Does Orthography Affect L2 Tone Production and Perception?

Donghui Zuo, Qiuyue Chen, Peggy Mok

Department of Linguistics and Modern Languages, The Chinese University of Hong Kong
donghui@cuhk.edu.hk, collachen@gmail.com, peggymok@cuhk.edu.hk

Abstract

Many studies have investigated the production and perception of second language learners. However, very few of them have discussed the effect of orthography. Although Mandarin and Cantonese share the same orthographic system, the phonological systems of the two languages are quite different. This preliminary study investigates the production and perception of Mandarin tones by Cantonese learners, and compares their error patterns in two conditions: when the subjects were presented with 1) stimuli written in Mandarin Pinyin (a transparent orthography), 2) stimuli written in Chinese characters (an opaque orthography). The result shows that orthography has different effects in tone production and perception: the Pinyin system facilitated tone production only, while the subjects performed significantly better with Chinese characters in perception. Possible accounts for the observed differences are discussed.

Index Terms: Mandarin, L2 production and perception of tones, orthography

1. Introduction

The Chinese writing system is regarded as an opaque orthography, as there is no isomorphic letter-to-phoneme correspondence in Chinese characters. For example, the word 馬 (/ma/ with a falling-rising tone) 'horse' was derived from a pictograph, and therefore gives no clue about its pronunciation. To annotate the pronunciations of Chinese characters, Mandarin Pinyin is used as the official Romanisation system of Mandarin Chinese. For example, the word 馬 'horse' is written as *ma3*, where the number 3 stands for the tone (the falling-rising tone). This system is a transparent orthography. It is the first thing every learner learns in his or her Mandarin class and it is used as one of the major input methods of Chinese characters. Nevertheless, although all Mandarin speakers and learners are familiar with the Pinyin system, Chinese character is used as the major writing system. As a result, the opaqueness of Chinese characters may be an obstacle for L2 production and perception. Some studies have investigated the role orthography plays in second language acquisition. For example, Young-Scholten and Archibald [1] stated that written representations help the learners retain the phonological information of words in memory. Silveria [2] found that orthography can be used to account for the pronunciation difficulties learners face in producing word-final consonants.

However, all the previous studies on L2 orthography focused on alphabetic writing systems, which more or less show some letter-to-phoneme correspondence. It is thus not surprising that orthography can facilitate pronunciation in their studies. Chinese characters, on the other hand, are mostly logographic, which adds to the opaqueness of the orthography. Particularly, tone is not represented in Chinese characters at all. It is unclear whether this type of orthography can hinder or help L2 production and perception of tones.

Many studies have investigated the L2 acquisition of Mandarin tones by Cantonese learners, including both production and perception [e.g. 3, 4]. Nonetheless, very few of them have discussed the role of orthography. Chu [4], in his work on the production and perception of Mandarin tones by Cantonese learners, suggested that the shared orthography of Cantonese and Mandarin is a source of negative transfer, because it can activate the L1 phonological representations. Due to the opaqueness of the Chinese characters, the learners may unconsciously take the homophones in their L1 as homophones in L2. This is where mispronunciations or misrecognition may stem from.

Despite having the same orthography, Mandarin and Cantonese have two different phonological systems. In terms of tones, Mandarin has a four-tone system while Cantonese has a six-tone system (the three checked tones of syllables with final stops are regarded as allotones of T1, T3 and T6 respectively). Table 1 shows the inventories and the pitch values of tones in Cantonese and Mandarin respectively,

Table 1: *Pitch values of the Cantonese and Mandarin Tones on a 5-point scale from low (=1) to high (=5)*

Tones	Cantonese	Mandarin
T1	55	55
T2	25	35
T3	33	214
T4	21	51
T5	23	-
T6	22	-

Although the tone systems differ from each other, there are some regular correspondence rules between Mandarin and Cantonese tones [5] (see Table 2). For example, 93% of Cantonese T1 syllables are pronounced as T1 in Mandarin, and 89% of Cantonese T2 words are pronounced as T3 in Mandarin. Checked tones are not listed here, as they do not have clear corresponding relationship with any particular Mandarin tones.

