# THE ROLES OF TONAL AND SEGMENTAL INFORMATION IN SPOKEN WORD RECOGNITION FOR L2 SPEAKERS: EVIDENCE FROM DUTCH LEARNERS OF MANDARIN

Ting Zou [1], Yiya Chen [2,3]
[1] Department of English, Beijing Foreign Studies University
[2] Leiden University Center for Linguistics, [3] Leiden Institute for Brain and Cognition
tingzou@bfsu.edu.cn, yiya.chen@hum.leidenuniv.nl

## ABSTRACT

An eye-tracking experiment with the visual world paradigm was conducted to investigate the role of segmental and tonal information during on-line lexical activation by native Mandarin speakers and Dutch learners of Mandarin with different levels of proficiency. The results suggest that native Mandarin speakers use tonal information effectively to constrain lexical activation at an early stage, similar to the way they exploit segmental information. Dutch learners cannot use tonal information for word recognition as effectively as native Mandarin speakers. A clear developmental path, however, was observed when we compared the results of advanced learners (who are more native-like) to that of the beginners (who made much less use of tonal information).

**Keywords**: tone perception, lexical activation, L2 learners of Mandarin, eye-tracking

## 1. INTRODUCTION

### 1.1 Tone processing by native Mandarin speakers

In Mandarin Chinese, tones are used to distinguish lexical meanings, and the role of tone in spoken word recognition by native speakers has received much attention in previous studies. Studies using online measures such as eye-tracking and event related brain potentials investigated the on-line utilization of tonal information during the time course of spoken word recognition. Parallel processing of segments and tones was found in these studies, suggesting the tones can be used to constrain lexical access and the role of tonal information is comparable to that of segmental information [5,6,9].

### 1.2. Tone processing by non-tone language speakers

There have been an increasing number of studies that have tested the perception and production of lexical tone in beginning Mandarin L2 learners [3, 10] as well as the phonetic learning of new tonal categories by speakers without prior tone language learning experience [1, 12]. Although the research questions varied across these studies, their results led to a convergent finding that naive non-native speakers of Mandarin can gain significant improvement in tonal identification and discrimination with a proper amount of training and can learn to use tones in a lexically contrastive way. Some studies also found a training-induced change in participants' neural system [4]. It is important to note that most of the studies mentioned above focus on the learning of lexical tones by naive non-native Mandarin speakers and beginning learners of Mandarin, the performance of advanced second language (L2) learners and the developmental trajectory during the time course of tone acquisition have not been studied systematically.

### 1.3. The present study

To fill the knowledge gap, an eye-tracking experiment with the visual world paradigm (VWP) (with pictures as visual stimuli) was designed. Within each trial, an auditory stimulus was played while four pictures were presented on the computer screen. Participants were asked to perform an auditory-visual picture matching task. With this paradigm, we aimed to test (1) the relative role of segmental and tonal information in lexical activation and selection by native speakers and Dutch learners of Mandarin, and (2) the developmental trajectory for Dutch learners of Mandarin in using segmental and tonal information effectively for spoken word recognition.

## 2. Method

### 2.1. Participants

15 Mandarin control participants and 26 Dutch learners of Mandarin participated in the experiment. The native Mandarin control comprised of 5 males and 10 females (mean age = 26.9, SD = 2.5). All were born and grew up in the Northern part of China and spoke Standard Chinese on a daily basis. All Dutch learners of Mandarin received formal Chinese training from a BA Chinese Studies program. The beginners had never lived in China (4 males, 7 females; mean age = 20.1, SD = 2.3). The other 15 participants (4 males and 11 females; mean age = 23.5, SD = 3.0) were advanced Mandarin learners, who had Mandarin experience between 3 and 14

years (mean = 4.9, SD = 2.6), and had spent at least one year in China.

## 2.2. Material

The stimuli in the eye-tracking experiment consisted of 12 sets of monosyllabic Mandarin words. Each set includes one critical word e.g., *chuang1* 'window' and four competitors similar to the design in [5]. The segmental competitor shared all phonemes with the critical word, but differed in tone e.g., *chuang2* 'bed'; the cohort competitor shared the initial consonant and tone, but differed in the rest of the syllable e.g., *che1* 'car'; the rhyme competitor shared the rime and tone, but differed in initial consonant, e.g., *guang1* 'light'; the tonal competitor shared the tone, but differed in all phonemes, e.g., *ji1* 'chicken'. Since the contribution of tonal information in word recognition is one of the main issues we aimed to examine, the stimuli (target words vs. segmental competitors) in the current study included all 12 possible tone contrasts.

