Priming Effects of Tones and Segments in Lexical Processing in Mandarin

Lan Shuai [1], *Bin Li* [1], *Tao Gong* [2]

[1] Department of Chinese, Translation & Linguistics, City University of Hong Kong, Hong Kong
[2] Department of Linguistics, University of Hong Kong, Hong Kong

susan.shuai@gmail.com, binli2@cityu.edu.hk, tgong@hku.hk

Abstract

In tone languages, pitch variation is one of the primary cues to distinguish lexical meanings. Similar to consonants and vowels, lexical tones play a role in accessing to lexical information, but these segmental and suprasegmental features may have different roles in this process. In this study, we examine the priming effects of tones, consonants, and vowels in an auditory task of lexical decision in Mandarin, and compare such phonological priming effects with the direct and semantic priming effects. We find that consonants and vowels play a major role at the stage of lexical access, whereas tones facilitate lexical processing afterwards. In addition, partially conflicting segments, not tones, interfere with lexical selection. We discuss and compare these results with the previous findings to shed some lights on the follow-up research.

Index Terms: lexical tone, lexical decision, priming, Mandarin

1. Introduction

Human speech comprises different types of phonemes (the minimal distinguishable sound units). Consonants and vowels are universal phonemes in languages. They differ not only in constriction size in general, but also in the relevant neural processing [1, 2]. For example, differences roles of consonants and vowels were found in the online auditory processing of lexical items [3].

As stated in the TRACE model [4], in accessing to lexical meanings, one has to recognize the phonemes that build up the words, by assembling all sorts of acoustic features, each occupying a particular temporal place in the speech sounds of words. According to this model, the acoustic features at the initial position of lexical items could activate all words having the same onset phoneme, and in monosyllabic words, the acoustic features at the rime position could be used to select and constrain word meanings via lexical competition. This model connects phonemes with lexical processing in speech perception, and predicts different roles of phonemes at the onset and rime positions in lexical processing of monosyllabic words. However, it ignores tones, whose roles in lexical processing remain largely unknown.

Mandarin, as a tone language, serves as a good candidate to examine the role of tones in lexical processing. There have been no consensuses in the previous work on the role of tones in lexical processing. Some scholars tried to incorporate toneme processing into the TRACE model [5]. They assumed that tonemes were separated nodes from other segmental phonemes, and there should be no differences between initial consonants and vowels. This view advocated a relatively less important role of tones in isolated syllables than in syllables with context. However, other scholars observed an important role of tones in lexical processing, distinct from the role of segments in isolated syllables [6]. In addition, no previous studies reported any difference between Mandarin initial consonants and vowels, as observed in other languages. In Mandarin, characters are monosyllabic, with simple syllable structures V, CV, CVC, VC (C for consonant, V for vowel). According to the TRACE model, initial consonants should activate more lexical meanings than vowels in medial position. Since tones are superimposed largely on vowels and have similar temporal positions as vowels, they should have similar roles in lexical processing. Furthermore, lexical processing can be divided into two stages: lexical access and lexical selection, respectively taking place around 200 ms and 400 ms after the onset of lexical items [7, 8]. Predictions from the TRACE model would support the role of initial consonants in lexical access, and the role of both vowels and tones in lexical selection. But, it could also be possible that the early proportions of vowels take part in lexical access, whereas tones are referred to only for lexical selection, since pitch variation is usually recognized late in time.

Noting these controversies and uncertainties, in this study, we plan to investigate the roles of initial consonants, vowels, and tones in Mandarin lexical processing at the stages of lexical access and lexical selection. Our prediction is that: at the stage of lexical access, initial consonants have an important role, whereas at the stage of lexical selection, vowels and tones become more important.

2. Perception Experiment

We adopted the priming paradigm and embedded it in an auditory lexical decision task to examine the roles of initial consonants, vowels, and tones in lexical processing. This paradigm has been widely used in studies of Mandarin tone processing [6, 9]. In addition, we recruited two ISI (inter-stimuli interval) conditions, 200 ms and 400 ms, to match the stages of lexical access and lexical selection, respectively.

