

RESEARCH

Open Access



Language-dependent cue weighting in distinctive feature: evidence from the perception of Mandarin high vowels by native English speakers

Wenhui Zhu¹, Sun-Hee Lee^{2*} and Xinting Zhang¹

*Correspondence:
lishanxi@cufs.ac.kr

¹ School of Humanities
and Social Sciences, Southern
University of Science
and Technology, Shenzhen,
China

² Department of Chinese, Cyber
Hankuk University of Foreign
Studies, Cyber BD Rm#716, 107
Imun-ro, Dongdaemun-gu,
Seoul 02450, Republic of Korea

Abstract

This study investigates the perception of the three Mandarin high vowels /i, u, y/ after dental, retroflex, and palatal fricatives and affricates (/s/-/ʃ/-/tʃ/; /ʂ/-/ʈʂ/-/tʂʰ/, and /tʂ/, /tʂʰ/, /ʈʂ/) by native English speakers. The results of the perceptual identification and categorization experiments show that among the three target vowels, the high front rounded vowel /y/ presents the greatest challenge for native English speakers. They have a significantly higher tendency to confuse /y/ with the Mandarin high-back rounded vowel /u/ compared to the Mandarin high-front unrounded vowel /i/, as they perceptually classified /y/ and /u/ into the same English vowel category /u/. The findings of the study suggest that native English speakers adopt a perceptual strategy that differs from that of native Japanese and Korean speakers, relying heavier on the feature of roundness as opposed to backness in perceiving the Mandarin /y/. This study contributes to the perceptual cue weighting field by examining the weighting of phonetic cues (i.e., distinctive features) in Mandarin high vowels by native English speakers. These results hold pedagogical significance as they highlight the importance of targeted perception training for learners of different language backgrounds to enhance both their recognition and reproduction of second language sounds.

Keywords: Perceptual cue weighting, High vowel, Distinctive feature, Mandarin, English, Japanese, Korean

Introduction

Perceptual cues, broadly speaking, can be any information that has consistent effects on how listeners perceive a certain phonological contrast (Schertz & Clare, 2020). Perceptual cue weighting is a fundamental component of language acquisition (Holt & Lotto, 2006), as it entails the determination of the relative significance of various phonetic cues in the perception of speech sounds.

Perceptual cue weighting in SLA

The results of multiple studies have indicated that the relative significance assigned to phonetic cues may vary between native and non-native listeners (Bohn & Flege, 1990; Casillas, 2015; Escudero & Boersma, 2004; Escudero et al., 2009; Flege et al., 1997; Guo & Chen, 2017; Shultz et al., 2012; Wang & Munro, 1999; Zhou, 2007). For instance, Flege et al. (1997) have found that Mandarin-speaking learners of English tend to place greater weight on duration cues than on spectral cues (such as formant frequencies), whereas native English-speaking listeners place more weight on spectral cues. Escudero et al. (2009)'s study also demonstrated that native Spanish speakers learning Dutch prioritized vowel duration over vowel spectrum, in contrast to native Dutch and German listeners who prioritized vowel spectrum. This suggests that cross-language differences can play a role in the way learners weigh different phonetic cues. Two influential second language (L2) speech perception models: the Perceptual Assimilation Model for L2 learners (PAM-L2: Best et al., 2007) and the Speech Learning Model (SLM: Flege, 1995) both proposed that the difficulty of L2 sounds is largely determined by their relationship to first language (L1) sound categories. For instance, SLM proposed that L2 sounds which do not exist in the learner's L1 will be more challenging for L2 learners to perceive, particularly during the beginning stages of acquisition. PAM-L2 proposed that two L2 phones that have been classified into the same L1 sound category will incur the most perceptual confusion for L2 learners. Studies have shown that the perception of speech sounds in an L2 depends on the learner's native language. For example, English-speaking learners of Spanish may have difficulty with the Spanish phonemes that do not exist in English (Flege, 1995). Kondaurova and Francis (2008) proposed that the duration distinction existed in the vowel allophonic level of Spanish and Russian, and L1 Spanish and Russian learners transferred this use of duration to the learning of L2 vowels. This could be seen in the tendency to rely more heavily on the vowel temporal/duration cue than the spectral cue. Data from Escudero and Boersma (2004)'s research further supported the role of L1 in determining the importance of cues during L2 speech perception. These suggest that the phonetic cues that are perceived as important by the learner depend on the native language.

It has been observed that the transfer from L1 can account for a significant portion of perceptual cue weighting. However, some research findings indicate that not all perceptual cue weighting can be attributed to the influence of L1. For instance, as demonstrated in the study by Flege et al. (1997), Mandarin-English learners placed a higher value on vowel duration even though Mandarin vowels do not differ in duration. Furthermore, despite the findings of Wang and Munro (1999) that L1 Mandarin learners showed a preference for vowel spectrum over vowel temporal/duration cues in their perception of the English /u/-/ʊ/ contrast, which differs in both tenseness and duration, subsequent studies, including the same one, indicated that listeners tend to prefer vowel duration over the spectrum in other vowel contrasts that vary in both duration and spectrum, such as the English /i/-/ɪ/, /e/-/æ/ and Dutch /a:-/a/ contrasts (Bohn & Flege, 1990; Escudero et al., 2009; Escudero & Boersma, 2004; Flege et al., 1997; Shultz et al., 2012; Wang & Munro, 1999). Given that L2 learners who exhibit a greater reliance on the duration over spectrum do not possess the temporal/duration feature in their L1, it is suggested that this weighting cannot be attributed to L1 transfer, but instead is regarded

as “a general speech perception strategy” (Bohn & Flege, 1990, p. 326) or duration is a more salient cue in vowel perception compared to spectrum (Bohn, 1995; Escudero et al., 2009).

