



## The role of orthography in L2 segment and tone encoding by learners at different proficiency levels

Yen-Chen Hao<sup>1</sup>, Chung-Lin Yang<sup>2</sup>

<sup>1</sup>University of Tennessee, U.S.A.

<sup>2</sup>Indiana University, U.S.A.

yhao6@utk.edu, cyl@indiana.edu

### Abstract

The present study compared the effect of two orthographic inputs on English speakers' phonological encoding of Mandarin words: the Chinese Romanization *Pinyin* and Chinese characters. English speakers at three proficiency levels participated in a word-learning experiment. During the learning phase, half of the participants saw Pinyin of the target words (PY group), while the other half saw characters (CH group). After learning, the participants judged the matching of sound and meaning of 128 pairs, half of which were complete matches, while the other half mismatched the target either in segment or tone. The results showed that the Advanced learners in the CH group were more accurate than their counterparts in the PY group in rejecting tonal mismatches. In contrast, the Naïve participants in the PY group were more accurate in accepting the matches than those in the CH group. Moreover, participants in the PY group were overall insensitive to tonal mismatches regardless of their proficiency. In the CH group, however, the Advanced learners scored significantly higher with the tonal-mismatched items than the less experienced participants. This study suggests that characters are more effective than Pinyin in helping Advanced L2 learners encode the tones of new Mandarin words.

**Index Terms:** L2 orthography, phonological encoding, segments and tones, different proficiency levels.

### 1. Introduction

One of the most crucial tasks in second language (L2) acquisition is to build a lexicon. This process requires forming association between the sound, meaning, and very often the orthographic form of L2 words. L2 orthography is not only essential in L2 reading and writing, it has been found to affect learners' encoding of the phonological forms of L2 words [1-9]. Some studies suggest that the availability of the written form aids learners' encoding of L2 sound and meaning. For example, Dutch speakers who were provided with the written forms of English non-words containing the vowels /ε/ and /æ/ during training could distinguish these two sounds to some extent in the post-training test, while those who were not provided with the written forms could not discriminate this vowel contrast at all [3]. In contrast, other studies have provided evidence for a negative effect of L2 orthographic form on L2 phonological encoding, especially when the L2 and L1 share the same orthographic symbols but employ different grapheme-to-phoneme correspondence. For example, English speakers were found to use their L1 correspondence rules to interpret Pinyin, the Chinese Romanization system. Hence they often failed to pronounce or perceive the vowel

that was not represented in Pinyin, even when it was present in the audio stimulus [1, 2]. In sum, some studies found that the availability of L2 orthographic form may help learners encode the sounds of novel words. However, other studies showed that the divergent grapheme-to-phoneme correspondence between the L1 and L2 orthographic systems may negatively interfere with the learning of L2 phonological forms.

A few studies have examined the effect of novel L2 orthographic symbols on the learning of L2 sounds. One study showed that the group of English speakers who were presented with Mandarin tonal diacritics could discriminate Mandarin tonal contrasts more accurately than the group that did not see the diacritics, suggesting that novel orthographic symbols can help English speakers establish new categories for Mandarin tones [7]. In another study conducted by the same authors [8], English speakers who saw the Arabic orthographic symbols did not discriminate the Arabic velar-uvular contrast more accurately than those who received no visual input. A third group who were presented with the Romanization of the target Arabic contrast performed worse than the other two groups. The authors concluded that the novel script failed to provide facilitation probably because it is relatively complicated and the target contrast is difficult as well. As for the Romanization, it had a negative effect on their learning because the alphabets could not adequately reflect the sound difference of this Arabic contrast. In yet another study that compared the effect of familiar and novel orthographic symbols on L2 sound learning, three groups of English speakers who received different visual inputs during a brief Mandarin sound instruction were compared in their ability to distinguish English and Mandarin consonants after the instruction [5]. One group was instructed in Pinyin, one group in Zhuyin (Mandarin phonetic symbols utilizing character-like symbols), and one group without any visual input. It was found that the Pinyin group performed better than the other two groups, suggesting that familiar orthographic symbols may be more beneficial for the learners at the beginning stage, because they are spared the task of learning new symbols and can focus on the learning of sounds. In sum, previous studies have not reached a consensus regarding the relative efficacy of familiar and novel orthographic symbols in L2 phonological encoding.