2.1. Materials

The testing materials included real monosyllabic Mandarin words and non-words. Real words were chosen from the mid-frequency Mandarin Characters [10]. All consonants, vowels, and tones in Mandarin were included in the materials. Non-words (NW) were syllables in conformity with Mandarin phonotactic rules but having no meanings or orthographic forms, e.g. [tʃua˥]. Participants were asked to judge whether the heard syllable was a real Mandarin word or not. Real words (RW) were used both as targets and as primes, whereas non-words were only used as targets.

In each trial, two syllables were presented in a pair. The first syllable served as the prime and the second as the target. In the priming trials, the prime syllable shared certain phonological or semantic similarity with the target. In the control trials, there was no similarity between the prime and the target. The target syllable was for the auditory lexical decision. Participants needed to respond to the target syllable by deciding whether it was a real Mandarin word or not.

There were 11 priming conditions. In the direct priming condition (Dir. Pri.), the prime was identical to the target. In the two consonant priming conditions (C. Pri.), the initial

consonant of the prime syllable was identical to that of the target syllable. Both RW and NW were used in these conditions. In the two vowel priming conditions (V. Pri.) or the two tone priming conditions (T. Pri.), the vowel or the tone of the prime syllable was identical to that of the target. Both RW and NW were used as target syllables. In the two semantic priming conditions (Se. Pri.), the target syllable could form a disyllabic Mandarin word with the prime syllable. In one of the two conditions, the target did not share any phoneme or toneme with the prime syllable. In the other condition, the target shared an identical tone with the prime. The purpose of introducing the second semantic priming condition was to investigate if the tone resemblance would introduce a greater semantic priming effect than the first semantic priming condition, where there was no phonological similarity. In the two control conditions, neither phoneme nor meaning was shared between the control prime and the target syllables. Both RW and NW syllables were used as targets.

Examples of the target syllables in each condition with [pi˩˩] as the prime syllable are shown in Table 1. There are 21 items in each condition. Considering the two ISI conditions, we generated 462 trials (11 conditions × 21 items × 2 ISI) for the experiment.

Table 1. *Examples of the target syllable of each condition with [pi˩˩] as the prime syllable.*

pi˩˩	Dir. Pri.	C. Pri.	V. Pri.	T. Pri.	Se. Pri.	Ctrl.
RW	pi˩˩	pa˥˩	$t^hi˥$	ʂu˩˩	tʂɹ˩˩ kaŋ˥	χə˥
NW	--	pe˥	$k^hi˥$	luaŋ˩˩	--	ky˥

All real word syllables were recorded from a female native Mandarin speaker, a trained phonetician. Non-word syllables were also recorded from the same speaker after practice. The recording was done in a sound proof room with a SHURE USB microphone. The microphone was approximately 20 cm apart from the speaker's mouth and pointed to the speaker with a 45 °horizontal angle. Praat [11] was used to record and digitize sounds, with a sampling rate of 44.1 kHz. There were two ISI (inter-stimuli interval) conditions. In the long (400 ms) ISI condition, only syllables around 400 ms long (±10 ms) were used. In the short ISI (200 ms) condition, all prime syllables were recorded for a second time with the duration kept around 200 ms (±10 ms). Intensity of each token was normalized to a RMS amplitude of 70 dB. The intensity of the prime syllables was adjusted to 55 dB, to avoid attracting much attention to the primes.

2.2. Participants and Procedure

Twenty-two University students (19 female, 3 male), who are native speakers of Mandarin, participated in this experiment. The average age was 23.64 (±2.19). They were tested in a quite classroom. They were instructed to focus on the louder sounds (targets), but to ignore the 'interfering' sounds with a smaller volume (primes), and then, to judge whether the louder sounds were Mandarin. They were encouraged to make the judgment as quickly and accurately as possible. They made the judgment by pressing the left and right buttons of a mouse using their right index and middle fingers. They were asked to adjust their seats and the mouse to a comfortable position at the beginning of the experiment, and keep their fingers on the mouse throughout the block. Half of the participants pressed the left mouse button for 'yes', and the right mouse button for 'no'. The order was reversed for the other half of the participants.

The presentation was designed and delivered by E-Prime. The whole experiment contained 6 blocks, each involving 77 trials and lasting around 4 minutes. Participants took rests between blocks. The whole experiment lasted around 30 minutes.