Further inquiry into the phenomenon of perceptual cue weighting in L2 acquisition is necessary, as the current understanding of this topic remains inconclusive and has primarily been based on research with European languages (such as English, French, and German). This study seeks to contribute to this field by examining the perceptual weighting of phonetic cues in Mandarin high vowels as perceived by native English speakers through a series of perceptual experiments.

Mandarin high vowels

There are five or six¹ contrastive vowel phonemes in Mandarin: one low vowel phoneme (/a/), one mid vowel phoneme (/ɤ/), and three high vowel phonemes (/i/, /y/, /u/) (Duanmu, 2007; Huang & Liao, 1983; Lin, 2007). The selection of target vowels is primarily based on their considerable level of difficulty for L2 learners, particularly after the dental, retroflex, and palatal fricatives and affricates (/s/-/ʃ/-/tʃ^h/; /ʂ/-/tʂ/-/tʂ^h/; and /tɕ/, /tɕ^h/, /ç/) in Mandarin. This difficulty has been noted in previous studies by Lu (1984), Zhu and Wang (1997), Wang (2001), Wang and Deng (2009), and Yao (2017). Furthermore, research on the Mandarin high vowels after the dental, retroflex, and palatal fricatives and affricates is relatively limited. As for English, it is generally agreed that it has 11² nonrhotic distinctive monophthongs /i, ɪ, e, ɛ, æ, ʌ, ɑ, ɔ, o, ʊ, u/ (Bauer et al., 2007; Hao, 2018). In comparing the vowel systems of Mandarin and English, it is found that the only Mandarin vowel that has no close equivalent in English is the front rounded vowel /y/. According to SLM, /y/ will incur the greatest difficulty for L1 English learners.

Empirical research indicates that native English speakers encounter challenges in distinguishing between the high vowels /y/ and /u/ in Mandarin. Specifically, the syllable pairs zhu [tʂu] and ju [tɕy], chu [tʂ^hu] and qu [tɕ^hy], and shu [ʂu] and xu [ɕy] are particularly challenging for English speakers to differentiate. Moreover, they tend to pronounce both Mandarin vowels /u/ and /y/ as the English vowel /u/ (Bi, 2001; Lu, 1984; Ni & Wang, 1992; Zhu & Wang, 1997). Similar research findings were obtained from experimental studies. For instance, Wang (2001) explored the perception of Mandarin high vowels by L1 Japanese and Korean learners with elementary proficiency in Mandarin. The study revealed that the occurrence of the vowel /y/ after certain consonants (j/tɕ/, q/tɕ^h/, x/ç/) caused a significantly higher error rate, compared to /y/ after consonants l/l/ and n/n/ (32.5% vs. 13.1% for Korean learners; 34.6% vs. 13.1% for Japanese learners); the predominant pattern of confusion was identified to exist between the vowels /y/ and /i/, rather than between the vowels /y/ and /u/ or between /i/ and /u/. Li and Liu (2008) investigated the acquisition of Mandarin vowel categories by L1 American English learners with elementary Mandarin proficiency. They found that the /y/ category was the last to be developed.

¹ There may be also a retroflex vowel phoneme /ɤ/ in Mandarin, but it has limited distribution and lacks of a clear phonetic description (Huang & Liao, 1983; Lin, 2007). It is beyond the scope of the present study.

² Non-rhotic standard English varieties, such as RP, Australian English, and New Zealand English have 11 monophthongs (Bauer et al., 2007). American English has 9 monophthongs (Ladefoged & Johnson, 2014).

Hao (2018) explored the ability of L1 English speakers to discriminate between the Mandarin vowel contrasts /li-ly/ and /lu-ly/, but did not delve into the onsets of dental, retroflex, and palatal fricatives and affricates that caused the most confusion. Although there is limited experimental evidence on English speakers' perception of Mandarin high vowels, previous research on high front rounded vowels in other languages suggests that /i-y/ and /u-y/ contrasts are expected to be more difficult to perceive for English speakers. The former pair is acoustically similar (Strange et al., 2004, 2007), and lip-rounding is not a primary feature for distinguishing vowels in English (Bauer et al., 2007; Hao, 2018). The latter pair /u-y/ is easily confused by English speakers who are not familiar with French or German (Levy, 2009; Levy & Strange, 2008; Strange et al., 2004, 2007). In view of these findings, the current study selects the three high Mandarin vowels following the dental, retroflex, and palatal fricatives and affricates as the target stimuli to investigate the weighting of phonetic cues in the perception of Mandarin high vowels. The objective of the investigation is to provide precise responses to inquiries concerning the topic at hand. These inquiries include determining the manner in which individuals who are native English speakers perceive high vowels in Mandarin, identifying the particular group of Mandarin high vowels that may result in the most confusion, ascertaining whether perceptual cue weighting is a factor in such perception, and specifying the types of cue weighting that may be involved as well as the underlying rationales.

