin and characters have influence on the perception of Cantonese listeners. While Pinyin facilitates the recognition in the monosyllabic task, but characters facilitate the disyllabic task, indicating the two tasks may have involved different processes.

Index Terms: L2 perception, orthography, tone, Mandarin

1. Introduction

1.1. Chinese character and Pinyin

The primary writing system for modern Chinese is Chinese characters. It is widely used in Chinese communities no matter which Chinese language is mainly spoken in that area. Although there are differences between the simplified and traditional versions of Chinese characters, they share the attribute of being opaque, in the sense that they give little cue to what how the characters should be pronounced. Therefore, several romanization systems were proposed to directly represent the sounds of Mandarin Chinese. The most widely used one among them is Hanyu Pinyin, known as Pinyin for short, designed and promoted by the Chinese government. Pinyin is a transparent system with a one-to-one mapping to the sounds of Mandarin syllables. For example, in the syllable *quán* in Pinyin, "q" denotes the onset, "uan" denotes the nasal rhyme, and the diacritics above "a" shows that the syllable should be pronounced with Tone 2 (T2), a rising tone.

While Chinese characters are the major writing system in everyday life, their opaqueness gives rise to problems in learning. Pinyin, as a result, is introduced at the beginning of Chinese learning for both native and nonnative speakers to assist the acquisition of Mandarin sounds. Some studies have found that orthography played an important role in L2 learning. For instance, it helped the learners to retain the phonological representation of words in their memory [1]; it can also be used to account for the pronunciation difficulties at the sublexical level [2]. However, few studies focus beyond the alphabetic writing systems. Therefore, it is yet to find out whether an opaque system like Chinese characters will have a similar effect to other systems with a letter-to-sound mapping for L2 learners. In particular, tones are poorly represented in Chinese characters. This study will focus on the effect of orthographies of Mandarin Chinese, character and Pinyin, on the perception of tones by Cantonese listeners, to find out if Cantonese L2 learners of Mandarin would have different accesses to the Mandarin lexicon when they are presented with Pinyin or characters.

1.2. Tones in Mandarin and Cantonese

Although sharing the same orthographic system in Chinese characters, Mandarin and Cantonese have different phonologies. In terms of lexical tones, Mandarin has a contour tone system, while Cantonese has a more complicated system that differs in both tone contour and tone register.

There are four lexical tones in Mandarin [3], shown in Table 1. All of them differ in pitch shape. Tone 1 (T1) is a high level tone. T2 is a high rising tone. T4 is a falling tone. There are two allotones for T3. Produced as a citation form or before a boundary, it is a dipping tone with a rising tail; otherwise it is produced as a low tone [4].

There is a tone sandhi rule in Mandarin. In a T3-T3 sequence, the first T3 is produced as a high rising tone [5], perceptually indistinguishable from T2[6, 7], especially when the linguistic context that implies the occurrence of tone sandhi is lacking [8].

Table 1. *Lexical tones in Mandarin.*

	T1	T2	T3	T4
Tone Shape	Level	Rising	Dipping	Falling
Tone Letter	55	35	21(4)	51

Cantonese has a more complicated tonal system with six lexical tones [9, 10, 11]. As shown in Table 2, they differentiate in both pitch shape and pitch register, with three level tones (T1: high level, T3: mid level, T6: low level), two rising tones (T2: high rising, T5: low rising) and one low falling tone (T4).

Table 2. *Lexical tones in Cantonese.*

	T1	T2	T3	T4	T5	T6
Tone register	High	High	Mid	Low	Low	Low
Tone shape	Level	Rising	Level	Falling	Rising	Level
Tone Letter	55	25	33	21	23	22

Despite the difference in the tonal system, the two historically related languages have some regular rules of correspondence [12], as shown in Table 3 and Table 4. Table 3 shows the major correspondence of Cantonese tones to Mandarin tones; while Table 4 shows the opposite. Chu [13] has also shown that these sublexical (onset, rhyme, tone) mappings between Mandarin and Cantonese were important to Cantonese speakers' production and perception of Mandarin, especially for speakers with low proficiency in Mandarin, to whom the concept route (mappings through meanings) was often unavailable. This was evident in the results of his experiments as the production and perception error patterns correlated with the mapping percentages in [12].