2.3. Analysis

The responses and reaction times were analyzed. The correct percent in the judgment of word status were used for screening. Data of the participants having over 80% accuracies were used for statistical analysis. Reaction time (RT) was used as a criterion for data screening. Only RT data falling in the range of 400 ms (after the target sound ends) and 3000 ms were used for analysis. Within each participant, RT exceeding 2 times of the standard deviation (S.D.) was replaced by the average RT of this participant. The calculation of average RT and S.D. and the replacement of outliers were all based on the data within the 400-3000 ms time limits. The P-P test revealed that a logarithm transformation of the RT data matched a normal distribution. Since normal distribution is a prerequisite for the following statistical analysis, the natural logarithm scores of the raw RT data (ln(RT)) were calculated.

A two-way repeated-measures ANOVA was conducted. One within-subject factor was the priming conditions, using the 11 priming types as its levels. The other within-subject factor was the ISI condition, having 2 levels (200 ms and 400 ms). ln(RT) served as dependent variables. Effect size (η^2) was calculated based on the sum of squares. Both by-subject and by-item analyses were performed.

2.4. Results

The overall accuracy in the judgment of word status was 90.7%. The accuracy of RW (real word) was 93.1%, and the accuracy of NW (non-word) was 86.7%. No participants had less than 80% accuracy. So the data of all participants were included. For RT (reaction time) data, only 4.96% that fell beyond the 400-3000 ms range were excluded. The averaged RT of all responses was 1039.62 (±447.48) ms, and the averaged RTs of RW and NW were 1018.79 (±447.58) ms and 1076.52 (±444.99) ms, respectively.

The by-subject repeated-measure ANOVA of ln(RT) revealed a significant interaction between ISI and priming conditions ($F(10, 210) = 3.162$, $p < .001$, $\eta^2 = .035$). Priming conditions had a main effect ($F(2.491, 52.311) = 8.359$, $p < .0005$, $\eta^2 = .201$), but not ISI. Differences between the 11 priming conditions were reported in the post-hoc pairwise comparisons. The direct priming condition showed significant priming effect, compared to the RW control condition ($p < .0005$). Both semantic priming conditions had significant priming effects ($p < .049$ with tone matching; $p < .033$ without phonological matching). Semantic priming with tone resemblance also showed significantly shorter RT than that of the pure tone priming condition ($p < .044$).

As for the NW conditions, tone priming had a significant priming effect ($p < .019$). The NW tone priming condition also had significantly shorter RT than that of the NW consonant ($p < .006$) or vowel priming conditions ($p < .008$). Post-hoc pairwise T-tests revealed significant differences between the two direct priming conditions under 200 ms and 400 ms ISIs

($p < .001$), the two semantic priming conditions with tone matching ($p < .013$), and the NW control conditions ($p < .005$).

Under the 200 ms ISI condition, direct priming ($p < .007$) and NW tone priming ($p < .009$) had significant priming effects. Semantic priming with tone matching had significant shorter RT than that of pure tone priming condition ($p < .009$). Under the 400 ms ISI condition, direct priming ($p < .0005$) and semantic priming without tone overlapping ($p < .023$) showed significant priming effects, whereas the NW consonant ($p < .032$) and vowel ($p < .007$) priming conditions had significant interference effects. RT of these two conditions was also much longer than that of the NW tone priming conditions ($p < .008$ for NW consonant priming; $p < .003$ for NW vowel priming). These results are shown in Figure 1.

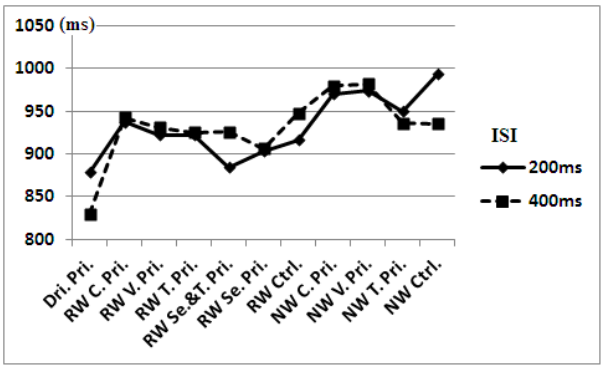

Figure 1: *RT of the eleven priming conditions under 2 ISIs (200 ms and 400 ms). RT over 2 times of S.D. was replaced by the average RT of each participant.*

