The role of pitch contours in tonal languages processing

Udo Will and Nick Poss

Cognitive Ethnomusicology, School of Music
Ohio State University, Columbus, OH
{will.51 & poss.3}@osu.edu

Abstract

In the present study we examined the role of pitch contours in constraining lexical activation with three priming experiments in Mandarin. Primes based on pitch contours were found to significantly facilitate vocal responses in naming tasks but did not seem to affect lexical processing. Primes containing segments unrelated to the target, but with a tone that matched the target significantly delay responses and, presumably, lexical selection. The delay was found in a shadowing task as well as a lexical decision task. This is the first report of inhibitory lexical effects of tones when prime and target differ in form; it can be understood as a group activation of candidates with the same tone that increases competition with the representation of the target word.

1. Introduction

The role of tonal information in perceptual processing [1,2] and in spoken and visual word recognition of Chinese (e.g. [3-8]) has become a prominent issue in recent years. However, these studies in general did not provide clear answers to the questions of how tonal information is represented in the mental lexicon, how tonal information in speech input is mapped onto underlying lexical representations, to what extent tonal information is used to constrain lexical activation, and how lexical competition influences the processing of tonal information. In addition, the phenomenon of speech surrogates in tonal language areas suggests the possibility that pitch contours alone are able to influence lexical activation or identification [9]. The present study investigates some of the effects produced by pitch contours - in isolation as well as in combination with segmental information- in Mandarin speakers.

2. Experiments

In the following experiments we used a unimodal priming paradigm in which primes and targets are presented auditorily and in which we analyze how the tonal information of primes affects responses to targets in a shadowing task (experiment 1 & 2) and a lexical decision task (experiment 3).

2.1. Experiment 1

In the first experiment we separated processing of tonal information from that of segmental information by using primes that only contained tonal information. Vocal tone primes were constructed by extracting the pitch contour from Mandarin syllables spoken in the four tones and resynthesizing them with a formant spectrum corresponding to the resonance of the human vocal tract. A second set of primes were included to test the influence of timbre. These were generated by playing the four Mandarin pitch contours on the Chinese vertical flute, xiao. These instrumental tone primes

were adjusted to match the vocal primes in frequency range, dynamic and duration. A control prime was produced by extracting the amplitude envelop from a spoken syllable and convolving it with white noise. This resulted in a sound that had the dynamic and duration of spoken syllables but did not contain any tonal or segmental information. All primes had identical duration and mean amplitude.

Auditorily presented stimulus sequences consisted of a prime followed by a target with 250 ms SOA. Participants were asked to repeat the target sound as quickly and accurately as possible. The target stimuli consisted of 41 Mandarin words and 39 pronounceable nonwords, each appearing twice, once paired with a matching tone prime and once with the control prime; half of the matching tone primes were vocal primes the other half were instrumental primes. Trials were presented in two sessions with each target only occurring once in a session.

14 Native mandarin speakers participated in this experiment. Stimulus presentation was controlled by DMDX (K.I. Forster, Arizona State University). To account for variable target duration RTs were measured from the target sound offsets in the following analyses.

2.1.1. Results

Statistical analysis of the response latencies shows significant effects of the tone matching condition for both the words and non-words. In the by-subject anova we obtained $F(2,13)=21.16$; $p=0.0005$ for nonwords and $F(2,13)=21.06$; $p=0.0005$ for words. The mean RT difference between the matching and the control condition was -44.77 ms for nonwords and -35.89 ms for words. The by-item analysis resulted in $F(2,39)=51.62$; $p<0.0001$ for nonwords and $F(2,39)=42.72$; $p<0.0001$ for words, and the mean difference between matching and control condition was -41.84 ms and -35.67 ms for nonwords and words, respectively.

There was no significant RT difference between words and nonwords (F<1; mean difference 3.98 ms in the matching and -4.8 ms in the control condition).

Both prime types showed a significant effect. For instrumental primes we obtained $F(2,13)=12.82$; $p=0.0034$, and for vocal primes $F(2,13)=23.50$; $p=0.0003$. The difference from the effect of the control primes was -27.87 ms for instrumental and -50.21 ms for vocal primes.