Inspired by these previous findings, the present study will compare the effect of Pinyin and Chinese characters on English speakers' phonological encoding of Mandarin words. These two orthographic representations are of interest because they differ in the following aspects. Pinyin utilizes Roman alphabets to represent the segments of each syllable and diacritics above the segments to mark the tone (e.g. mā for T1: high-level tone; má for T2: rising tone; mǎ for T3: low dipping tone; mà for T4: falling tone). Characters, on the other hand, are not sound-marking symbols. Each character corresponds to

a syllable but cannot be decomposed into individual segments and tones. Pinyin is widely used in Mandarin instruction for English-speaking beginning learners due to its shared symbols with English orthography. As the learners' proficiency increases, however, the use of characters in language instruction becomes more prevalent. Considering the potentially different effect of Pinyin and characters for English speakers at different learning stages as a result of the instructional method, we compare three L2 proficiency groups in our study. Furthermore, we will examine the effect of Pinyin and characters on L2 segment and tone encoding because individual segments and tones are marked in Pinyin but not in characters. Specifically, our study aims to address the following questions: 1) Is the character or Pinyin input more helpful for English speakers to encode the sounds of new Mandarin words? 2) Do these two orthographies differ in their facilitation of segment or tone encoding? 3) Does the effect of these two orthographies on Mandarin phonological encoding differ for English speakers at different proficiency levels?

## 2. Methods

### 2.1. Participants

Three groups of native English speakers differing in their Mandarin experience participated in the current study. The Naïve group was composed of 20 college students who were naïve to Mandarin (18F, mean age=21.25). The Intermediate (Inter) group consisted of 29 learners recruited from 2<sup>nd</sup>- and 3<sup>rd</sup>-year Chinese classes (16F, mean age=19.62). The Advanced (Adv) group consisted of 17 learners recruited from 4<sup>th</sup>-year Chinese classes (6F, mean age=20.56). The participants at each level were randomly assigned to the Pinyin input (PY) group and Character input (CH) group (Naïve: 10 vs. 10; Inter: 16 vs. 13; Adv: 9 vs. 8 in the PY and CH groups, respectively.)

### 2.2. Target words

Sixteen disyllabic Mandarin words were selected from the vocabulary of the highest HSK (a standardized Chinese proficiency test) levels (i.e. level 5 and 6), which include low frequency words that are more likely to be unknown to L2 learners. On the other hand, the individual characters used to form the target words are introduced in the textbooks of Elementary Chinese (*Integrated Chinese* Level 1). The purpose of this selection criterion is to ensure that the characters are known to the learner groups.

After selecting the target words, variants differing from the target words in segments (segmental mismatch) or in tones (tonal mismatch) were created for each target word. Half of the variants were made in the first syllable while the other half in the second syllable. The segmental mismatches included both consonantal and vowel divergence. The divergence was created by replacing the target segment non-existent in English with a segment that exists in English, e.g. replacing [uo] with [oo]; or by replacing the target segment that does not conform to the English grapheme-to-phoneme correspondence with one that is closer to the English spelling conventions, e.g. replacing [eiɑŋ], which is spelled as "xiang" in Pinyin, with [ʃɑŋ], which is spelled as "shang". The tonal mismatches focused on the contrast between T1 and T4 and that between T2 and T3. Table 1 provides some examples that illustrate the design of segmental and tonal-mismatched items.

Table 1: *The English meaning (Eng), characters (CH), Pinyin (PY) and IPA of two target words and their segmental and tonal mismatches. The mismatches are highlighted in bold.*

	Target words	Segmental mismatch	Tonal mismatch
Eng	'theory'		
CH	学说		
PY	xuéshuō	xuéshōu	xuéshuò
IPA	[ɣɛT2 ʃuoT1]	[ɣɛT2 <b>ʃoo</b> T1]	[ɣɛT2 ʃuo <b>T4</b> ]
Eng	'day dream'		
CH	空想		
PY	kōngxiǎng	kōngshǎng	kōngxiàng
IPA	[koŋT1 eiɑŋT3]	[koŋT1 <b>ʃɑŋ</b> T3]	[koŋT1 eiɑŋ <b>T2</b> ]

One female native Mandarin speaker from northern China produced the 16 target words, 16 segmental mismatched variants, and 16 tonal mismatched variants. These recordings were used as the stimuli of this experiment.