To gain insight into how native English speakers prioritize differences in non-native speech sounds and investigate the factors that impact how listeners weigh various speech cues, the present study examines the weight given to different cues when perceiving high vowels in Mandarin. A crucial aim of the study is to understand how cue weighting influences the acquisition of speech categories in L2 learners. Specifically, the study focuses on how English speakers without previous knowledge of Mandarin perceive the /y/ vowel category. This approach allows for strict control over participants' linguistic backgrounds and enables a precise determination of the distributional properties linking cue informativeness to categorization responses.

Experiments

Experiment 1: identification task

Participants

A group of native English speakers participated in this study. The group consisted of 19 university freshmen who were beginning Mandarin learners (12 males and 7 females). Their age was 21.5 on average. The Mandarin learning period was 178.7 h on average. None of these students were heritage speakers of Mandarin. All of them reported having normal hearing.

Stimuli

Previous research has shown that the combination of Mandarin high vowels with dental, retroflex, and palatal fricatives and affricates pose significant difficulties. To address this, the study presented the target Mandarin high vowels /i, u, y/ in monosyllabic Chinese characters that began with three different sets of consonants (the dental, retroflex, and palatal fricatives and affricates /s/-/ʃ/-/ʃʰ/; /ʂ/-/ʈʂ/-/ʈʂʰ/, and /tʃ/-/tʃʰ/-/tʃʰ/) along with

the fourth Mandarin tone,³ resulting in a total of nine⁴ syllables (ju/ɿy/-qu/ɿy/-xu/ɿy/-ji/ɿi/-qi/ɿi/-xi/ɿi/-zhu/ʈʂu/-shu/ʂu/-tʂu/ʈʂu/). This was done to standardize the impact of onset consonants on both English and Chinese target vowels in a subsequent perceptual mapping task as well (as described in section "Stimuli").

On a reading sheet, these syllables were printed in Chinese characters. Two native Mandarin speakers (a male and a female) were recruited and recorded individually in an audio laboratory at a sampling rate of 44.1 kHz using the Praat software version 5.1.05 (Boersma & Weenink, 2009). The speakers were instructed to speak at their natural tempo and to self-correct any errors or incoherent sentences. Only isolated Mandarin monosyllables were chosen using Praat after recording. Before recording, the author ensured that each syllable was correctly pronounced by all participants and provided instructions for them to read each of the three Mandarin syllables in a carrier sentence: [wo ʈʂʰ uo ____] ("I say ____"). Each talker read the list once. The quality of the native talkers' productions was evaluated by two additional native Mandarin speakers to ensure they were good exemplars of the target stimuli. The syllables with the highest quality ratings were selected and used in a task where participants had to identify them. The study utilized a total of 9 different vowel syllables, with 3 syllables for each of the 3 different vowel types.

Procedure

The forced identification task was conducted using Praat software. Participants listened to all the stimuli in comfortable auditory conditions and were instructed to select one of the three Mandarin high vowels - /i/, /u/, or /y/ - by writing down the corresponding digit that represented the Chinese characters containing the target vowels they heard to avoid Pinyin spelling bias. They were advised to guess if they were unsure and to take as much time as necessary to make a decision. Each participant heard all nine vowel syllables three times, and the 27 stimuli were randomly presented in a different order to each participant. The accuracy of the participants' perception of the target was assessed by comparing their responses to the originally recorded syllables. A total of 513 responses were obtained from the participants, which corresponded to 19 (participants) × 27 (target syllables) responses. The target syllables comprised nine syllables for each of the three different vowel types. The collective errors for each vowel type from all participants were aggregated to determine the mean error percentage. This involved summing up the errors made by each listener for each specific vowel type and then dividing the total by the respective number of responses for that vowel type. To illustrate, if the 19 English learners collectively made two errors in the perception of the vowel /i/, the mean error percentage for that vowel among them would be calculated as 2 divided by 171, resulting in 1.2%.

³ Mandarin has four patterns of pitch changes (tones) to distinguish word meanings (Lin, 2007). Tone 1 is high level; tone 2 is high rising; tone 3 is low falling-rising and tone 4 is high falling. Since the change of Mandarin tones can affect the meaning of Mandarin words, it is necessary to standardize the tone of target syllables.

⁴ The target vowels have limited distribution after chosen consonants. For instance, /y/ cannot appear after consonants zh/ʈʂ/-sh/ʂ/-tʂ/ʈʂ/ and /u/ cannot appear after consonants j/ɿ/-q/ɿ/-x/ɿ/. To ensure a balanced number of stimuli for each target vowel, we selected three consonants for each type of target vowel, resulting in the formation of 9 syllables instead of 27.

Table 1 Identification error rates for target vowels by L1 English learners

Mandarin vowel	English learner (n = 19)	
	Total error number	Mean error rate (SD; range)
y	61	35.7% (0.24; 0–78%)
u	28	16.4% (0.16; 0–56%)
i	0	0%

Table 2 Confusion matrix between target vowels by 19 participants

Stimuli	i	y	u
i	–	0 (0.0%)	0 (0.0%)
y	5 (2.9%)	–	56 (32.7%)
u	0 (0.0%)	28 (16.4%)	–

Results

The average identification error rates for the target three Mandarin high vowels across all participants were as follows: 0% for /i/ (no errors⁵), 16.4% for /u/ (28 errors out of a total of 171/u/ stimuli), and 35.7% for /y/ (61 errors out of a total of 171 /y/ stimuli). Among the three target Mandarin high vowels, the high front rounded vowel /y/ is the hardest for learners to identify correctly. Please see Table 1 for details.