2.1.2. Discussion

Tone contour primes show a significant facilitatory effect on vocal responses in the auditory naming task. The effect size appears to be the same for words and nonwords in both the by-subject and the by-item analysis. The two tonal prime types produced different effect sizes: in comparison with the
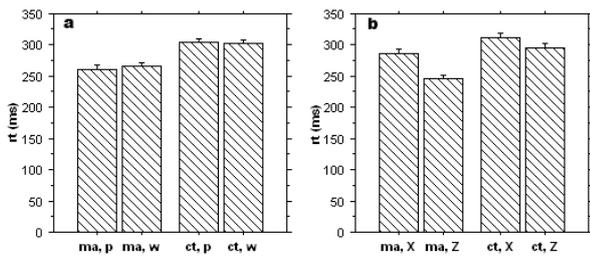
Fig.1: *Group means and standard errors for experiment 1. a: Words (w) and Pseudowords (p) in the tone match (ma) and control (ct) condition. b: Instrumental (X) and vocal (Z) primes in the tone match (ma) and control (ct) condition.*

control prime (white noise) vocal tone contours caused nearly double (50 ms) the effect of instrumental contours (28 ms). The observed facilitation does not seem to be a lexical effect as words and nonwords were not differentially affected. Furthermore, Poss & Will [9] recently reported that, in a comparable shadowing experiment that included also a nonmatching condition for the tone primes, there were no differences in facilitation between the match and nonmatch condition. A possible explanation for the observed effect might be that the tonal information of the primes acts directly on pre-motor areas involved in speaking [10] to produce this facilitation. Activations in these areas in passive listening task to music have recently been observed in a fMRI study [11].

2.2. Experiment 2

With tone primes only producing facilitation but no detectable lexical effect, we were curious to explore whether and how this changes when we add segmental information to the primes. In the following experiment we therefore used Mandarin words and pronounceable nonwords as primes. Target stimuli were 46 Mandarin words and 48 pronounceable nonwords. These were preceded (250 ms SOA) by the primes whose onset and rhyme segments were not related to the target segments (*i.e.*, no segmental match). Each prime/target pair was presented twice, once with matching and once with nonmatching tones. The prime target sequences were arranged in two lists so that no prime and target occurred more than once.

14 native mandarin speakers participated in the experiment. Stimulus presentation was controlled by DMDX and RTs were again calculated from the target sound offsets to account for variable target durations.

2.2.1. Results

Statistical analysis shows a significant effect of the tone matching condition for words but not for non-words. In the by-subject anova we obtained $F_{(2,13)}=5.33$; $p=0.038$ for words and $F_{(2,13)}=2.72$; $p=0.123$ for nonwords. The mean difference between the matching and nonmatching condition was 10.01 ms for words and 9.08 ms for nonwords.

The by-item analysis did not show a significant effect ($F_{(2,47)}=3.46$; $p=0.07$ for nonwords and $F_{(2,5)}=3.11$; $p=0.08$ for words. The mean difference between matching and non-matching condition was 9 ms and 8.6 ms for nonwords and words, respectively).

The mean RT difference between words and nonwords was 30.72 ms in the matching and 31.69 ms in the nonmatching condition.(w/ma 254.79ms; n/ma 285.52ms) and is highly significant.
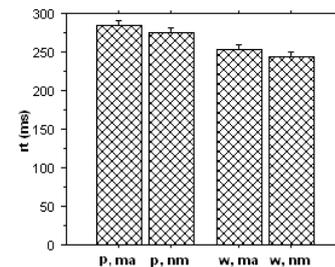

Fig.2: *Group means and standard errors for experiment 2. Words (w) and Pseudowords (p) in the tone match (ma) and non-match (nm) condition.*

2.2.2. Discussion

The tone match condition led to a significant delay of 10 ms of the vocal response compared to the nonmatch condition. This contrasts with findings of Poss & Will [9] who report no difference between the match/nonmatch condition in a shadowing task with tone primes (no segmental information). Although nonwords showed a comparable mean difference (9 ms) for the two conditions, this difference was not significant. Furthermore, responses to nonwords were on the average produced 31 ms later than that to words. This contrasts with the results from experiment 1 for which the overall RT difference between words and nonwords was 0.4ms. The addition of segmental information changes the effect of the primes on the target responses from a facilitatory to an inhibitory one. The fact that words and nonwords are affected differentially seems to indicate that the inhibition in the tone match condition involves lexical processes.