### 2.3. Procedure

This experiment consists of a target word assessment, word-learning phase, a criterion test, and a perception test. All the tasks except for the target word assessment were programmed and administered in *E-Prime* 2.0.

#### 2.3.1. Target word assessment

The participants listened to the recording of the 16 target words and tried to write down their English meaning. The purpose of this task was to check whether all the targets are new to the participants.

#### 2.3.2. Word-learning phase

The participants listened to the audio recording of each target word through headphones and saw its English meaning on the screen. They were instructed to repeat after each audio stimulus. The PY group saw Pinyin with tone marks below the English meaning, while the CH group saw characters. Each target word appeared four times at a random order.

#### 2.3.3. Criterion test

The purpose of the criterion test was to ensure that the participants have memorized the association between sound and meaning of the target words. The participants were instructed to judge the matchedness of 32 pairs of audio stimulus and English meaning. No Pinyin or character input was provided. In half of the trials the audio stimulus matched the meaning on the screen. In the other half the audio stimulus corresponded to a target word different from the one shown on the screen. If the participants failed to achieve 90% or higher accuracy in the criterion test, they had to repeat the learning phase until they reached the threshold.

#### 2.3.4. Perception test

In the perception test, the participants were instructed to judge the matchedness of 128 pairs of audio stimuli and English meanings. Different from the criterion test, in the mismatched trials the audio stimulus differed from the target word only by one segment or tone, rather than corresponded to a different word. 64 pairs were complete matches (16 target words × 4 repetitions), 32 pairs were segmental mismatches (16

segmental mismatches  $\times$  2 repetitions), and 32 pairs were tonal mismatches (16 tonal mismatches  $\times$  2 repetitions).

### 2.3.5. Analysis

Individual participants' accuracy rates in the Perception test were compared via a Repeated Measures ANOVA across Proficiency levels and Orthographic input (between-subject factors) and the Matching condition (within-subject factor). The target words correctly identified by each participant in the pre-learning assessment were removed from his/her data, so that all the target words analyzed were novel to the learners prior to the experiment. This way we can ensure that the learners' accuracy in the perception test reflects their learning outcome rather than their pre-knowledge of the target words. Furthermore, given the binomial nature of the accuracy data, the accuracy rates were first transformed to the Rationalized Arcsine Unit before running the ANOVA [10].

## 3. Results

The mean accuracy rates of the three proficiency groups who saw Pinyin during the learning phase are plotted in Figure 1, while those who saw characters are plotted in Figure 2. The accuracy rates of the Naïve group are represented by the bars with slanted lines; those of the Inter group by the grey bars; and those of the Adv group by the black bars. From left to right are the accuracy rates in the matched, segmental-mismatched, and tonal-mismatched conditions. The error bars denote standard errors.

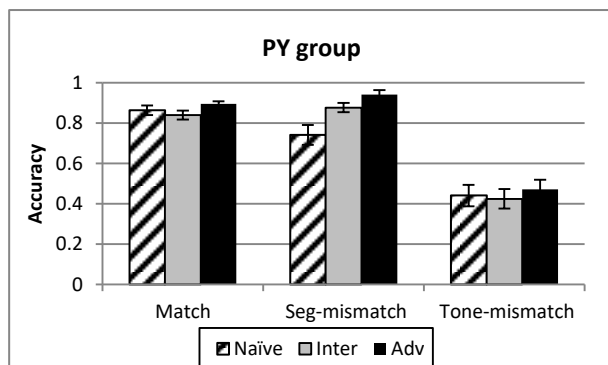


Figure 1: Accuracy rates of the three proficiency groups who received Pinyin input during the learning phase.