To better reveal the confusion among the target three Mandarin vowels, the confusion patterns among the target vowels across all participants are presented in Table 2. The results showed 0% confusion for /i-u/ (no errors), 2.9% confusion for /i-y/ (5 errors out of a total of 171 /i-y/ contrasts), and 49.1% confusion for /u-y/ (84 errors out of a total of 171 /u-y/ contrasts). It was observed that the Mandarin vowel /y/ was more frequently confused with the Mandarin vowel /u/ than with the vowel /i/ across all contrast errors. Out of the 171 /y/ stimuli, /y/ was mistakenly identified as /u/ 56 times by all participants (mean error rate: 32.7%), and as /i/ only 5 times (mean error rate: 2.9%). The results of paired-sample t-tests showed a significant difference between the percentages of the two types of /y/ errors (/u/ type error and /i/ type error) ($t = -2.249$, $df = 18$, $p = 0.037$). Therefore, participants displayed cue weighting when identifying the Mandarin vowel /y/. The other two target Mandarin vowels showed no cue weighting in the identification task, as vowel /i/ had no error at all, and vowel /u/ was only mistaken as vowel /y/.

Figure 1 presented an analysis of individual differences in the patterns of errors made by participants in identifying Mandarin high vowels. The data clearly indicate that the majority of learners tended to mistake /y/ for /u/ (blue columns). However, one participant (PLe) made more errors in mistaking the Mandarin /y/ for /i/ (red columns) than for /u/ (as the participant only made a total of 3 mistakes).

⁵ The absence of errors observed here indicates that the errors made by L2 learners were not caused by difference in the preceding consonants.

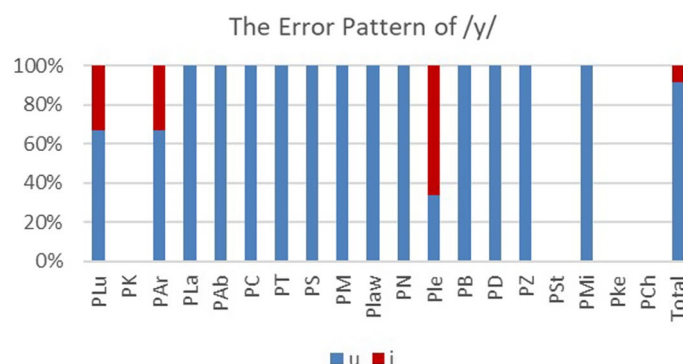


Fig. 1 Participants' error pattern of the Mandarin vowel /y/

Experiment 2: categorization task

The study aimed to explore the factors contributing to the weighting of cues in identifying the Mandarin vowel /y/ among participants. To achieve this, a categorization task was conducted to see how L1 English speakers perceive the vowel. The PAM-L2 uses the perceptual mapping task to evaluate the initial perception of the target L2 sounds among L2 learners, which is a well-established method for categorizing L2 vowels based on L1 vowels. This was accomplished by measuring the mean correspondence percentages and mean goodness-of-fit ratings, as described in previous studies by Guion et al. (2000), Strange et al. (2004), and Tyler et al. (2014). Hence, a perceptual mapping task is utilized in this context.

Participants

In this task, 11 naïve English speakers (6 male; 5 female) with no knowledge of Mandarin and limited immersion experience with any L2 participated. None of the participants considered themselves able to converse in or understand the non-native language in conversational settings. Their average age was 25.4. Two native speakers of Mandarin were recruited as speakers, one female and one male, aged 25 and 36 years. All listeners and speakers reported having normal hearing.

Stimuli

In this task, the same isolated Mandarin monosyllables syllables (ju/ɬy/-qu/ɬy/-xu/ɬy/-ji/ɬi/-qi/ɬi/-xi/ɬi/-zhu/ɬu/-shu/ɬu/-tɕ/ɬu/) used in Experiment 1 were presented as stimuli (as described in section "Stimuli"). English reference vowels are the four English high vowels (i/, /ɪ/, /u/, and /ʊ/). To ensure that the onset consonant had a standardized effect, English reference vowels were added to words beginning with the alveolar fricative /s/ (as in "sea") or the palato-alveolar fricative /ʃ/ (as in "sheep"), thus creating English syllables. The participants were provided with a piece of paper containing capitalized English reference words that represented the four English high vowels: shEEp /i/, shIp /ɪ/, shOE /u/, and shOOk /ʊ/.

Procedure

Each participant was given directions on how to complete the assignment. Participants studied the guidelines and inquired as necessary. Participants were able to complete the

Table 3 Perceptual categorization of Mandarin /y/ in terms of English vowels

Target Mandarin vowel	English vowel				
	i	ɪ	u	ʊ	Others
y	20% (3)	3.3% (2.5)	70.0% (3.6)	3.3% (3.5)	3.4%

The table presented the percentage of categorization (%) and goodness-of-fit ratings are in parentheses

tasks in the same sequence since Praat delivered the various stimulus tokens in the same order. Each stimulus was presented to the participants once, and they were asked to choose the closest English vowel by writing down the digits that represented it. After that, participants were asked to choose a number (1–5) to represent whether the Mandarin vowel stimulus was a bad example of the English vowel/approximant they had selected (1) or a good example (5) of the same. Participants would select the “others” option if they think that none of the given English vowels are similar to the targeted Mandarin vowel stimulus, and there is no need to assess the goodness of their selections.