2.3. Experiment 3

The results from experiment 2 raise the question why previous studies using decision tasks have not identified a priming effect of matching tones. In principle this could be due to specifics of the task, subject response strategies, or differences in stimuli sets. To test this we performed the following experiment in which we used exactly the same stimuli sets as the previous experiment; the task, however was changed from shadowing to lexical decision: participants had to indicate via button presses whether the target sounds they were hearing were Mandarin words or not.

19 native Mandarin speakers participated in this experiment, 6 of which had also participated in experiment 2 (with 2 to 4 weeks between their two participations). This allowed for a direct comparison of their responses in the two experiments.

2.3.1. Results

Again, we found a significant effect in the tone matching condition for words but not for non-words. With the by-subject anova we obtained $F_{(2,18)}=7.64$; $p=0.013$ for words and $F_{(2,18)}=1.04$; $p=0.32$ for nonwords. The mean difference between the matching and the nonmatching condition was 17.38 ms for words and -10.76 ms for nonwords.

There was no significant difference between the overall reaction time for words and nonwords (-2.4 ms).

The by-item analysis did not show significant effects (F<1 for both words and nonwords).

The comparison of the reaction times with those of the shadow experiment 2 shows significant differences for the six subjects that participated in both. For the matching condition we get $F_{(2,5)}=8.48$; $p=0.03$; with a mean difference of 183.7 ms, and for the nonmatching condition we get $F_{(2,5)}=9.50$; $p=0.027$; with a mean difference of 185.4 ms. On the average RT is 184.5 ms longer in the decision task than in the shadowing task.
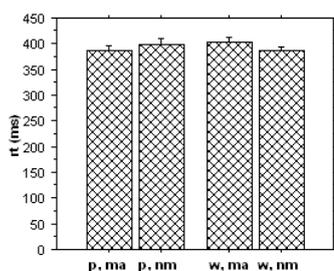

Fig.2: *Group means and standard errors for experiment 3 (decision task). Words (w) and Pseudowords (p) in the tone match (ma) and non-match (nm) condition.*

2.3.2. Discussion

In this decision experiment, as in the previous experiment, primes with matching tones (and unrelated segments) lead to an inhibition for the responses to words. The mean inhibition 17.4 ms is slightly larger than in the shadowing experiment (10 ms). Other results contrast with those from experiment 2, indicating task related effects: First, the response times for words and nonwords do not differ significantly. The 31 ms difference in experiment 2 may be due to the fact that nonwords, having no lexical entries, need to be prepared for

naming following the lexical search; this process is obviously not required in the decision task. Second, in experiment 2 RTs for nonwords in the tone match condition were longer than in the nonmatch condition. No such difference was found in the decision task, though the reason for this is not clear at present. Third, the by-item analysis is worse (F<1) than in experiment 2 (F=3.5 and 3.1 for nonwords and words respectively). It seems likely that the deterioration in the response to tonal information may be due to influences of postaccess processes like response strategies, because the decision task involves an explicit motor response (button presses). The involvement of postaccess processes is also indicated in the longer RTs of the decision experiment in comparison with those of the shadowing task.

3. General Discussion

The first experiment revealed a novel effect, the facilitation of vocal responses by the tone contour of primes. Due to the fact that no lexical involvement was identified, this finding may not be immediately relevant for the understanding of lexical processing of tonal information. However, this effect may well be involved in other language phenomena like e.g. speech surrogates in tonal language areas. In a comparative study on Mandarin, Hmong and English speakers, Poss & Will [9] have recently shown, that the facilitatory effect is present in all three groups, but only in tonal language speakers is there a stronger facilitation effect for voice contours than for instrumental contours; i.e. this effect is shaped by speakers language experience. Several studies provide evidence for cortical mechanisms activated more by vocal than by non-vocal stimuli (for review, see [12]). However, the observed effect is most likely related to the direct activation of pre-motor areas involved in speaking that have been observed following tonal stimulation [10, 11].

