Proceedings of the 22nd North American Conference on Chinese Linguistics (NACCL-22) & the 18th International Conference on Chinese Linguistics (IACL-18). 2010. Vol 1. Clemens, L.E. & C.-M. L. Liu, eds. Harvard University, Cambridge, MA. 429-437.

Perceiving Vowels and Tones in Mandarin: The Effect of Literary Phonetic Systems on Phonological Awareness

Hsin-Ni Lin[1], Chien-Jer Charles Lin[2]
[1]*National Taiwan Normal University*, [2]*Indiana University*

A tone-vowel monitoring task similar to Ye and Connine's (1999) experiment was conducted in Taiwan to examine how literary phonetic systems affect people's perception of vowels, and then the relative temporal availability of vowel and tonal information. The results demonstrated that although participants were sensitive to the acoustic nature of monosyllabic stimuli, they mainly resorted to the literary systems they learned during the perception. Therefore, the impact of literary systems on the perception was taken into consideration when the relation of vowels and tones was investigated in this study. With the consideration, it was found that vowel information was available relatively earlier than the tonal information.

1. Introduction

A syllable in Mandarin is composed of segments and a tone. Not only segments but also the tone contributes to distinguishing meanings of words. For instance, in Mandarin, there are four distinct tones; when they are respectively combined with the syllable *ma*, each combination has a different meaning: *ma* means 'mother'; *ma2* means 'hemp'; *ma3* means 'horse'; *ma4* means 'scold.' Therefore, to understand how the auditory information of a Mandarin syllable is processed, the perception of segments and tones is a fundamental issue.

In recent studies concerning the perception of Chinese segments and tones, there has been growing interest in the relationship of vowels and tones (i.e. which information among the two is processed first and which information is more important during perception) (Ye & Connine, 1999; Lee, 2007; Liu & Samuel, 2007; Tong, Francis & Gandour, 2007; Malins & Joanisse, 2010). Although considerable research has been devoted to the relationship of vowels and tones, rather little attention has been paid to the possibility that the literary education about Chinese phonetics may influence the relationship. According to Cheung and Chen (1997), and Shu et al. (2008), learned phonetic systems affect people's phonological awareness.[1] For example, Cantonese

[1] Phonological awareness refers to a person's ability to segment and manipulate sound units. Depending on the smallest unit a person can segment and manipulate, phonological awareness includes three hierarchical levels: (1) syllable awareness, (2) demi-syllable awareness (i.e. onset and rime awareness), and (3) phoneme awareness (Treiman, 1987). In the hierarchy,

speakers who have learned Pinyin[2] can recognize and manipulate phonemes while those who did not are able to manipulate only syllables. With this regard, it is highly possible that a phonetics system would impact the perception of segments, and then further influence the relationship between vowels and tones. The primary objective of this paper is to verify this possibility.

Our research is built on previous research by Ye and Connine (1999), who conducted a tone-vowel monitoring task. In the task, Mainland Chinese participants were administered to respond whether Mandarin monosyllabic stimuli contained both the vowel /a/ and Tone 2 (e.g. *ba2*, *lai2*, *yan2*). By manipulating the types of mismatches in the syllables (including tonal mismatches such as *ba4*, and vowel mismatches such as *bi2*), they found the mean reaction time for vowel mismatches was significantly shorter than that for tonal mismatches. Based on the result, Ye and Connine concluded that vowel information has "perceptual advantage" over tonal information.

Given the discrepancies between the phonetic systems taught in schools in China and Taiwan (relevant examples summarized in Table 1), it is likely that Mainlanders and Taiwanese would perform differently in Ye and Connine's perceptual task.

Table 1. Phonetic representations of syllables with /a/ in Mainland China and Taiwan

	a. /ba/	b. /ban/	c. /baη/	d. /bai/	e. /bau/
Mainland Pinyin	ba	ban	bang	bai	bao
Taiwan Zhuyin[3]	ㄅㄚ	ㄅㄢ	ㄅㄤ	ㄅㄞ	ㄅㄠ

(note: ㄅ = /b/, ㄚ = /a/, ㄢ = /an/, ㄤ = /aη/, ㄞ = /ai/, ㄠ = /au/)

Hence, this study aims to find the answers to three research questions:

- Do the literary phonetic transcription systems affect the perception of vowels? (i.e. phonological awareness)
- If the answer to the previous question is yes, does the acoustic nature of sounds still play a role during the perception?

phoneme awareness is the most difficult level to achieve; once one is capable of it, he is surely able to achieve the other two levels.