Responses from participants were not timed and could take as much time as they chose. Each item could only be heard once; however, they might hear it again by requesting permission from the researcher. 11 participants ultimately answered 18 tokens, with 3 tokens for each of the 6 varieties of vowels, totaling 198 responses. There were 66 total responses, 66 for each type of vowel. The number of times each English vowel was matched to a specific Mandarin vowel was divided by 66 to determine the mean correspondence percentage for each Mandarin-English vowel match. By adding up all the goodness-of-fit ratings and dividing by the total number of matches, the mean goodness-of-fit rating was determined.

Results

The findings indicated that the English vowel category that was most commonly used to classify the Mandarin vowel /y/ was the English vowel /u/. Specifically, the Mandarin vowel /y/ was most frequently categorized as the English vowel /u/ (with a mean goodness-of-fit rating of 3.6 and a frequency of 70%), followed by the English vowel /i/ (with a mean goodness-of-fit rating of 3 and a frequency of 20%) (Table 3).

Based on the findings of a paired sample t-test ($t = -3.778$, $df = 10$, $p = 0.004$), the proportions of the two types of classifications (i.e., /u/ type and /i/ type) for the vowel /y/ differed significantly. The categorization of the Mandarin vowel /y/ in relation to English vowels was analyzed based on individual outcomes. Figure 2 presented the proportions of the Mandarin vowel /y/ that were classified as either English /i/ or /u/ (refer to Table 2). The results indicated that most participants categorized the Mandarin high-front rounded vowel /y/ into the English high-back rounded vowel /u/ category, which is absent in Mandarin. However, two listeners (PNa and PCa) classified /y/ as English /i/ more frequently than /u/, while one listener (PAN) classified /y/ equally into both English /i/ and /u/.

Discussion

The perceptual identification and categorization tasks conducted in the study indicate that native English speakers categorize the Mandarin high vowel /y/ perceptually into the Mandarin high vowel /u/ category, and make more errors when distinguishing between /y-u/ than /y-i/. These findings are consistent with previous research on the

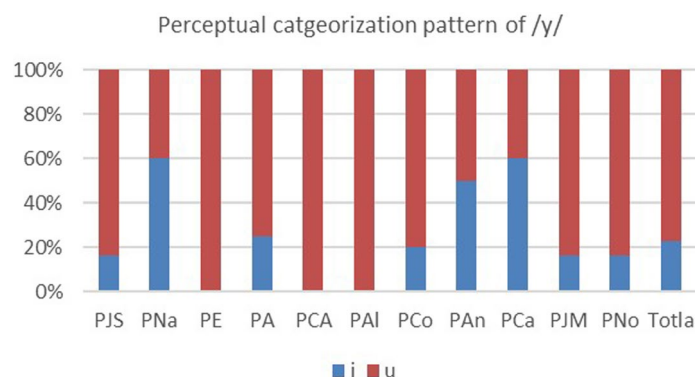


Fig. 2 Perceptual categorization pattern of Mandarin /y/

categorization of the German and French high-back rounded vowel /y/ by native English speakers. An acoustic comparison by Strange et al. (2004) showed that although the German vowel /y/ is categorized by native American English speakers into the American English back high-rounded vowel category /u/, it is most similar to the American English front high unrounded vowel /i/. In addition, according to Strange et al. (2009), native American English speakers categorize the French front high-rounded vowel (/y/) as an American English back high-rounded vowel /u/, rather than an American English front high unrounded vowel /i/. However, Japanese and Korean learners of Mandarin mostly confuse the Mandarin vowel /y/ with the Mandarin vowel /i/ (Wang, 2001; Wang & Deng, 2009). According to Wang (2001), Japanese and Korean L2 learners of Mandarin perceptually categorize the Mandarin /y/ as the Japanese vowel /i/ 44% of the time and /u/ 28% of the time, as well as the Korean vowel /i/ 58% of the time and /u/ 11% of the time. This makes it difficult for them to differentiate between the vowels /y/ and /i/ in Mandarin (Lee, 2010; Wang & Deng, 2009).

Given that fricatives and affricates are the only preceding consonants selected to normalize the influence of onset consonants, and that no errors were observed in identifying the /i-u/ stimuli (as described in Sect. "Results"), it can be concluded that the learners' perceptual confusion is not attributable to onset consonants. Therefore, in order to understand the reasons for the conflicting tendencies in the perception of Mandarin /y/ among L1 learners, it is necessary to compare L2 and L1 vowel patterns. Mandarin has three high vowels, of which /i-y/ are front and high and distinguished by lip-rounding, while /u-y/ are high and rounded, distinguished by backness. In contrast, Japanese has five vowels, including only one back mid-rounded vowel, and does not distinguish vowel pairs by lip-rounding (Akamatsu, 1997; Labrune, 2012; Nishi et al., 2008; Vance, 1997). Consequently, L1 Japanese learners of Mandarin exhibit lower sensitivity to the lip-rounding feature and instead rely more on the backness feature to perceive target vowels. These learners demonstrate better proficiency in distinguishing the Mandarin vowel contrast /y-u/ as compared to /y-i/. Acoustic studies suggest that the frequency of the second formant (F2) is closely linked to the backness feature, while the frequency of the third formant (F3) is closely linked to the roundness feature (Bao & Lin, 2014). Yamada and Tohkura (1992) and Iverson et al. (2003) reported that Japanese listeners pay greater attention to F2 frequency than F3 frequency, indicating that L1 Japanese learners are more sensitive to backness than roundness. However, it is not obvious why L1 Korean