Table 2: *Cantonese-Mandarin tone correspondence [5]*

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%

Chu [4] used a set of experiments to show that Cantonese learners of Mandarin are aware of the tonal correspondence rules, and they use them in Mandarin word production and perception. Chu argued that the learners make use of three different routes in Mandarin word production and perception – concept route, lexical route and sub-lexical route

(see Figure 1). The concept route is a direct link between the concept and the L2 Mandarin phonological representation. This route is usually used only by advanced learners. Beginning learners rely more on the lexical or sub-lexical routes in which the L1 Cantonese phonological system are involved. The lexical route first activates the L1 Cantonese phonological representation and then the L2 Mandarin phonological representation. The sub-lexical route retrieves the L1 Cantonese sub-lexical information (i.e. onsets, rimes and tones) after the L1 Cantonese phonological representation is activated. Then the Cantonese sub-lexical representations activate their Mandarin counterparts to varying degrees according to the different probabilities of the Cantonese-Mandarin correspondence rules, and finally the L2 Mandarin sub-lexical representations are combined to form the L2 Mandarin phonological representations.
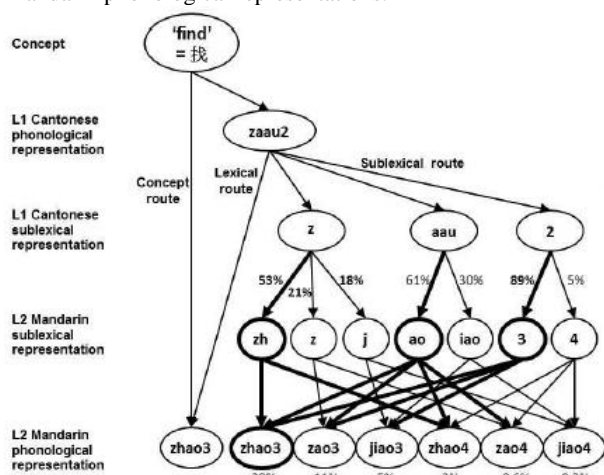

Figure 1: *Mandarin word production model of Cantonese learners [4]*

The perception process is similar to the production one. The only difference is that the representation levels are upside down in the perception model. The one at the top is the L2 Mandarin phonological representation and the one at the bottom is the concept level.

Chu's model can account for the negative homophonic transfer from Cantonese. We can see that the shared orthography plays an important role in the production and perception of L2 Mandarin, as in an opaque orthographic system, the one-to-one relationship activates the L1 Cantonese lexical system and triggers transfer. Then, a logical question arises: does such activation always hinder L2 production and perception? The present study is a preliminary investigation of the effects of orthography on L2 tonal production and perception. It compares the production and perception of Mandarin tones by Cantonese learners in two conditions: 1) when the subjects are presented with Mandarin Pinyin, a transparent orthography; 2) when the subjects are presented with Chinese characters, an opaque orthography. The error patterns in the two experiments are discussed, and the differences between the two conditions are highlighted.

2. Method

2.1. Subjects

Sixteen native speakers of Hong Kong Cantonese (9 females, 7 males; aged from 19 to 29, mean = 23) participated in this study. All the subjects were university students who were beginning learners of Mandarin. Most of them learned Mandarin through formal education in primary school. Their average length of formal training in Mandarin is 2.6 years.

2.2. Materials

The study was composed of two experiments – production and perception. The materials of the two experiments are different in order to avoid practice effect.

2.2.1. Production

The production experiment consisted of two tasks: a Pinyin task and a Chinese character task.

In the Pinyin task, two sets of Pinyin stimuli were used. The syllables *mi* [mi] and *na* [na] with all four tones were selected as the target sounds, which resulted in 8 tokens (2 syllables × 4 tones). All the eight tokens are attested syllables in Mandarin and can stand alone as monosyllabic words except for *na1* which is of relatively low frequency and can only appear as part of a word. All the stimuli were written in Pinyin.

In the Chinese character task, two sets of Chinese character stimuli were used. The syllables *ya* [ja] and *wu* [wu] with all four tones were selected as the target sounds. That resulted in 8 tokens (2 syllables × 4 tones). All the syllables were written in Chinese characters which share the same meanings in Mandarin and in Cantonese.