## 2.3. Procedure

Participants were tested individually in a sound-attenuated room. Before the eye-tracking recording, there was a familiarization session to ensure that participants knew the names of all picture stimuli. In the test session, participants performed an auditory-visual picture matching task and the eye movements were recorded via an Eyelink 1000 device with a 16mm lens. The sampling rate was set at 500-Hz. On each trial, participants were presented with a four-picture display, which contained a target, a phonological competitor (one of the four competitor conditions: segmental, cohort, rhyme and tone), and two phonologically unrelated distractors. We also included a baseline condition which included the target and three distractors. At the beginning of each trial, a fixation cross ("+") appeared in the centre of the screen for 500 ms. Then, a four-picture display was presented on the screen. Participants were instructed to first look at the fixation cross and then select the word they heard by mouse clicking on the corresponding picture. After the response, the next trial proceeded after a 1000-ms pause.

## 2.4. Data analysis

Within each trial, the visual display was divided according to the areas of the four items (i.e., target, competitor and distractors), and only looks within the corresponding areas were included in analyses. The reciprocal trials and the trials with incorrect responses were excluded from data analysis. Eye movement data was analyzed from 200 ms to 1400 ms after stimulus onset. The data recorded at 2-ms intervals was resampled so that the proportion of looks was calculated every 16 ms (62.5 Hz), following the sampling rate used in [5], for ease of comparison.

The statistical analysis of the eye movement data curves was conducted with growth curve analysis (GCA) [7,8] with linear mixed modeling in R. The changing of the gaze distribution probability over time was fitted using four-order orthogonal polynomials. The intercept term indicates the average overall fixation proportion; the linear term indicates a monotonic change in the general direction of the curve, while the quadratic, cubic and quartic terms reflect the details of the steepness of the curve [5, 7].

## 3. Results

For the fixation data of three groups of participants in four conditions, linear mixed models were built for both target looks and competitor looks. Both models had Time components (up to the fourth-order component), Participant Group, Competitor Condition and their interactions as fixed effects. By-subject and by-item intercepts were entered as random effects. Post-hoc comparisons of the interactions of the three factors (Participant Group, Condition and Time components) were conducted using the *glht* function with Bonferroni adjustment in the Multcomp package in R.

The statistical results of the fixed effects are presented in Appendix 1. For each participant group, results of the target looks are discussed first, followed by the results of competitor looks. Since we were concerned with the degree of activation of different types of competitors, the curves of looks to target/competitor in each competitor condition were compared with baseline. Therefore, in the following, the post-hoc results of time components between a particular competitor type and the baseline curve are reported in details.

For native Mandarin speakers, the proportion of looks to target as a function of time gradually ramped up and reached the maximum of approximately 90% around 1100ms in all conditions, exhibiting a sigmoidal curve. Specifically, the proportion of looks to target in the segmental condition (**Fig.1*a***, *chuang1 vs. chuang2*) showed a slightly delayed raising pattern compared to baseline (ca. 500-700ms), indicated only by the significant difference of quadratic component (Est. = 0.10, $z$ = 3.58, $p < 0.05$). The proportion of looks to the competitor exhibited a slightly higher peak compared to baseline (at ca. 400ms), as suggested by a significant difference in intercept (Est. = 0.01; $z$ =