The by-item analysis disclosed a significant interaction between ISI and priming conditions ($F(10, 200) = 2.841$, $p < .003$, $\eta^2 = .025$) and a main effect of priming conditions ($F(10, 200) = 5.330$, $p < .0005$, $\eta^2 = .163$). Post-hoc pairwise comparisons only revealed a significant direct priming effect ($p < .005$). Differences between the 2 ISIs under the direct priming condition ($p < .009$), semantic priming with tone matching ($p < .002$), and the NW control condition ($p < .008$) were confirmed. The direct priming effect was not significant under the 200 ms ISI, but under the 400 ms ISI ($p < .0005$).

Lacking correspondent effects in the by-item analysis, compared to the by-subject analysis, reflected a great variation between items. Further analysis of the priming conditions following the same primes revealed that there seemed to be a strong correlation between consonant duration and RT.

3. Discussion

This experiment explored the priming effects of initial consonants, vowels, and tones in Mandarin lexical processing. The priming effects were measured by comparing the priming conditions with the control conditions where there was nothing in common between primes and targets. Direct priming (repetition) condition served as a monitor, since it provided the strongest priming effect. The semantic priming effects were also examined with or without tone resemblance, to elaborate whether tone priming enhances semantic processing. Since the task was lexical decision, targets in the 4 out of 11 conditions belonged to non-words. The phonological priming effects were not only examined in real words, but also compared in the non-word phonological priming conditions and control conditions. The 2 ISI conditions were designed to reflect the stages of lexical access and lexical selection.

3.1. Tone and Segment Processing

Consistent with the previous research [6] on tone and segment priming, the current study showed that strong priming effects were presented when the segments were used as primes. In [6], the tone priming effect was absent. However, in our study, by exploring the phonological priming in real words and non-words separately, we found the tone priming effect for non-words especially under the 200 ms ISI condition. This effect did not show up under the 400 ms ISI condition. Instead, non-word consonant and vowel priming conditions showed interference at this stage. These effects suggest that real word primes inhibit the judgment of non-word targets that share initial consonants or vowels with primes. Such effect occurs at the stage of lexical selection.

Tonal information does not have such effect. But, under 200 ms ISI, this information can facilitate the rejection of non-words, if there is a matching in tones between primes and targets. This could also be due to the interference of the non-word control condition. 200 ms after the onset of real word primes is in the lexical access stage and the judgment of non-words may be inhibited. However, tones can affect this process by eliminating the interference.

The results in the non-word control conditions under 200 ms and 400 ms ISI are also significantly different, which further indicates that, at the stage of lexical access, the real word prime activates a lot of possible options, making the non-word judgment difficult, even if there is no interference from phonological matching between the non-word target and real word prime. Tone resemblance of the non-word target in the tone priming condition may narrow down the number of candidates in the lexical access, thus facilitating the lexical selection afterwards. Such facilitation can also shorten the RT, compared to the non-word control condition. A shorter RT of the semantic priming with tone matching than that of the pure tone priming condition under 200 ms ISI also supports that the available tonal information propels a lexical selection. In contrast, strong semantic priming effect without facilitation of tonal information only occurs under the 400 ms ISI condition. However, at the stage of lexical selection, comparing the RT of the NW tone priming and NW control conditions under 400 ms ISI, tone matching does not affect processing. At this stage, syllables with the same onset and medial should have been activated. In this experiment, target syllables in the segmental priming conditions only share initial consonants or vowels with the primes, but differ in all other aspects. So, introducing syllables starting with the same initial consonant or vowel but in different syllable structures and bearing different tones actually confounds lexical selection.

All these results indicate a similar role of initial consonants and vowels, but a different role of tones, during lexical access and lexical selection. It seems that lexical access before 200 ms is mainly initiated by segments, whereas tones help the access onward by narrowing down the possibilities. Soon after this stage, during lexical selection around 400 ms, a restricted number of syllables with the same syllable structure are selected. Additional tonal information will not greatly affect this process. However, syllables sharing partial segmental features will make lexical selection more difficult.