Table 3. *Mandarin-Cantonese tone correspondence [12].*

Mandarin	Cantonese tones
T1[55]	T1 (84%), T3 (9%)
T2[35]	T4 (76%), T6 (13%)
T3[214]	T2 (60%), T5 (25)
T4[51]	T6 (47%), T3 (40%)

Table 4. *Cantonese-Mandarin tone correspondence [12].*

Cantonese Tone	Mandarin Tone	%Correspondence
T1[55]	T1[55]	93%
T2[25]	T3[214]	89%
T3[33]	T4[51]	91%
T4[21]	T2[35]	93%
T5[23]	T3[214]	76%
T6[22]	T4[51]	94%

1.3. Mandarin tones perceived by Cantonese listeners

Some previous studies have investigated the perception of Mandarin tones by Cantonese listeners. T2-T3 and T1-T4 have been found the most confusing tone pairs [14, 15, 16]. However, few studies have focused on the role of orthography. Chu [17] suggested that the shared orthography (Chinese characters) might be a source of negative transfer, because it might activate the L1 phonological representations, in conflict with those of L2. Because of the opaqueness of the Chinese characters, the listeners may unconsciously take the homophones in their L1 (Cantonese) as homophones in their L2 (Mandarin), giving rise to mispronunciations or misrecognition of Mandarin words.

Since Chinese characters and Pinyin provide two different approaches to represent Mandarin words, it is very possible that L2 learners may have different processes of access to their Mandarin lexicon when these two types of orthographies are presented. Therefore, this study aims to investigate this possible difference through the perception of Mandarin tone by Cantonese listeners.

2. Methodology

2.1. Materials

Monosyllabic and disyllabic words were prepared for the perception experiment. Within each group, there were two conditions: words presented in traditional Chinese characters, which Hong Kong speakers were accustomed to, and words presented in Pinyin.

For monosyllabic words, there were four minimal quartets, where each item was a possible word in Mandarin. Two groups were presented in Chinese characters, the other two in Pinyin. In total the listeners went through 16 items in the monosyllabic recognition task.

The disyllabic words were minimal pairs that only differ in the tone of one syllable, either the first one or the second. There were six possible tone-pair combinations in Mandarin, each presented visually in either characters of Pinyin. There were eight items for each condition, making the total number of the tested items 192 (2 different syllables × 6 tone pairs × 2 visual presentations × 8 items).

2.2. Participants

The participants were undergraduate students at the Chinese University of Hong Kong. All of them were native speakers of Hong Kong Cantonese. Most of them started to learn Mandarin through formal education at primary school, but with varying amount of Mandarin instruction. In addition, since Mandarin is not a compulsory subject in public examinations in Hong Kong, there was no objective measure to assess their Mandarin proficiency. They speak Cantonese in everyday life, English in classroom settings, and only speak Mandarin when necessary. None of the participants reported any history of hearing or speech impairment or loss. They all received course credits for taking part in this experiment. 49 listeners participated in the monosyllabic task, and 39 listeners among them also participated in the disyllabic task.

2.3. Procedure

The experiment took place in a sound-treated room using the E-Prime software. Instructions were given both visually on the screen and orally by the experimenter. The participants listened to a Mandarin word, and chose the corresponding word on screen in characters or Pinyin by pressing buttons on a response box.

After a practice session, the participants finished four blocks of different tasks. The first two blocks, one presented in characters, the other in Pinyin, required the listeners to choose the monosyllabic word they heard from four choices on the screen. For each trial, the fours choices were segmentally identical and differed in the tones only. In the following two blocks for disyllabic words, the two words in the minimal pairs (characters in one block, Pinyin in the other) were shown on the screen for the listeners to choose from. All the items were randomized in the blocks.