2.2.2. Perception

The perception experiment also consisted of a Pinyin task and a Chinese character task.

In the Chinese character task, two sets of disyllabic Chinese character stimuli which differed only in the tones of the second syllable were used (e.g., T1-T3 pair: 麻衣 *ma2 yi1* 'linen cloths' vs. 螞蟻 *ma3 yi3* 'ant'). Each set contained 6 minimal pairs of all possible confusable tone combinations (i.e., T1-T2, T1-T3, T1-T4, T2-T3, T2-T4, T3-T4). This resulted in 12 minimal pairs (2 sets × 6 pairs).

In the Pinyin task, one set of disyllabic Pinyin stimuli was used. Similar to the Chinese character task, the stimuli differed only in the tones of the second syllables (e.g. T1-T3 pair: can1 *guan1* 參觀 'visit' vs. can1 guan3 餐館 'restaurant'). The stimulus set contained 6 minimal pairs of all possible confusable tone combinations. This resulted in 6 minimal pairs (1 sets × 6 pairs). Chinese characters are not shown in the Pinyin task.

All the 36 stimuli (18 pairs × 2 members) were legal words in both Mandarin and Cantonese, and shared the same lexical meanings in these two languages. The stimuli were produced by a female native Mandarin speaker.

2.3. Procedures

The experiments were conducted in a quiet room at the Chinese University of Hong Kong. The subjects completed the production experiment before the perception experiment.

2.3.1. Production

The stimuli were presented to the subjects on a piece of paper. They were asked to produce the Pinyin stimuli first and then the Chinese character stimuli, all in isolation. The recordings were taken with a solid state recorder with a sampling rate of 44100Hz.

128 Pinyin tokens (2 syllables × 4 tones × 16 speakers) and 128 Chinese character tokens (2 syllables × 4 tones × 16 speakers) were collected. All the recordings were transcribed by three native Mandarin speakers with training in phonetics. The agreed transcriptions of at least two transcribers were accepted as the actual tones produced by the speakers. If the three transcribers did not agree with each other, the file was further checked by the second author who decided which of the three transcriptions was mostly auditorily similar to the sound produced by the subjects. Table 3 shows the numbers of tokens the transcribers agreed on.

Table 3: *Numbers of the tokens the transcribers agreed on*

	No. of tokens	Percentage
3 transcribers agree	173	67.6%
2 transcribers agree	80	31.3%
No agreement	3	1%

2.3.2. Perception

The stimuli were played to the subjects through headphones. They were asked to write down their answers after listening to the sound files.

The Pinyin task was a transcription task. The subjects were given a set of Pinyin without tones (e.g. *can guan*) and were asked to write down the tones of the second syllables (i.e. *can guan1* or *can guan3*) after listening to the stimuli. Both words of each minimal pairs were presented once. 192 responses in total (6 pairs × 2 words × 16 subjects) were collected in this task.

The Chinese character task was a two-alternative forced choice task. The subjects were asked to choose the words they heard from two given options (e.g. 字跡 / 自己 [*zi4 ji4* 'handwriting' / *zi4 ji3* 'oneself']). Both words of each minimal pairs were presented once. The order of the words was randomised. 384 responses (12 pairs × 2 words × 16 subjects) were collected in this task.

3. Results

3.1. Production

Only three production errors were found among the 128 tokens collected in the Pinyin task – one T1 was mispronounced as T4, one T3 as T2, and one T4 as T1. The overall error rate (%Err) is 2.3%. By contrast, the overall %Err in the Chinese character task is 43.4%. The %Err of each tone in the Chinese character task is shown in Table 4.