Table 4 High vowels in Mandarin, English, Korean, and Japanese (Qian, 2017)

	Front		Back	
	Unrounded	Rounded	Unrounded	Rounded
Mandarin	i	y		u
English	i:(tense) ɪ(lax)			u:(tense) ʊ(lax)
Korean	i	(yɾl)	ɯ—	uɯ
Japanese	iɪ		ɯɯ	

speakers also rely more on backness than lip-rounding, or why L1 English speakers rely on roundness, given that these two languages' vowels are distinguished by both backness and lip-rounding. Please see Table 4 for the high vowels in the four languages.

The number of monophthongs in Korean is a subject of debate, with varying claims ranging from seven to ten (Ahn & Iverson, 2005; An, 1998; Brown et al., 2015; Franklin & Stoel-Gammon, 2014; Ha et al., 2009; Heo, 2013; Jin, 2012; Yang, 1996). In the ten-vowel system (/i y ɯ u e ɔ ʌ o æ a/), there is a distinction between roundness and non-roundness for high front and back vowels. In the seven-vowel system (/i ɯ u e ʌ o a/), there is no high front rounded vowel /y/. Furthermore, the weakening of the lip-rounding distinction can be observed in the high front vowels of the Korean language, as the /y/ and /ø/ diphthongs have evolved into /wi/ and /we/, as noted by previous studies (Heo, 2013; Jin, 2012). The weakening of the lip-rounding distinction in high front vowels suggests that Korean speakers currently place more emphasis on backness than roundness features. Compared to L1 Japanese learners, L1 Korean learners are less likely to confuse the vowels /y/ and /i/ due to the presence of a rounded/unrounded contrast in the Korean vowel system. Evidence from Wang (2001)'s study shows that low-experienced Korean learners of Mandarin are less likely to make discriminating errors between /y/ and /i/ than moderately experienced Japanese learners of Mandarin. Therefore, it is reasonable to assume that low-experienced Japanese learners would have higher mean discrimination error percentages than low-experienced Korean learners.

L1 English learners of Mandarin likely classify high vowels based on the roundness characteristic more than backness. This is because the rounded vowels /y/ and /u/, which only differ in backness but share the same roundness, are grouped in the same L1 English vowel category. The confusion between /y/ and /u/ is also common among L1 English learners due to their shared roundness. Although backness is a distinctive feature, it appears to be less important than roundness, which is traditionally considered to be the redundant or amplified quality of backness in English. However, data from L2 learning suggests that roundness may be more significant than previously thought, and acoustic evidence supports this claim. According to the study conducted by Iverson et al. (2003), it was found that F3 frequency is more influential than F2 frequency for American English speakers. This suggests that L1 English learners prioritize roundness over backness in their perceptual attention. The position of English /u/ in the acoustic vowel space is quite central. Certain varieties of English, such as New Zealand English, consider its high vowel /u/ to be a central rather than a back vowel (Hay et al., 2008), which suggests that the distinction in backness between high New Zealand English

vowels has been diminishing. According to Stevens et al. (1986), the enhancement feature theory suggests that enhancement features may substitute for the distinctive feature they enhance. De Jong (1995) claimed certain dialects of American English demonstrate a shift in the function of backness, where roundness assumes its role. Therefore, as the distinctive feature of backness weakens, L1 English speakers may rely more on the lip-rounding feature.

Conclusion

In this study, the perception of three Mandarin high vowels by L1 English speakers was investigated using two tasks: a perceptual identification task and a cross-language perceptual mapping task. Results showed that L1 English learners confused the high back rounded vowel /u/ and the high front rounded vowel /y/ the most, with a high mistaken rate for /y/ as /u/. The perceptual mapping experiment also revealed that /y/ was frequently classified as the English vowel /u/ and, to a lesser extent, as the English vowel /i/. According to PAM-L2, L1 English speakers perceptually categorized Mandarin high front rounded vowel /y/ and high back rounded vowel /u/ into the same vowel category, which lead to large perceptual confusion of the /y-u/ contrast. This confusion is consistent with L1 English learners' perception of other high-front rounded vowels but not with the perception of L1 Japanese and Korean learners. The study proposes a perceptual weighting of roundness and backness distinctive features in L2 learners' perception of Mandarin high vowels, which varies based on the learner's L1 background. L1 English speakers rely more on the round/unround contrast, while L1 Japanese and Korean learners rely more on the backness feature. This weighting is important for investigating L2 learners' perception of nonnative contrasts and for language development. Using other languages as target L2s, such as Mandarin, allows for an examination of the role of different cues in language learning. The evidence of perceptual weighting adds to the study of weighting in language learning and may inspire further research in this area.