[2] Pinyin is a phonetic system used to transcribe segments in Cantonese or Mandarin in Mainland China. In the Pinyin system, onsets and rimes of Cantonese or Mandarin are represented by 25 Roman alphabets or the combinations of some of them, such as "b", "p", "a", "ai", and "an".

[3] Unlike Pinyin, Zhuyin includes 37 symbols to represent each of the onsets and rimes in Mandarin. For example, "b", "p", "a", "ai", and "an" in Pinyin are respectively represented as "ㄅ", "ㄆ", "ㄚ", "ㄞ" and "ㄢ" in Zhuyin. The biggest difference between the Pinyin system and then Zhuyin system is in the part of rimes (Wang, 1997). In the Zhuyin system, a rime is represented by only a single symbol; in Pinyin system, every phoneme in a rime is explicitly presented.

- When literary phonetic differences are taken into account, what is the relationship between vowel and tonal information during the perception?

An experiment similar to Ye and Coninne's was conducted in Taiwan. Participants were also asked to decide whether a syllable contained the combination of tone 2 and the vowel /a/. The method of the experiment is presented in Section 2. Section 3 provides results and discussion on the collected data. Finally, Section 4 concludes the current study.

2. Method

2.1. Participants

Twenty-one undergraduate students (3 males and 18 females) were recruited as participants from departments other than Department of English and Department of Chinese[4] in National Taiwan Normal University. However, only the data collected from 18 of them (3 males and 15 females) were analyzed since the data-recording software ran into some problem when the other three participated in the experiment.

All of the participants were native Mandarin speakers in Taiwan. They were all right-handed and had no history of brain lesions. In addition, none of them had learned Pinyin phonetic system.

2.2. Materials

Since this study is built on the research of Ye and Connine (1999), the materials in our experiment were similar to those in their study. In this experiment, seven types of stimuli were included. 6 types of them were critical items, and the stimuli in the other type were filler items. The types of critical items are displayed in the following.

[4] Students in Department of English and Department of Chinese were not included because they had learned Mandarin Phonetics, which may render them to possess general knowledge about how Mandarin sounds are represented in the Pinyin phonetic system.

Table 2. Types of critical stimuli in the experiment

Syllable Type		Rime and Tone within the Syllable: rime	Rime and Tone within the Syllable: tone	Example	Expected response
Type 1	Base Syllable	a (ㄚ)	Tone 2	*ba2* (ㄅㄚˊ)	Yes
Type 2	Rime-mismatched Syllable <i>	i or u (ㄧ or ㄨ)	Tone 2	*bi2* (ㄅㄧˊ)	No
Type 3	Tone-mismatched Syllable	a (ㄚ)	Tone 4	*ba4* (ㄅㄚˋ)	No
Type 4	Double-mismatched Syllable <i>	i or u (ㄧ or ㄨ)	Tone 4	*bi4* (ㄅㄧˋ)	No
Type 5	Rime-mismatched Syllable <ii>	an, ang, ai, or ao (ㄢ, ㄤ, ㄞ, or ㄠ)	Tone 2	*bai2* (ㄅㄞˊ)	No
Type 6	Double-mismatched Syllable <ii>	an, ang, ai, or ao (ㄢ, ㄤ, ㄞ, or ㄠ)	Tone 4	*bai4* (ㄅㄞˋ)	No

Type 1 stimuli were the target-bearing stimuli which contained the combination of the vowel /a/ and Tone 2. This type was named "Base Syllable" because stimuli in Type 2 (Rime- mismatched Syllable <i>), Type 3 (Tone-mismatched Syllable), and Type 4 (Double-mismatched Syllable <i>), which were non-target-bearing stimuli, were all created based on stimuli in Type 1. There were twelve stimuli in each of Type 1- 4.

Type 5 (Rime-mismatched Syllable <ii>), like Type 2, also consisted of rime-mismatched syllables; Type 6 (Double-mismatched Syllable <ii>), like Type 4, included double-mismatched syllables. However, unlike the rimes in Types 2 and 4, the rimes in Types 5 and 6 were an (ㄢ), ang (ㄤ), ai (ㄞ), and ao (ㄠ), all of which contained an /a/ vowel that is not explicitly presented in Zhuyin. We chose these rimes in order to examine whether Taiwanese Mandarin speakers were able to detect the assimilated vowel /a/ in these four rimes. In addition, unlike Types 2 and 4, Types 5 and 6 were not created based on Type 1 because the rimes in Types 5 and 6, an (ㄢ), ang (ㄤ), ai (ㄞ), and ao (ㄠ), could not be combined with some of the onsets in Type 1 stimuli due to accidental syllabic gaps in Mandarin. Therefore, we created Type 5 (Rime-mismatched Syllable <ii>), by choosing 4 syllables for each of the four rimes, an (ㄢ), ang (ㄤ), ai (ㄞ), and ao (ㄠ). As for Type 6, it was the counterpart of Type 5 by altering Tone 2 in Type 5 stimuli into Tone 4. Consequently, there were 16 items in each of Types 5 and 6. All stimuli in the six types introduced above are given in Appendix. The mean word frequency[5] of each type was controlled to be around 310 (ranging from 290-323), except for Type 4.