6.85, $p < 0.001$). The small delay in target looks and the slightly higher proportion of looks to the competitor compared to baseline suggests a marginal competition effect for segmental competitors. **Fig.1*b*** compares the cohort (*chuang1 vs. che1*) and the baseline conditions. The looks to target in the cohort condition differed significantly from the baseline condition in the cubic component (Est. = −0.16, $z = -5.65$, $p < 0.001$). The looks to competitor, however, did not differ significantly between cohort and baseline conditions, indicating that the participants did not launch more looks to cohort competitors than to the baseline condition. The rhyme condition (**Fig.1*c***, *chuang1* vs. *guang1*) showed a lower proportion of looks to the target than the baseline condition, as indexed by a significant difference in the mean height (Est. = −0.02, $z = -5.22$, $p < 0.001$). The looks to the rhyme competitors showed a significantly higher peak (at ca. 400ms) than baseline (for overall height, the linear and the quadratic component, all $p$ values < 0.001), indicating that the rhyme competitors were activated effectively to complete with the targets. In the tonal condition (**Fig.1*d***, *chuang1* vs. *ji1*), looks to target were not significantly different from baseline in any aspect. Looks to competitor only differed significantly from baseline in the quadratic (Est. = 0.07, $z = 4.03$, $p < 0.01$) component, which indicated that tonal competitors were not activated adequately.

Compared to native speakers, the beginners generally showed a less proportion of looks to the target and more looks to the competitor in all conditions. **Fig.2*a*** shows that in the segmental condition, there was a slightly less proportion of looks to the target than baseline (ca.200-400ms and 800-1200ms), which was reflected by a marginally significant intercept difference (Est. = −0.01, $z = -3.18$, $p = 0.07$). There is a greater rate of looks to competitor than baseline, confirmed by the significantly different overall height and quadratic component (all p values < 0.001). The proportion of looks to segmental competitor remained high (about 20%) during the whole trial, reflecting the difficulty in distinguishing minimal tone pairs. As shown in **Fig.2*b***, the target looks in the cohort condition increased faster than baseline before 800ms, indicated by a significantly different cubic component (Est. = 0.16, $z = 5.15$, $p < 0.001$). No significant difference in competitor looks was found between the cohort and baseline conditions, suggesting that the cohort competitor was not activated adequately for competition. In **Fig.2*c***, there was clearly a smaller proportion of looks to the target in the rhyme condition than baseline, which is supported by significant overall height, quadratic, and cubic components (all p values < 0.001). There was also a greater proportion of looks to the competitor in the rhyme condition than baseline, confirmed by the significant overall height, linear and quartic components (all p values < 0.001). These differences suggest that like native speakers, rhyme competitors were activated adequately as potential candidates for word recognition. The proportion of looks to target in the tonal condition (**Fig.2*d***) showed an early and lower peak compared to the baseline curve, as reflected by significant differences in the linear, quadratic and cubic components (all p values < 0.001). The proportion of looks to the competitor remained above 10% in the tonal and the baseline conditions for the whole process, with a significant difference only in the linear component (Est. = −0.09, $z = -4.74$, $p < 0.001$).

Compared to the beginners, the advanced learners generally exhibited an improvement towards the native direction. **Fig.3*a*** shows that for advanced learners, there was a smaller proportion of looks to the target in the segmental condition than baseline, suggested by significant differences in the overall height, the linear component and the quadratic component (all $p$ values < 0.001). The fixation curve of the segmental competitor was higher than baseline, showing significant differences in the overall height, linear component, quadratic component and quartic component (all $p$ values < 0.05). The lower rate of looks to target and higher looks to competitor conjointly indicate a strong competition effect for segmental competitor.

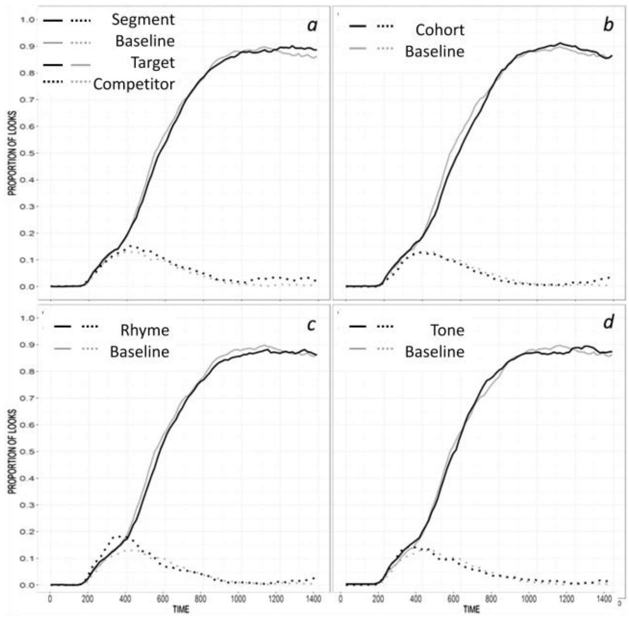


**Figure 1:** Mean proportion of looks to the target (solid line) and the competitor (dotted line) in different conditions for native Mandarin speakers. Different conditions were plotted against the baseline condition for comparison.