The combination of tonal and segmental information drastically changes the effects of primes on the target responses. In the case of tonal match (segments never matched in our experiments) we found a significant delay in the response. This inhibition may be due to the activation of groups of candidates with the same tone as the target that increases competition, resulting in a delayed response. This is the first time that a significant effect of tonal information in processing of Mandarin words has been identified. Interestingly, data of a recent study by Lee [8] seem to indicate the same effect, though it was not significant in his study. Lee used four types of primes, primes identical with the target (ST), with matching segments (S), with matching tone (T), and completely unrelated (U) primes. In all his experiments the T primes produced a longer reaction time than any of the other, though the differences were not significant. The reason for the different significances in his and our experiments may be due to combined effects of experiment design and experimental task. The comparison of experiment 2 and 3 above clearly reinforces concerns, already raised more than 20 years ago [13], that decision tasks may not be best suited to test effects of lexical access. It is conceivable that these concerns are more relevant when testing the apparently small (10 - 17 ms in the above experiments) influence of tonal information than in form priming.

The fact that we observed an inhibitory effect of matching tones clearly speaks against the idea that tonal information could act as or at the level of the phonological code in lexical

access. If the inhibition it is taken to reflect activation of a class or category of entries (with the same tone as the target), then tonal information acts more like the 'access code', though, in this case not, as proposed by Cutler [14] on the basis of her segmental priming studies, to reduce number of possible candidates, but to activate them.

4. Conclusions

In this study we have shown that tonal contour primes have a significant facilitatory effect on verbal production. Conversely, primes containing tonal as well as segmental information produce a small but significant inhibition if tones of primes match those of targets. We interpreted this as a reflection of increased lexical competition due to activation of candidates by the pitch contour of the primes.

This is a first study to describe such effects and it indicates some venues to further our understanding of the way in which tonal information influences lexical activation, In addition, our study also reiterates some questions about the usefulness of decision tasks in exploring these question, especially in connection with the influence of tonal information.

5. References

[1] Shen, X. S., & Lin, M. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech, 34,* 145-156.

[2] Whalen, D. H., & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica, 49,* 25-47.

[3] Cutler, A., & Chen, H. C. (1997). Lexical tone in Cantonese spoken-word processing. *Perception & Psychophysics, 59,* 165-179.

[4] Ye, Y., & Connine, C. M. (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes*, *14,* 609-630.

[5] Zhou, X. (2000). Phonology in lexical processing of Chinese: Priming tone neighbors. *Psychological Science*, *23*, 133-140.

[6] Zhou, X., & Marslen-Wilson, W. (1997). The abstractness of phonological representation in the Chinese mental lexicon. In H.- C. Chen (Ed). *Cognitive Processing of Chinese and other Asian languages* (pp. 3-26). Hong Kong: The Chinese University Press.

[7] Yip, M. 2001. Phonological priming in Cantonese spoken-word processing. *Psychologia*, 44, 223-229.

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

[9] Poss, N. ; Will, U. 2007. The role of pitch in the processing of tonal languages: new evidence on the connection between music and language. Paper presentation at the SEM symposium 'New Directions in Cognitive Ethnomusicology', Columbus, OH.

[10] Poeppel, D. & Hickok, G. 2004 Towards a new functional anatomy of language. *Cognition* 92, 1-12.

[11] Koelsch, S., Fritz, T., v.Cramon, Y., Mueller, K., and Friederici, A.D 2006 Investigation Emotion With Music. An fMRI Study. *Human Brain Mapping* 27, 239-250.\

[12] Belin, P., Fecteau, S. and Bedard, C. 2004 Thinking the voice: neural correlates of voice perception. Trends in Cognitive Sciences, 8(3), 129-135.

[13] Balota, D.A. and Chumbley, J.I. 1984 Are Lexical Decisions a Good Measure of Lexical Access? The role of Word Frequency in the Neglected Decision Stage. *Journal of Experimental Psychology: Human Perception and Performance*, 10(3), 340-375.

[14] Cutler, A. 1986 Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech,* 29, 201-220.