3.2. Temporal Issues

Our results also suggest a critical period of tone processing from 200 ms onward. Together with the available information of segments, the tonal information after 200 ms helps lexical selection, but the tonal information introduced after 400 ms from the syllable onset does not affect lexical processing. This is in accordance with the previous study, which shows that the tonal information in the nucleus of monosyllabic words is important for tone perception [12], and the tonal information in the second syllables in disyllabic words is more important for lexical selection [5].

The reason for lacking a difference between the priming effects of initial consonants and vowels may lie in the deviant critical time periods in which the processing of initial consonants and vowels is dissociated, e.g. less than 100 ms. But, since the focus of this study is lexical access and lexical selection, full information of the whole syllable is provided in the priming tasks.

In the 2 ISI conditions, we found 3 significant differences. The first one is the RT difference in the direct priming conditions, indicating that lexical selection after 400 ms from the onset of the prime has a stronger facilitation effect on the processing of the target than lexical access. The second one is a shorter RT under 200 ms ISI in the semantic priming condition with tone matching, indicating that the available tonal information at the stage of lexical access accelerates the following lexical processing. And the third one is the RT difference between the two NW control conditions, probably due to the interference of RW primes during lexical access.

3.3. Limitations and Future Directions

In order to address the issue of Mandarin sounds, this experiment included all types of consonants, vowels and tones. The inherent differences in consonantal durations may result in different RTs. One solution is to measure the detailed phonetic parameters, such as the durations of the initial consonant, vowel, and pitch onset time, and use them as covariance in the analysis. Another solution is to conduct new experiments with these parameters well controlled. The facilitation or interference effect shown in non-words does not have any correspondence in real words. This may be caused by the fact that the lexical processing of the target word removes the effect from the primes. The reason for this deserves further investigation.

4. Conclusions

This work explored the roles of Mandarin phoneme and toneme processing in lexical access and lexical selection, and identified comparable roles of initial consonants and vowels at the both stages. Tones are found to only facilitate lexical processing after the stage of lexical access, and this effect diminishes after the stage of lexical selection. Meanwhile, syllables starting with the same consonants or vowels but bearing different syllable structures and tones confound lexical selection. These results reflect not only the important roles of consonants and vowels, but also the critical period of tonal processing during lexical access and lexical selection.

5. References

[1] Caramazza, A., Chialant, D., Capasso, D. and Miceli, G., "Separable Processing of Consonants and Vowels", Nature, 403:428-430, 2000.

[2] Carreiras, M. and Price, C.J. "Brain Activation for Consonants and Vowels", Cerebral Cortex, 18:1727-1735, 2008.

[3] Bonatti, L.L., Pena, M., Nespor, M. and Mehler J. "Linguistic constraints on statistical computations: the role of consonants and vowels in continuous speech processing", Psychological Science, 16:451-459, 2004.

[4] McClelland, J. and Elman, J. "The TRACE model of speech perception", Cognitive Psychology, 18:1-86, 1986.

[5] Ye, Y. and Connine, C.M. "Processing Spoken Chinese: The Role of Tone Information", Language and Cognitive Processes, 14:609-630, 1999.

[6] Lee, C.-Y. "Does Horse Activate Mother? Processing Lexical Tone in Form Priming", Language and Speech, 50(1): 101-123, 2007.

[7] Levelt, W.J.M. "Spoken Word Production: A Theory of Lexical Access", PNAS, 98(23):13464-13471, 2001.

[8] Friederici, A.D. "Towards a Neural Basis of Auditory Sentence Processing", Trends in Cognitive Sciences, 6(2):78-84, 2002.

[9] Poss, N., Hung, T.-H. and Will, U. "The Effects of Tonal Information on Lexical Activation in Mandarin", Proceedings of NACCL-20(1):205-211, 2008.

[10] "Xian Dai Han Yu Chang Yong Zi Pin Du Tong Ji Biao", Guo Jia Yu Yan Wen Zi Gong Zuo Wei Yuan Hui Han Zi Chu (Eds.), 1989.

[11] Boersma, P. and Weenink, D. "Praat: doing phonetics by computer", [Computer program], version 5.1.29. Online: http://www.praat.org., accessed on 11 Mar 2010.

[12] Kaan, E., Barkley, C.M., Bao, M. and Wayland, R. "Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: an event-related potentials training study", BMC Neuroscience, 9:53, 2008.