The study's implications extend to pedagogy, with a proposed method for teaching Mandarin high vowels to L2 learners who struggle with perceiving/producing the Mandarin /y/ vowel in class. The current method focuses on teaching the Mandarin vowel /i/ first, which is considered one of the easiest vowels for L2 learners, and then adding lip-rounding onto it. Students are asked to perceive/produce /i/ first and then to add lip-rounding onto it. This emphasizes the distinctive feature of 'roundness', which is suitable for L1 English students who are sensitive to roundness and have no difficulty in rounding their lips. However, L1 Korean and Japanese learners may find this method ineffective due to their insensitivity to the roundness feature, and should instead be reminded of the backness between Mandarin vowels /u/ and /y/. The proposed method highlights the importance of adapting teaching strategies to L1 learners' perceptual cue weighting and backgrounds when teaching Mandarin high vowels.

In the future, several areas could benefit from further research. One such area is the investigation of how acoustic properties, particularly those related to roundness and backness, affect the perception of Mandarin high vowels by L2 learners from various language backgrounds. This could involve analyzing vowel F2 values in synthetic and/or natural stimuli. Another area for future research is the examination of individual differences among L1 and L2 listeners. While most L1 English participants prioritize backness

over roundness, a few participants in the current study exhibited the opposite strategy, categorizing and identifying /y/ as /i/ instead of /u/. To further validate the results and explanations of the current study, future research may also need to involve a larger number of participants.

Abbreviations

SLA	Second language acquisition
L1	First language
L2	Second language
SLM	Speech learning model
PAM-L2	Perceptual assimilation model for L2 learners
F1	The first formant
F2	The second formant

Acknowledgements

We would like to thank Dr. Helen Charters and Dr. Jason Brown from the University of Auckland for their questions and remarks. This research is sponsored by Grant 22YH61D from the 2022 International Chinese Language Education Research Topic Youth Project Funding and grant XJZLGC202206 from the Undergraduate Teaching Quality and Teaching Reform Project of Southern University of Science and Technology.

Author contributions

WZ and S-HL are in charge of conceptualization, methodology, and editing. WZ wrote the main manuscript text. XZ validated data and prepared figures. All authors reviewed the manuscript.

Funding

This work was supported by the 2022 International Chinese Language Education Research Topic Youth Project Funding (Grant Number 22YH61D) and the Undergraduate Teaching Quality and Teaching Reform Project of Southern University of Science and Technology (Grant Number XJZLGC202206).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

All authors report no competing interests.

Received: 25 March 2023 Accepted: 13 July 2023

Published online: 15 September 2023

References

- Ahn, S. C., & Iverson, G. K. (2005). Structured imbalances in the emergence of the Korean vowel system. *Historical Linguistics*.
- An, S. C. (1998). *An introduction to Korean phonology*. Hansin Munhwasa.
- Akamatsu, T. (1997). *Japanese phonetics: Theory and practice*. Lincom Europa.
- Bao, H., & Lin, M. (2014). *实验语音学概要 (增订版) [eEssentials of experimental phonetics]* (revised). Peking University Press.
- Bauer, L., Warren, P., Bardsley, D., Kennedy, M., & Major, G. (2007). New Zealand English. *Journal of the International Phonetic Association*, 37(01), 97–102.
- Best, C. T., Tyler, M., Bohn, O., & Munro, M. (2007). Nonnative and second-language speech perception. *Language Experience in Second Language Speech Learning*, 13–34.
- Bi, Y. (2001). 美国学生学习汉语的声韵难点分析 [Analysis of difficulties on Mandarin initials and finals by American students]. *辽宁工学院学报(社会科学版) [Journal of Liaoning Institute of Technology] (Social Sciences Edition)*, 3(2), 39–41.
- Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (version 5.1.05) [computer program]. <http://www.praat.org>
- Bohn, O. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 279–304.
- Bohn, O., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, 11(3), 303–328.
- Brown, L., & Yeon, J. (Eds.). (2015). *The handbook of Korean linguistics*. Wiley.
- Casillas, J. (2015). Production and perception of the /i/-/I/ vowel contrast: The case of L2-dominant early learners of English. *Phonetica*, 72(2–3), 182–205.
- De Jong, K. (1995). On the status of redundant features: The case of backing and rounding in American English. *Phonology and Phonetic Evidence: Papers in Laboratory Phonology, IV*, 68–86.