Table 4: *Percentage of each tone being mispronounced as other tones in the Chinese character production task*

	T1	T2	T3	T4
T1		0	0	34.4
T2	3.1		37.5	3.1
T3	3.2	45.2		6.5
T4	28.1	12.5	0	

The data from the Chinese character task confirms that Cantonese learners of Mandarin are most likely to confuse T1 with T4, and T2 with T3 – T1 was most often mispronounced as T4 (34.4%), and T4 was most often mispronounced as T1 (28.1%). T2 was pronounced as T3 in 37.5% cases, and T3 as T2 in 45.2% cases.
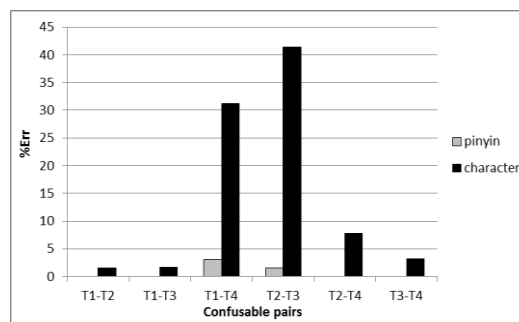

Figure 2: *Collapsed %Err of all the confusable tone pairs in the production experiments*

Figure 2 shows the collapsed %Err of all the confusable tone pairs (e.g., T1-T2 pair includes both mispronouncing T1 as T2 and T2 as T1). Comparing the production data from the two tasks, we find that the subjects made almost no mistake in the Pinyin task, while their error rates in the Chinese character task were quite high. Paired sample t-test shows that orthography has a significant effect on the %Err in the production experiment (t(15) = -7.678, p < 0.001), which suggests that the Pinyin system helped the subjects to produce the tones correctly.

3.2. Perception

The perception experiment also shows that Cantonese learners of Mandarin confuse T1 with T4, and T2 with T3 most, especially the latter pair (see Table 5 and Table 6 for the error rates of each tone in the two tasks). 33.3% tokens of T2 were misperceived as T3, and 31.3% tokens of T3 were misperceived as T2 in the Pinyin task. The error rates of the Chinese character task were lower than those of the Pinyin task. But T2 and T3 was still found to be the most confusable pair in the Chinese character task– 12.5% tokens of T2 were misperceived as T3 and 11.5% tokens of T3 were misperceived as T2. One interesting finding in the Chinese character perception task is that T1 and T4 did not appear to be confusing, which is quite different from the findings in the production tasks.

Table 5: *Percentage of each tone being misperceived as other tones in the Pinyin perception task*

	T1	T2	T3	T4
T1		18.8	4.2	20.8
T2	16.7		33.3	14.6
T3	6.3	31.3		14.6
T4	18.8	4.2	4.2	

Table 6: *Percentage of each tone being misperceived as other tones in the Chinese character perception task*

	T1	T2	T3	T4
T1		0	3.1	0
T2	0		12.5	4.2
T3	5.2	11.5		1.0
T4	6.3	1.0	2.1	

Regarding the effect of orthography, Figure 3 shows the collapsed %Err of all the confusable tone pairs in the two perception tasks. It is found that there is also asymmetry between the Pinyin task and the Chinese character task in perception. But contrary to the production part, the subjects performed significantly better in the Chinese character task than in the Pinyin task. The overall %Err of the Pinyin task

was 47.9%, while the %Err of the Chinese character task was only 11.7%. Paired sample t-test shows significant difference between the two tasks (t(15) = 5.495, p < 0.001). The large discrepancy between the error rates of the two tasks suggests that, contrary to production, the Chinese character system helped the subjects to perceive the tones better than the Pinyin system.
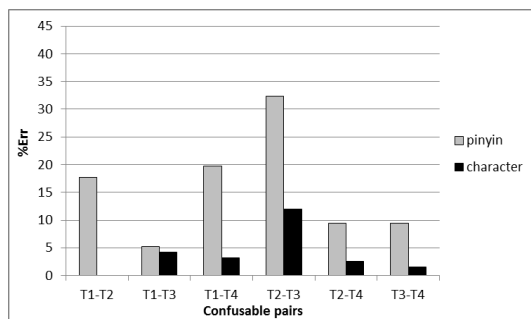

Figure 3: *Collapsed %Err of all the confusable tone pairs in the perception experiments*

As the Pinyin task required the subjects to write down the tone instead of asking them to choose from two options, there were actually four possible answers (i.e. T1, T2, T3 and T4) in the Pinyin task, while the Chinese character task only provided the subjects with two options. In order to make the two sets of data more comparable, we excluded the unintended errors (i.e., misperceiving the target words as neither member of the respective minimal pairs, which does not construct meaningful words, e.g., perceiving the second syllable as T2 in the T1-T3 pair *can1 guan1* 參觀 'visit' vs. *can1 guan3* 餐館 'restaurant') and calculated the %Err again. The overall %Err was reduced to 37.5%, but it was still significantly higher than that of the Chinese character task (t(15) = 4.667, p < 0.001).