In addition to the aforementioned six types, the fillers made of a (ㄚ) and Tone 2 were added to make the proportion of "Yes stimuli" to "No stimuli" 1 to 1. Most of the fillers were non-existent syllables in Mandarin so as to avoid the experimental design having the tendency to guide participants to monitor the sound stimuli by relating those sounds to the Zhuyin phonetic representations, thus obtaining artificially biased results.

Each stimulus in the seven types was recorded five times in a sequence by a female native Mandarin speaker in a quiet recording studio. Later, we cut one sound from the five tokens through visual and auditory scrutiny. To enable the volumes of the sounds to be similar, some of the sounds were amplified through SoundEdit Pro 2.1.

2.3. Design

Each participant heard all of the stimuli in random orders (presented by E-prime 2.0). Following Ye and Connine's design, a participant was allowed as much time as they needed to press a response button; in other words, only after a participant pressed the "Yes" or "No" response button did the next trial appear. Also, between a participant's response and the next trial, there was a pause of 250ms.

2.4. Procedure

Participants were tested individually in a quiet experimental studio. They were instructed to judge whether a combination of a (ㄚ) and Tone 2 were contained in each Mandarin monosyllable presented. If yes, they should press the "Yes" button; if no, they had to press the "No" button.[6] The instructions were auditorily presented instead of visually; therefore, the instruction can be avoided misleading participants to perceive a (ㄚ) sound in the instruction by retrieving the Zhuyin phonetic representations.

In addition, participants were told to press a response button as soon as possible since their accuracy and reaction time would be recorded. Furthermore, they were also informed about the inclusion of non-existent syllables in Mandarin in the experiment stimuli.

Prior to the experiment, there was a practice session. The session included 10 trials which were distinct from stimuli included the experiment. Participants were asked to judge whether a stimulus contained a combination of e (ㄜ) and Tone 2.

[5] Word frequency was calculated from Academic Sinica's website "Sou Ci Xun Zi".

[6] For half of the participants, "Yes" button was on the right in the response box; for the other half, "Yes" button was on the left.

3. Result & Discussion

Mean accuracy and reaction time for Type 1-6 are provided in Table 3. The accuracy is based on the expected responses of each type (see Table 2). Reaction time (RT, henceforth) was calculated both from the onset and the offset of a stimulus, and those less than 200 ms and greater than 2000 ms from the onset were removed from the analysis (1%). In addition, in the calculation of the mean reaction time to each Type, only those of correct responses (96%) were included.

Table 3. Accuracy and Mean RTs of Type 1-Type 6

Syllable Type		**Examples**	**Accuracy** (percentage)	**RT measured from Onset** (ms)	**RT measured from Offset** (ms)
Type 1	Base Syllable	*ba2* (ㄅㄚˊ)	95.33%	850	291
Type 2	Rime-mismatched Syllable <i>	*bi2* (ㄅㄧˊ)	96.74%	903	300
Type 3	Tone-mismatched Syllable	*ba4* (ㄅㄚˋ)	94.37%	880	403
Type 4	Double-mismatched Syllable <i>	*bi4* (ㄅㄧˋ)	100.00%	781	273
Type 5	Rime-mismatched Syllable <ii>	*bai2* (ㄅㄞˊ)	91.52%	966	364
Type 6	Double-mismatched Syllable <ii>	*bai4* (ㄅㄞˋ)	99.65%	844	351

A striking difference between Ye and Connine's results and ours is shown in the responses to Type 5 here. As predicted, Taiwanese participants (91.52%) responded "no" to those monosyllables containing Tone 2 and rimes *an*, *ang*, *ai*, or *ao*. The Mainlanders in Ye and Connine (1999), however, responded "yes" to the kind of auditory inputs. The contrast revealed that the literary phonetic transcription systems affect the perception of vowels. Mainland participants responded by resorting to Pinyin whereas Taiwanese participants resorted to Zhuyin.