3. Results

3.1. Perception of monosyllabic words

A pair-wise t-test was conducted to compare the means of the average accuracy rate of the Pinyin condition and character condition by all the listeners. The results show that the listeners perceived significantly better ($p<0.001$) in the Pinyin condition (M=91.1%, SD=0.11) than in the character condition (M=76.3%, SD=0.22). This pattern indicates that Pinyin can help the listeners in recognizing the monosyllabic tones.

Table 5 shows the overall error rate by all the listeners. In general, the Cantonese listeners performed well in identifying monosyllabic tones, making less than 3% of errors even in the most confused tone pair (T3 mistaken as T2). Most errors occurred between T2 and T3, T2 and T4, and T1 and T4.

Table 5. *Overall error rate (%) by all the listeners.*

		Response tone			
		T1	T2	T3	T4
Stimulus tone	T1		0.89	1.02	0.26
	T2	0.38		2.04	1.40
	T3	1.66	2.93		0.38
	T4	2.04	2.04	0.77	

Table 6 and Table 7 show the error rates by all the listeners under the Pinyin and character condition. Comparing data presented in the two tables, it can be observed that listeners did better in the Pinyin condition, with fewer mistakes. In the Pinyin condition, listeners misidentified T4 as T1. However, in the character condition, listeners made most confusions in the T2-T3 tone pair.

Table 6. *Error rate (%) in the Pinyin condition.*

		Response tone			
		T1	T2	T3	T4
Stimulus tone	T1		0	0	0.26
	T2	0		0.51	0
	T3	0.77	1.53		0.26
	T4	3.06	0.77	1.28	

Table 7. *Error rate (%) in the character condition.*

		Response tone			
		T1	T2	T3	T4
Stimulus tone	T1		1.79	2.04	0.26
	T2	0.77		3.57	2.81
	T3	2.55	4.34		0.51
	T4	1.02	3.32	0.26	

3.2. Perception of disyllabic words

A three-way ANOVA with repeated measures was conducted based on the perception accuracy of the disyllabic words. The three within subjects variables are TonePair (6 levels), Syllable (which syllable, A or B, has the different tone, 2 levels), and Orthography (character or Pinyin, 2 levels). Corrections for violations of sphericity were made, where appropriate, using the Greenhouse–Geisser estimates of sphericity. Bonferroni correction was applied to make pairwise comparison. With consistent results from non-parametric analysis concerning the variables Orthography and TonePair, the skewedness of the current data is expected to be compensated by the robustness of ANOVA.

The orthography of presentation (character or Pinyin) was found to have a main effect on the perceptual results [$F(1,37)=33.54$, $p<0.001$]. However, contrary to the results of the monosyllabic words, the tones of the disyllabic words presented in characters ($M=96.33\%$, $SD=0.08$) are better discriminated than those in Pinyin ($M=91.84\%$, $SD=0.12$).

The results also show that TonePair has a significant main effect [$F(3.21,118.59)=27.39$, $p<0.001$]. The T2-T3 pair has the lowest recognition accuracy. In pair-wise comparisons, the accuracy is significantly lower than the other tone pairs ($p<0.001$ in every comparison). In addition, the accuracy of T1-T4 pair and T3-T4 pair is significantly lower than the T2-T4 pair, which is the best discriminated tone pair ($p=0.01$ for the comparison between T1- T4 pair and T2-T4 pair, and a marginal difference at $p=0.053$ for the comparison between T3-4 pair and T2-4 pair).

There is a significant effect of Syllable [$F(1,37)=11.50$, $p=0.002$]. The accuracy of the words where the second syllable differs in tone ($M=94.92\%$, $SD=0.10$) is slightly better recognized than those where the first syllable differs ($M=93.25\%$, $SD=0.11$).

The interaction between TonePair and Orthography is significant [$F(5,185)=15.873$, $p<0.001$], indicating that the difference between orthographies of presentation depends on the tone pairs. Besides, the two-way interaction Orthography × Syllable [$F(1,37)=6.35$, $p=0.016$], as well as the three-way interaction TonePair × Orthography × Syllable [$F(3.24,119.86)=4.35$, $p=0.005$] are also significant.