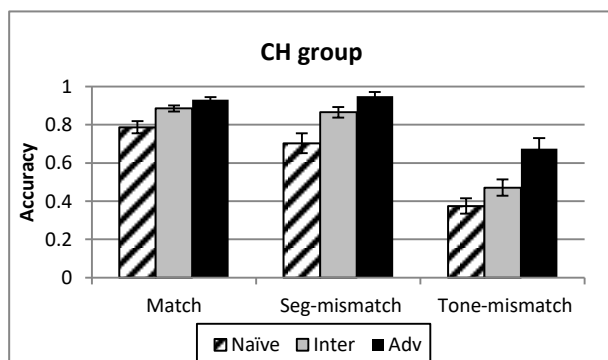


Figure 2: Accuracy rates of the three proficiency groups who received Character input during the learning phase.

As can be seen from Figure 1 and 2, one obvious similarity between the PY and CH groups is that the participants were generally least accurate in the tonal-mismatched condition. One noticeable dissimilarity between the PY and CH groups, on the other hand, is the greater accuracy difference in the tonal-mismatched condition between the participants from the three proficiency levels in the CH group than in the PY group.

The results of the Repeated Measures ANOVA showed that the participants' proficiency significantly affected their accuracy ( $F(2,60)=22.23$ ,  $p<.001$ ,  $partial \eta^2=.43$ ). Post-hoc analysis (Bonferroni adjustment) indicated that the Adv group achieved significantly higher accuracy than the Inter group, who in turn outperformed the Naïve group ( $ps<.005$ ). The Matching condition was also found to be significant ( $F(2,120)=225.09$ ,  $p<.001$ ,  $partial \eta^2=.79$ ). The source of the difference was identified to be the lower accuracy of tonal-mismatched trials than both the matched and segmental-mismatched trials ( $ps<.001$ ). Orthographic input, however, was not found to be a significant factor that impacted the accuracy ( $F(1,60)=.89$ ,  $NS$ ).

In addition to the main effects, the interaction between Proficiency and Orthographic input ( $F(2,60)=3.85$ ,  $p=.027$ ,  $partial \eta^2=.11$ ) was found to be significant. Post-hoc analysis (Bonferroni adjustment) indicated that the CH group achieved significantly higher accuracy than the PY group at the Advanced level ( $p=.03$ ). Another significant interaction was found between Proficiency and Matching condition ( $F(4,120)=3.41$ ,  $p=.011$ ,  $partial \eta^2=.10$ ). Post-hoc tests revealed that in the matched condition, the only difference is that the Adv group was significantly more accurate than the Naïve group ( $p=.004$ ). In the segmental-mismatched condition, the Adv group performed significantly more accurately than the Inter group, who outperformed the Naïve group ( $ps\leq.009$ ). Finally in the tonal-mismatched condition, the Adv group was more accurate than both the Inter and Naïve groups ( $ps\leq.037$ ), while there is no difference between the latter two groups.

Although the three-way interaction was not found to be significant ( $F(4,120)=1.05$ ,  $p=.38$ ,  $partial \eta^2=.03$ ), the difference in accuracy between the three proficiency levels seems to differ in the PY and CH groups, based on the visual comparison of Figure 1 and 2. To verify this observation, post-hoc tests (Bonferroni adjustment) were performed on the three-way interaction. It was found that within the PY group, the only significant difference found is that the Adv learners performed significantly better than the Naïve participants in the segmental-mismatched condition ( $p<.001$ ). As for the CH group, the Adv learners were significantly more accurate than the Naïve participants in the matched condition ( $p=.001$ ). In the segmental-mismatched condition, the Adv learners were more accurate than the Inter learners, who were more accurate than the Naïve participants ( $ps\leq.048$ ). In the tonal-mismatched condition, the Adv learners were more accurate than both the Inter and Naïve groups ( $ps\leq.021$ ), while there is no difference between the latter two. Across the two input groups, in the matched condition, the Naïve participants in the PY group were more accurate than those in the CH group ( $p=.037$ ). In the tonal-mismatched condition, in contrast, the Adv learners in the CH group achieved higher accuracy than those in the PY group ( $p=.013$ ).