**Fig.3*b*** shows a smaller proportion of looks to target in the cohort condition than baseline during the whole process, reflected in their significantly different overall height (Est. = −0.03, z = −9.71, $p < 0.001$).

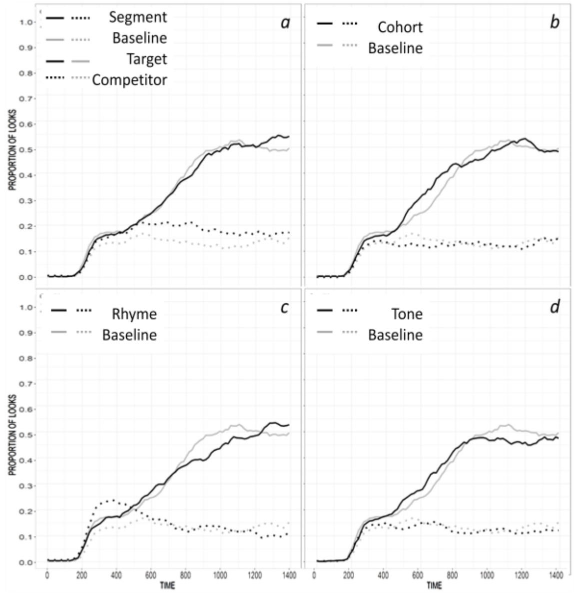


**Figure 2:** Mean proportion of looks to the target (solid line) and the competitor (dotted line) for beginning learners of Mandarin.

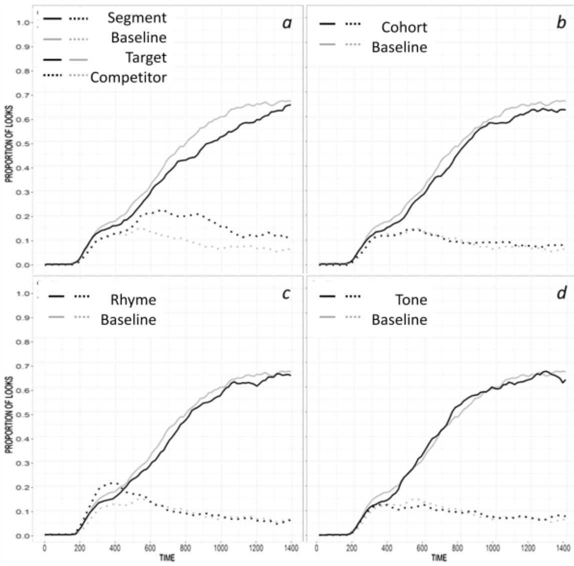


**Figure 3:** Mean proportion of looks to the target (solid line) and the competitor (dotted line) for advanced learners of Mandarin.

The curve for the cohort competitor did not differ significantly in any aspect compared to baseline. **Fig.3*c*** shows that during the whole process, the proportion of looks to the target in the rhyme condition was smaller than baseline. The proportion of looks to the competitor showed a higher and earlier peak compared to baseline, confirmed by significant differences in the overall height, linear component, and quadratic components (all p values $< 0.001$). The smaller proportion of looks to the target and greater proportion of looks to the competitor indicate a strong competition effect of the rhyme competitor. For the tonal condition (**Fig.3*d***), the proportion of looks to target was slightly lower than baseline around 200-400ms, as indexed by a significant difference in the quadratic component (Est. = −0.10, z = −3.69, $p < 0.05$). The curve of the tonal competitor was not significantly different from baseline in any aspect, which indicates that for advanced learners, a competitor with only tonal overlap cannot be activated effectively and compete with the target.