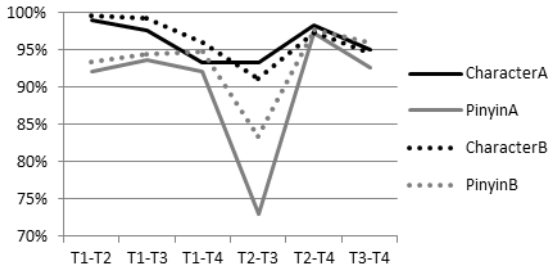

Figure 1: *Average accuracy for each condition. (A/B: tone differed in the first/second syllable)*

The results of the statistical analysis are reflected in Figure 1. The accuracy rates of the character items (lines in black) are higher than those of the Pinyin items (lines in grey) in most cases. The most striking feature in the figure is that the low accuracy for the T2-T3 pair, showing that this is the most difficult tone pair to be discriminated by Cantonese listeners, especially when the words were visually presented in Pinyin and when the T2-T3 minimal difference occurred in the first syllable.

4. Discussion

The results of the monosyllabic task and the disyllabic task both show significant effects of orthography. However, a discrepancy between the two tasks is the importance of Pinyin or Chinese characters. For the monosyllabic tones, items presented in Pinyin were significantly better perceived than those in characters. However, for the disyllabic tones, items presented in characters were significantly better recognized than those in Pinyin. This contrary pattern indicates that listeners might retrieve information differently in monosyllabic and disyllabic words.

In the monosyllabic task, what the listeners were asked to do was to choose the corresponding tone from all four possible answers. Therefore, they did not need to access the conceptual route; rather, they only paid attention to the different tones they were hearing. As a result, Pinyin, a transparent way to represent the sounds of the syllables, would have an advantage in being correctly recognized. In addition, the four choices on the screen had the same letters for onset and rhyme. The only difference lay in the diacritics that signify lexical tones, which was the only thing listeners needed to pay attention to. The characters, on the other hand, would be unrelated to sounds, and look quite different from each other orthographically. It might take more effort for the subjects to find the characters corresponding to the sounds they heard.

In contrast, the facilitating effect of pinyin over characters has disappeared in the disyllabic task. In fact, although words in characters were better recognized than those in Pinyin, the difference in the disyllabic task is much smaller than the difference in the monosyllabic task. One possible reason is that when listening to disyllabic words which were meaningful, listeners tended to perceive the words holistically. As a result, they matched the characters with the meaning they heard. In this way, unlike the monosyllabic task, the disyllabic task involved both a bottom-up process and a top-down process. Another possibility, derived from Chu's study [17], is that the high correspondence between Mandarin tone and Cantonese tone as shown in Tables 3 and 4 may facilitate the perception with characters, as listeners may instantly gained access to the Cantonese pronunciation from the written characters and mapped them onto the Mandarin sounds they heard. On the other hand, Pinyin, which only denotes the sounds, does not

allow the listeners to take advantage of such tone mapping. Furthermore, reading Pinyin requires an understanding of the phonological knowledge of Mandarin as well as adept skills in transcribing Mandarin sounds in Pinyin. Both skills might be lacking by the Cantonese listeners, as they probably have neither enough phonological awareness for Mandarin, nor enough familiarity with Pinyin to do well in this task. In addition, their phonological knowledge in Cantonese did not help them to relate to the Mandarin phonological patterns presented in Pinyin. More importantly, since there are many more disyllabic tokens, and they involved more complicated segmental combinations than the monosyllabic task (8 tokens vs. 96 tokens), the listeners had to put in much more effort in forming the final pronunciation from Pinyin. As a result, recognizing words in Pinyin in the disyllabic task is much more difficult than that in the monosyllabic task, resulting in the low accuracy in the disyllabic recognition. In a parallel project [18], it was found that Cantonese speakers with both high and low proficiency in Mandarin read disyllabic words in Pinyin slower than those in characters. This side evidence supports the idea that the Cantonese speakers really needed some effort to put together the Pinyin symbols in order to figure out the pronunciations. In sum, the different effects of orthographies on monosyllabic and disyllabic words show that they have an influence on the information retrieval of lexical items.