## 4. Discussion

The main purpose of this study was to assess the effect of two types of orthographic inputs, Pinyin and Characters, on English speakers' Mandarin sound encoding. Our data suggest that the effect varies depending on the participants' Mandarin proficiency level and the type of sound contrast. Specifically, at the advanced level the CH group was more accurate than the PY group in rejecting tonal mismatches, while the Naïve participants who received the PY input during the learning phase were more accurate in accepting the matched stimuli than those who received the CH input. The finding on the Naïve participants is compatible with that in [5], who found that English speakers with no prior knowledge of Mandarin could discriminate Mandarin and English sounds better when the instruction was conducted in Pinyin than in Zhuyin. It can be inferred from both of these studies that the L2 orthography with common symbols as the L1 may be more beneficial for learners at the beginning stage, because they are not distracted from the sound encoding by the additional task of learning new symbols. However, the current data also showed that for the more advanced learners, the novel orthographic symbols may have a greater facilitatory effect than the familiar symbols. There are several possible explanations. First, as suggested in [8], the familiar orthographic symbols may not adequately represent sound contrasts that do not exist in the learners' L1. Hence novel L2 symbols may become more helpful as they offer a more nuanced representation of the L2 sounds. Second, the individual characters of the target words in this study were carefully chosen so that they should all be known to the advanced learners. As a result, the learners in the CH group were not tasked with additional learning of new symbols and thus could focus on sound encoding just like the PY group. Third, as the learners' Mandarin proficiency increases, the use of characters in language instruction becomes more prevalent than that of Pinyin. Thus the CH group at the Adv level achieved significantly higher accuracy than the PY group probably because the presentation of characters along with English meaning of new words is closer to their accustomed way of vocabulary learning.

One intriguing observation is that the Adv CH group performed better than the PY group particularly in the tonal-mismatched trials, suggesting that the character input facilitated their tone encoding more than the Pinyin input. However, as described in the Introduction, the tone of a syllable is marked in Pinyin but not in characters. Then why is the character input more helpful than Pinyin for the Adv learners' tone encoding? We argue that this may be because tones are marked in Pinyin by diacritics, which are relatively unfamiliar symbols to English speakers compared to the Roman alphabets used to mark segments. Hence it is likely that the learners in the PY group attended more to segments than to tones because the former are marked by more familiar symbols. As for characters, there is no direct one-to-one grapheme-to-phoneme correspondence. Nonetheless, each character corresponds to one syllable, including both segments and tone. As a result, the character input probably fosters a more holistic representation of sounds encompassing the segmental and tonal dimensions. A similar hypothesis has been proposed by [11], who found that Mandarin speakers recalled words much faster than Vietnamese speakers when both being prompted by their L1 orthography. The author suggested that this is because Mandarin speakers have developed an integrated entry for each character with its pronunciation (segments & tones) and meaning, while

Vietnamese speakers may have separate entries for segments and tones, since these are marked separately in their orthography. Our findings showed that the Adv L2 learners of Mandarin could benefit from the integral representation of sounds in characters, similar to the native speakers in [11]. As for the participants with lower Mandarin proficiency, the tone encoding accuracy rates of the CH and PY groups were equally low. This suggests that it may require extensive experience with characters for the English speakers to become more balanced in segment and tone encoding. Moreover, the PY input seems to have a hindering effect on tone encoding regardless of the participants' Mandarin experience, which can be inferred from the below-chance accuracy of tonal-mismatched items by all participants in the PY group.

One tendency common to all participant groups is their higher accuracy in detecting segmental than tonal mismatches to the target words. Even for the Adv learners who received the CH input, their accuracy in the tonal-mismatched trials (.675) is still much lower than that in the segmental-mismatched trials (.95). A similar finding has been reported in previous research [12, 13], which showed that English speakers were more accurate in perceiving Mandarin vowels than tones. These authors attributed this tendency to their L1 influence, because vowels but not tones are contrastive in English. Similarly in the current study, the English-speaking participants appeared to attend more to segmental than to tonal information probably because they are accustomed to encoding segmental but not tonal information in their L1. As a result, they are far better at detecting segmental mismatches than tonal mismatches.