4. Conclusion and discussion

The results showed that, for native speakers, as the input unfolded (e.g., *chuang1*), the rhyme competitor (e.g., *guang1*) was activated adequately to compete with the target. The segmental (e.g., *chuang2*), cohort (e.g., *che1*) and tonal (e.g., *ji1*) competitors were not activated adequately. According to the TRACE model, the global goodness of fit between the input and the underlying representation will determine the degree of activation of the potential candidates. Therefore, the adequate activation of the rhyme competitors emphasizes that the combination of rhyme and tonal information plays an important role in lexical access in Mandarin Chinese. This finding is also in line with the results of [2] and [11]. The lack of activation of the segmental competitor suggests that native Mandarin speakers can use tonal information effectively to constrain lexical access in an early stage.

Compared to native Mandarin speakers, non-native learners generally showed fewer fixations to the target and more looks to the competitor in all conditions, indicating the target was not fully activated and the competitor was not fully suppressed. For both groups of learners, the rhyme competitor was activated most to compete with the target. The segmental competitor was also highly activated, indicating that tonal information cannot be used effectively to inhibit incompatible candidates. The origin of their inability of using tonal information lies in their difficulty in distinguishing tonal minimal pairs, as indexed by a late high plateau in the curve of the segmental competitor compared to baseline. Similar to native speakers, both cohort and tonal competitors were activated to a less extent for learners. Moreover, a clear developmental path was observed. Compared to the beginners, the advanced learners could activate the correct targets and suppress the competitors more effectively, showing a significant improvement towards native-like tonal perception and spoken word recognition.

# 5. REFERENCES


[1] Chandrasekaran, B., Sampath, P. D., & Wong, P. C. (2010). Individual variability in cue-weighting and lexical tone learning. *Journal of the Acoustical Society of America*, *128*(1), 456-465.

[2] Choi, W., Tong, X., Gu, F., Tong, X., & Wong, L. (2017). On the early neural perceptual integrality of tones and vowels. *Journal of Neurolinguistics*, *41*, 11-23.

[3] Hao, Y. C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, *40*(2), 269-279.

[4] Lu, S., Wayland, R., & Kaan, E. (2015). Effects of production training and perception training on lexical tone perception – A behavioral and ERP study. *Brain Research*, *1624*, 28-44.

[5] Malins, J. G., & Joanisse, M. 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.

[6] Malins, J. G., & Joanisse, M. F. (2012). Setting the tone: An ERP investigation of the influences of phonological similarity on spoken word recognition in Mandarin Chinese. *Neuropsychologia*, *50*(8), 2032-2043.

[7] Mirman, D. (2014). Growth curve analysis and visualization using R. Boca Raton, FL: CRC Press.

[8] Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. Journal of Memory and Language, 59(4), 475-494.

[9] Schirmer, A., Tang, S. L., Penney, T. B., Gunter, T. C., & Chen, H. C. (2005). Brain responses to segmentally and tonally induced semantic violations in Cantonese. *Journal of Cognitive Neuroscience*, *17*(1), 1-12.

[10] Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America*, *106*(6), 3649-3658.

[11] Wiener, S., & Turnbull, R. (2016). Constraints of tones, vowels and consonants on lexical selection in Mandarin Chinese. *Language and speech*, *59*(1), 59-82.

[12] Wong, P. C., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, *28*(4), 565-585.


**Appendix 1.** Summary of fixed effects for the models of Looks to target and Looks to competitor.

| Fixed effects | Looks to target | | | Looks to competitor | | |
|---|---|---|---|---|---|---|
| | df | $\chi^2$ | *p* | df | $\chi^2$ | *p* |
| Linear component | 1 | 46869.0 | <.001 | 1 | 122.0 | <.001 |
| Quadratic component | 1 | 3974.8 | <.001 | 1 | 6377.2 | <.001 |
| Cubic component | 1 | 3541.8 | <.001 | 1 | 4201.5 | <.001 |
| Quartic component | 1 | 586.2 | <.001 | 1 | 36.0 | <.001 |
| Group | 2 | 51.8 | <.001 | 2 | 33.7 | <.001 |
| Condition | 4 | 272.3 | <.001 | 4 | 1778.6 | <.001 |
| Group: Condition | 8 | 211.8 | <.001 | 8 | 512.4 | <.001 |
| Group: Time components | 8 | 13042.0 | <.001 | 8 | 2224.8 | <.001 |
| Conditon: Time components | 16 | 244.7 | <.001 | 16 | 1735.5 | <.001 |
| Group: Condition: Time components | 32 | 257.6 | <.001 | 32 | 543.1 | <.001 |