5. Conclusion

This study shows that orthography does have an influence on L2 perception. However, the effect of orthography varies based on the perception materials. In the task to recognize monosyllabic tones, Pinyin facilitated the identification of tones by transparently representing the sound. In the task to recognize disyllabic tones, where meaning becomes an important factor in retrieving the lexical items, characters helped the listeners to activate the L1 lexicon, which in turn would facilitate the access to L2 phonology because of the high correspondence of tones. This result points to future in-depth investigations into the role of non-alphabetical orthography in native and non-native lexicon access, in which more balanced number of materials should be used in order to minimize the bias in the results.

6. References

[1] Young-Scholten, M. and Archibald, J., "Second language syllable structure", in J. Archibald [Ed], Second Language Acquisition and Linguistic Theory, 64-97, Blackwell, 2000.

[2] Silveria, R., "Investigating the role of orthography in the acquisition of L2 pronunciation: a case study", in M. A. Watkins, et al. [Ed]. Recent Research in Second Language Phonetics/Phonology, 270-290, 2009.

[3] Chao, Y. R. "Tone and intonation in Chinese." Bulletin of the Institute of History and Philosophy 4:121–134, 1933.

[4] Duanmu, S. "The Phonology of Standard Chinese". Oxford University Press, 2007.

[5] Shih, C. "The Prosodic Domain of Tone Sandhi in Chinese". Ph.D dissertation, University of California, San Diego. 1986.

[6] Peng, S. "Lexical versus 'phonological' representations of Mandarin sandhi tones", in M. B. Broe and J. Pierrehumbert [ED], Acquisition and the lexicon: Papers in Laboratory Phonology V, 152–167, Cambridge University Press, 2000.

[7] Wang, W. S.-Y., and Li K. "Tone 3 in Pekinese", Journal of Speech and Hearing Research 10(3):629–636, 1967.

[8] Speer, S. R., Shih, C, and Slowiaczek M. L. "Prosodic Structure in Language Understanding: Evidence from Tone Sandhi in Mandarin", Language and Speech 32 (4): 337–354, 1989.

[9] Vance, Timothy J. "Tonal distinction in Cantonese", Phonetica 34: 93-107, 1977.

[10] Bauer, R. S., and Benedict P. K. Modern Cantonese phonology. Mouton de Gruyter, 1997.

[11] Kao, D. L. Structure of the Syllable in Cantonese. The Hague, Mouton, 1971.

[12] Zhang, L., & Gao, S. "Putonghua zi yin ren ji xun lian 12 jiang", [12 Lecture series of the recognition and memorization of Mandarin pronunciation]. Joint Publishing Co. Ltd. 2000.

[13] Chu, P. C. K. Interlanguage speech intelligibility benefit : implications for a model of second language word production and recognition. PhD Dissertation. New South Wales University, 2013.

[14] Lee, Y.-S, Vakoch, D. A., and Wurm, L. H. "Tone perception in Cantonese and Mandarin: a cross-linguistic comparison," J. Psycholinguist. Res., 25(5), 527–542, 1996.

[15] Hao, Y.-C. "Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers," J. Phon., vol. 40, no. 2, pp. 269–279, 2012.

[16] C. K. So and C. T. Best, "Cross-language Perception of Non-native Tonal Contrasts: Effects of Native Phonological and Phonetic Influences," Lang. Speech, vol. 53(2), 273–293, 2010.

[17] Chu, P. C. K. "Towards a Model of Second Language Word Production and Recognition in Mandarin". Proceedings of the International Conference on Chinese Language Learning and Teaching in the Digital Age, Hong Kong, China, 2011.

[18] Li, J., Xu, R. B., and Mok P. P. K. "Effect of orthography on L2 production of Mandarin tones", submitted to The 4th International Symposium on Tonal Aspects of Languages, 2014.