- Duanmu, S. (2007). *The phonology of standard Chinese*. OUP.
- Escudero, P., Benders, T., & Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics*, 37(4), 452–465.
- Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, 26(4), 551–585.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 92, 233–277.
- Flege, J. E., Bohn, O., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437–470.
- Franklin, A. D., & Stoel-Gammon, C. (2014). Using multiple measures to document change in English vowels produced by Japanese, Korean, and Spanish speakers: The case for goodness and intelligibility. *American Journal of Speech-Language Pathology*, 23(4), 625–640.
- Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America*, 107(5), 2711–2724.
- Guo, X., & Chen, X. (2017). 北京话和粤语背景学习者英语词重音产出研究 [Study on English stress production by learners with backgrounds in Beijing Mandarin and Cantonese]. *外语教学与研究 [Language Teaching and Linguistic Studies]*, 49(2), 188–201.
- Ha, S., Johnson, C. J., & Kuehn, D. P. (2009). Characteristics of Korean phonology: Review, tutorial, and case studies of Korean children speaking English. *Journal of Communication Disorders*, 42(3), 163–179.
- Hao, Y. C. (2018). Second language perception of Mandarin vowels and tones. *Language and Speech*, 61(1), 135–152. <https://doi.org/10.1177/0023830917717759>
- Hay, J., MacLagan, M., & Gordon, E. (2008). *New Zealand English*. Edinburgh University Press.
- Heo, Y. (2013). An analysis and interpretation of Korean vowel systems. *Acta Koreana*, 16(1), 23.
- Holt, A., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of America*, 119(5), 3059–3071. <https://doi.org/10.1121/1.2188377>
- Huang, B., & Liao, X. (1983). 现代汉语 [Modern Mandarin]. 甘肃人民出版社 [Gansu People's Publishing House].
- Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87(1), B47–B57.
- Jin, W. (2012). Variation and change in Mandarin Korean: The case of vowel/y. *Language Variation and Change*, 24(1), 79–106.
- Kondaurova, M. V., & Francis, A. L. (2008). The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *The Journal of the Acoustical Society of America*, 124(6), 3959–3971.
- Labruno, L. (2012). *The phonology of Japanese*. Oxford University Press.
- Ladefoged, P., & Johnson, K. (2014). *A course in phonetics*. Cengage learning.
- Lee, S. (2010). *Korean learners' cognitive process of Mandarin mono-vowel categorization (unpublished doctoral dissertation)*. Beijing Language and Culture University.
- Levy, E. S. (2009). On the assimilation–discrimination relationship in American English adults' French vowel learning. *The Journal of the Acoustical Society of America*, 126(5), 2670–2682.
- Levy, E. S., & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, 36(1), 141–157.
- Li, J., & Liu, J. (2008). 美国学生汉语中介语元音系统建构次序的实验研究 [Experimental Study on the construction order of interlanguage vowel systems in American students learning Mandarin Chinese]. *现代外语 [Modern Foreign Languages]*, (3), 310–316.
- Lin, Y. (2007). *The sounds of Chinese*. Cambridge University Press.
- Lu, J. (1984). 中介语理论与外国人学习汉语的语音偏误分析 [Interlanguage theories and an analysis of phonetic errors in Mandarin learning by foreigners]. *语言教学与研究 [Language Teaching and Linguistic Studies]*, 3, 44–56.
- Ni, Y., & Wang, X. (1992). 英语国家学生学习汉语语音难点分析 [Analysis of difficulties in mandarin Chinese pronunciation for English-speaking students]. *汉语学习 [Chinese Language Learning]*, (2), 47–50.
- Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., & Trent-Brown, S. A. (2008). Acoustic and perceptual similarity of Japanese and American English vowels. *The Journal of the Acoustical Society of America*, 124(1), 576–588.
- Qian, Y. (2017). *A study of Sino-Korean phonology: Its origin*. Routledge.
- Schertz, J., & Clare, E. J. (2020). Phonetic cue weighting in perception and production. *Wiley Interdisciplinary Reviews: Cognitive Science*, 11(2), e1521. <https://doi.org/10.1002/wcs.1521>
- Shultz, A. A., Francis, A. L., & Llanos, F. (2012). Differential cue weighting in perception and production of consonant voicing. *The Journal of the Acoustical Society of America*, 132(2), EL95–EL101.
- Stevens, K. N., Keyser, S. J., & Kawasaki, H. (1986). Toward a phonetic and phonological theory of redundant features. *Invariance and Variability in Speech Processes*, 426–449.
- Strange, W., Bohn, O. S., Trent, S. A., & Nishi, K. (2004). Acoustic and perceptual similarity of North German and American English vowels. *The Journal of the Acoustical Society of America*, 115(4), 1791–1807.
- Strange, W., Levy, E. S., & Law, F. F. (2009). Cross-language categorization of French and German vowels by naïve American listeners. *The Journal of the Acoustical Society of America*, 126(3), 1461–1476.
- Strange, W., Weber, A., Levy, E. S., Shafiro, V., Hisagi, M., & Nishi, K. (2007). Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *The Journal of the Acoustical Society of America*, 122(2), 1111–1129.
- Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, 71(1), 4–21. <https://doi.org/10.1159/000356237>
- Vance, T. J. (1997). *An introduction to Japanese phonology*. SUNY Press.

- Wang, Y. (2001). 韩国, 日本学生感知汉语普通话高元音的初步考察 [A preliminary study of the perception of high vowels in Mandarin by Korean and Japanese learners]. *语言教学与研究* [Language Teaching and Linguistic Studies], (6).
- Wang, X., & Munro, M. J. (1999). The perception of English tense-lax vowel pairs by native Mandarin speakers: The effect of training on attention to temporal and spectral cues. In *Proceedings of the 14th international congress of phonetic sciences* (Vol. 3, pp. 125–128). University of California.
- Wang, Y., & Deng, D. (2009). 日本学习者对汉语普通话“相似元音”和“陌生元音”的习得 [The acquisition of the “unfamiliar vowels” and “similar vowels” in Mandarin by Japanese learners]. *世界汉语教学* [Mandarin Teaching in the World], 2, 262–279.
- Yamada, R. A., & Tohkura, Y. (1992). The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Perception and Psychophysics*, 52(4), 376–392.
- Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, 24(2), 245–261.
- Zhu, C., & Wang, J. (1997). 对外汉语中介音类型研究 [A study on the sound type for Mandarin as an interlanguage]. 第五届国际汉语教学讨论会论文选 [the 5th international conference on mandarin language pedagogy].

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