Besides, within Type 5, syllables containing *an2* was easier to reject than syllables containing *ang2*. The accuracy of *an2* (95.56%) was significantly higher than that of *ang2* (80.00%) [t (137) = 2.89, $p < .01$]. Also, the average RT for *an2* was 116ms shorter than that of *ang2* [t (120) = -2.43, $p < .05$].[7] This inconsistency between the *an2*-*ang2* pair might be because the greater phonetic alternation from [a] was involved in *an2*, in which the back vowel /a/ preceded an alveolar consonant /n/, than in *ang2*, in which the vowel was followed by a back consonant /η/. This suggests that participants were still sensitive to the phonetic quality of the vowels, even though their perception of vowels was influenced by the literary phonetic systems.

Finally, we probed into the relationship of vowel and tonal information while considering the influence of phonological awareness on the perception of vowels. The result showed that Type 2 (Rime-mismatched Syllable <i>) was rejected 103ms earlier than Type 3 (Tone-mismatched Syllable) [t_1 (34) = -1.96, p = .059; t_2 (22) = -2.92, $p < .01$].[8] This revealed that vowel information was available earlier than tonal information during the processing, which is congruent with Ye and Connine's findings. In other words, the perceptual advantage for vowels still exists when the literary differences are taken into account.

4. Conclusion

Results of the current study provide answers to the three research questions listed in the introduction section. During the perception of vowels, Mandarin native speakers mainly resort to their literary phonetic transcription systems while their perception the acoustically vowel information still plays a role. Additionally, when the impact of phonetic systems on vowel perception is taken into consideration, tonal information is still available subsequent to vowel information.

REFERENCES

Cheung, Him. and Chen, Hsuan-Chih. 2004. Early orthographic experience modifies both phonological awareness and on-line speech processing. *Language and Cognitive Processes* 19(1). 1-28.

Lee, Chao-Yang. 2007. Does horse activate mother? Processing lexical tone in Form

[7] Only statistics of data measured from the offset is reported here. The onset data also showed significant differences.

[8] The statistical results were based on data measured from the offset. The onset data showed no significance, which might be due to the differences in sound length of Type 2 and Type 3 stimuli.

Priming. *Language and Speech* 50 (1). 101-123.

Liu, Siyun. and Samuel, Arthur, G. 2007. The role of Mandarin lexical tones in lexical access under different contextual conditions. *Language and Cognitive Processes* 22(4). 566-594.

Malins, Jeffery, G. and Joanisse, Marc, F. 2010. The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. Journal of *Memory and Language* 62(4). 407-420.

Shu, Hua. Peng, Hong. and McBride-Chang, Catherine. 2008. Phonological awareness in young Chinese children. *Developmental Science* 11(1). 171-181.

Tong, Yunxia. Francis, Alexander L. and Gandour, Jackson T. 2007. Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language and Cognitive Processes* 22(8). 1-20.

Treiman, Rebecca. 1987. On the relationship between phonological awareness and literacy. *Cahiers de Psychologie Cognitive* 7(5). 524-529.

Wang, H. Samual. 1997. Cong yu yin qie fen fang shi kan han yu yin biao de zheng yi. *Hua Wen Shi Jie* 86. 62-66.

Ye, Yun. and Connine, Cynthia. M. 1999. Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes* 14(5). 609-630.

APPENDIX[9]

Type 1 Base syllable (12)	Type 2 Rime-mismatched syllable <i> (12)	Type 3 Tone-mismatched syllable (12)	Type 4 Double-mismatched syllable <i> (12)	Type 5 Rime-mismatched syllable <ii> (16)	Type 6 Double-mismatched syllable <ii> (16)
ba2	bi2	ba4	bi4	bai2	bai4
pa2	pu2	pa4	pu4	lai2	lai4
ma2	mi2	ma4	mi4	hai2	hai4
fa2	fu2	fa4	fu4	zhai2	zhai4
da2	du2	da4	du4	pan2	pan4
na2	ni2	na4	ni4	man2	man4
la2	li2	la4	li4	nan2	nan4
ga2	gu2	ga4	gu4	chan2	chan4
ha2	hu2	ha4	hu4	fang2	fang4
zha2	zhu2	zha4	zhu4	lang2	lang4
cha2	chu2	cha4	chu4	hang2	hang4
sha2	shu2	sha4	shu4	nang2	nang4
				bao2	bao4
				mao2	mao4
				nao2	nao4
				shao2	shao4

[9] The number in the parenthesis indicates the number of stimuli in each type.