4. General discussion

The current study confirms the findings in previous studies on L2 production and perception of Mandarin tones: T1-T4 and T2-T3 are the two most confusing pairs for L2 learners, especially the T2-T3 pair. The subjects made most errors in T2-T3 pair in both production and perception. This can be explained by the acoustic similarities between the two tones, e.g., [6]. It is also possible that due to the presence of T3 Sandhi in Mandarin (i.e., when a T3 is followed by another T3, the first T3 is realised as T2), learners may take T2 as the underlying tone of some T3 syllables or T3 as the underlying tone of some T2 syllables. A third explanation of this phenomenon is related to transfer. According to the Perceptual Assimilation Model [7], L2 learners tend to perceptually assimilate the L2 phones to their native phonemes. This theory is also applicable to the acquisition of tones. Since Mandarin T3 [214] is a contour tone with no equivalence in the Cantonese tone inventory, Cantonese learners may assimilate the Mandarin T3 to either a low falling (i.e., Cantonese T4 [21]) or high rising tone (i.e., Cantonese T2 [25]) which is acoustically similar to part of the Mandarin T3. As the Mandarin T2 [35] is also a high rising tone, when the speakers produce a Mandarin T3 word with Cantonese T2, it will sound like a Mandarin T2 instead of T3.

It is worth noticing that Cantonese learners of Mandarin tend to confuse T1 and T4 in production (only the Chinese character task is discussed, as almost no errors were found in

the Pinyin task), while this pair is easy to discriminate by learners of other language backgrounds like English [6]. Matthews and Yip [8] suggest that this is due to the fact that the Cantonese T1 [55] has a high falling allotone [53] so that native speakers do not distinguish the two realisations (level vs. falling) in Cantonese. It is likely that the subjects were influenced by their native phonology, and did not treat these two tonal contours (i.e., high level vs. faling) as contrastive in Mandarin. Therefore, when they produce the Mandarin T1 or T4, they tend to mix these two tones up. However, the results from the perception experiment, the Chinese character task in particular, show a different picture: the error rates of the T1-T4 pairs were not significantly higher than those of the other confusing pairs. This may indicate that the subjects can hear at least some difference in the pitch directions of the two Mandarin tones. In addition, the correspondence rules between Cantonese and Mandarin tones mentioned above may have also helped the learners discriminate Mandarin T1 and T4 in perception.

Regarding the effect of orthography, the current study found that orthography plays an important role in both the L2 production and perception. The opaque orthography (i.e., Chinese characters) hinders tone production, while it facilitates tone perception.

It is easy to understand why the subjects made more errors in the Chinese character production task than in the Pinyin production task. According to Chu's model [4] mentioned in Introduction, most beginning learners go through the sub-lexical route before they retrieve the L2 Mandarin phonological representations in the Chinese character task. Although most of the words used in the experiment abide by the Cantonese-Mandarin correspondence rules, the beginning learners may have not yet established the correct correspondence. Therefore, the use of sub-lexical route causes a lot of errors in their production. In contrast, when the subjects were given the words in Pinyin, since the transparent orthography shows the pronunciations, the subjects could retrieve the correct L2 Mandarin phonological representations without involving L1 Cantonese phonology, which lead to a much lower %Err.