Finally, this study overall did not find a substantial difference in the performance between the PY and CH groups. Except for the tonal-mismatched condition at the Adv level and the matched condition at the Naïve level, these two groups patterned very similarly. This suggests that L2 orthography may not play a crucial role in English speakers' learning of Mandarin words. More generally speaking, it could be the case that orthographic input is not particularly helpful for English speakers in their word learning. Such a claim has found support in former studies which showed that English speakers who were provided with orthographic representation and audio input of the target words did not memorize the phonological forms better than those who were only provided with audio input [4, 14]. This could be due to the fact that the English orthography does not contain reliable grapheme-to-phoneme correspondence, and hence English speakers may rely more heavily on the audio input than on orthography in their L1 and L2 word learning.

## 5. Conclusions

This study demonstrated that the effect of L2 orthography on the phonological encoding of new words may depend on a number of factors. The L2 orthography adopting symbols similar to the L1 may be beneficial for beginning learners. However, for the advanced learners, novel orthographic symbols may become more beneficial. In our study, the benefit of the character input has been found specifically in tone encoding. We attributed this to the unique nature of the Chinese writing system, which promotes an integrated representation of segmental and tonal information. The current findings advocate for the use of characters rather than Pinyin in teaching L2 learners of Chinese, except for the beginning stage, because Pinyin seems to hinder students from attending to the tonal information of Mandarin words.

## Acknowledgements

This study was funded by the University of Tennessee Professional Development Award given to the first author.

## References

- [1] B. Bassetti, "Orthographic input and phonological representations in learners of Chinese as a foreign language." *Written Language & Literacy*, vol. 9, no. 1, pp. 95-114, 2006.
- [2] B. Bassetti, "Effects of hanyu pinyin on pronunciation in learners of Chinese as a foreign language". In A. Guder, X. Jiang, & Y. Wan (Eds.), *The cognition, learning and teaching of Chinese characters* (pp. 155-179). Beijing: Beijing Language and Culture University Press, 2007.
- [3] P. Escudero, R. Hayes-Harb, and H. Mitterer, "Novel second-language words and asymmetric lexical access." *Journal of Phonetics*, vol. 36, no. 2, pp. 345-360, 2008.
- [4] R. Hayes-Harb, J. Nicol, and J. Barker, "Learning the phonological forms of new words: Effects of orthographic and auditory input." *Language and Speech*, vol. 53, no. 3, pp. 367-381, 2010.
- [5] C. Pytlyk, "Shared orthography: Do shared written symbols influence the perception of L2 sounds?." *The Modern Language Journal*, vol. 95, no. 4, pp. 541-557, 2011.
- [6] E. Simon, D. Chambless, and U.K. Alves, "Understanding the role of orthography in the acquisition of a non-native vowel contrast." *Language Sciences*, vol. 32, no. 3, pp. 380-394, 2010.
- [7] C. Showalter and R. Hayes-Harb. "Unfamiliar orthographic information and second language word learning: A novel lexicon study." *Second Language Research*, vol. 29, no. 2, pp. 185-200, 2013.
- [8] C. Showalter and R. Hayes-Harb, "Native English speakers learning Arabic: The influence of novel orthographic information on second language phonological acquisition." *Applied Psycholinguistics*, vol. 36, no. 1, pp. 23-42, 2015.
- [9] C.-L. Yang, *Phonological variation and L2 word learning: the role of orthography in word recognition and production*. Unpublished PhD dissertation. Indiana University Bloomington, 2015.
- [10] G.A. Studebaker, "A rationalized arcsine transform." *Journal of Speech, Language, and Hearing Research*, vol. 28, no. 3, pp. 455-462, 1985.
- [11] N. Liu, "Implicit Priming Effects in Chinese Word Recall: The Role of Orthography and Tones in the Mental Lexicon." *Computational Linguistics & Chinese Language Processing*, vol. 18, no. 3, pp. 1-20, 2013.
- [12] T.L. Gottfried and T.L. Suiter, "Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones." *Journal of Phonetics*, vol. 25, no. 2, pp. 207-231, 1997.
- [13] Y.-C. Hao, "Second language perception of Mandarin vowels and tones." *Language and Speech*, vol. 61, no. 1, pp. 135-152, 2018.
- [14] P. Escudero, "Orthography plays a limited role when learning the phonological forms of new words: The case of Spanish and English learners of novel Dutch words." *Applied Psycholinguistics*, vol. 36, no. 1, pp. 7-22, 2015.