However, it needs to be pointed out that the Pinyin task may not involve the link between L2 phonological representations and the concepts at all, while the Chinese character task did. In other words, the subjects may not know the meaning of the words they were reading in the Pinyin task. It is the same situation with some transparent orthographies like German or Spanish – due to the consistent letter-to-phoneme mapping, many beginning learners can read given sentences fluently before they can understand the meaning of the sentences, but when being asked to make spontaneous speech, they make a lot of mistakes. So if we take the link between concept and sound into consideration, it is hard to say whether a more transparent orthography facilitates speech production unilaterally: it certainly helps in word list reading tasks, but if any spontaneous speaking task is involved, the subjects still need to start from the concept level and go through the L1 lexical system before they retrieve the correct L2 phonological representations. A more transparent orthography may not be helpful in these situations.

As for the perception part, the Pinyin perception task also involves fewer representation levels than the Chinese character perception task, as the subjects just need to extract the tones from the L2 sub-lexical representation level in the Pinyin task. In contrast, they have to further process the meaning down through their L1 lexical system in the Chinese character task so as to get the correct characters. It is expected

that the subjects will do better in the Pinyin task than in the Chinese character task. Nonetheless, contrary to our expectation, the subjects made fewer mistakes in the Chinese character task than in the Pinyin task, which suggests that the use of L1 lexical system actually facilitates perception.

There are two possible accounts for the better performance in the Chinese character perception task. The first possibility is that since the possible answers were given as options, when the subjects saw the Chinese characters, they retrieved the concepts at once and formed certain expectancy of the perceived sounds. That is to say, it is likely that the Chinese character task involves not only a bottom-up perception process, but also a top-down process, which activates the target Mandarin representations to a stronger degree and helps the subjects to perceive the stimuli as the target tone. Such a top-down process is not involved in the Pinyin task, as the Pinyin task might only activate the L2 Mandarin representations, not the L1 Cantonese representations or the concepts. The subjects are more susceptible to the auditory resemblance of some acoustically similar tone pairs.

The second possibility lies in the different settings of the Pinyin and Chinese character tasks – a transcription task for the Pinyin part and a two-alternative force choice task for the Chinese character part. Even though we excluded the unintended errors in the data analysis, it is still possible that the two tasks involve different processing styles, which renders the transcription task more difficult than the force choice task, and thus causes more errors.

Comparing with the similarities and differences between L1 and L2 phonology, the effect of orthography on L2 speech learning is less studied. However, our study demonstrates that orthography does play an important role in both L2 production and perception. Therefore, future models of second language acquisition may need to take orthography into consideration, especially when the L1 and L2 are two closely related languages which share the same orthography (e.g., Mandarin and Cantonese), or when the orthographies of the two languages have a set of relatively consistent corresponding rules (e.g., Spanish and Portuguese).

One limitation of this preliminary study is that the stimuli were not well controlled in terms of word frequency, segmental features, character strokes and Cantonese-Mandarin correspondence. It is possible that some difference in the word frequency between members of minimal pairs caused bias in production and perception. It is also possible that tonal errors are correlated with certain segmental features in addition to orthography. Possibilities like these cannot be eliminated with the present data. Also, the production and perception experiments used different stimulus sets (e.g., the production experiment used monosyllable words, while the perception experiment used disyllabic words), which undermines the comparability of the results from the two experiments. In future studies, all these factors should be better controlled, and more consistent experiment settings should be used in order to further explore the effects of orthography on L2 production and perception.

5. Acknowledgements

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, 2000.

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

[3] Leung, A., Tonal assimilation patterns of Cantonese L2 speakers of Mandarin in the perception and production of Mandarin tones. Proceedings of the 2008 CLA Annual Conference, 2008.

[4] Chu, P. C. K. Towards a Model of Second Language Word Production and Recognition in Mandarin. Young Scholar Award Competition for the International Conference on Chinese Language Learning and Teaching in the Digital Age, Hong Kong, China, 2011.

[5] Tsang-Cheung, L.Y. The Phonological Correspondences between Cognate Morphemes in Cantonese and Mandarin. Unpublished M. Phil. Thesis, University of Hong Kong, 1988.

[6] Kiriloff, C., "On the auditory discrimination of tones in Mandarin", Phonetica 20:63-67, 1969.

[7] Best, C. T., A direct realist view of cross-language speech perception, In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross- language research, 171-204, 1995.

[8] Matthews, S., and Yip, V., Cantonese: a comprehensive grammar, 1994.