THE PERCEPTION OF LEXICAL TONES AND TONE SANDHI IN L2: SUCCESS OR FAILURE?

Ao Chen & René Kager

Utrecht Institute of Linguistics, Utrecht University, the Netherlands
a.chen@uu.nl; r.w.j.kager@uu.nl

ABSTRACT

In the current study, we explore the issues of the acquisition of lexical tones and the acquisition of tonal phonological rules by means of categorical perception experiments with advanced L2 learners of Mandarin whose native language is Dutch. By comparing their performance with native Mandarin listeners and native Dutch listeners, we found that advanced L2 learners may build up native-like representations of lexical tones, and show native-like perception of tone sequences vis-a-vis the tone sandhi rule of Mandarin. However, learners only succeed if the task is simple; when it becomes more complicated, they fail to approach native-like perception, and even lose their original sensitivity to the subtle acoustical differences of pitch change.

Keywords: lexical tone acquisition, categorical perception, T3 Sandhi

1. INTRODUCTION

It has been shown that lexical tones are very difficult for non-native listeners to acquire Kiriloff [5], Shen [6]. However, it is not clear whether poor production is due to failure of motoric control of the articulators or due to failure to perceive tones in a native-like way. Moreover, the acquisition of phonological grammars for tones has been largely neglected. In order to obtain a clearer picture of the acquisition of lexical tones, we run perception experiments involving Mandarin lexical tones as well as the Mandarin tone sandhi rule with advanced learners of Mandarin whose native language is Dutch, a non-tone language (henceforth, AL), and we compare their performance with both native Mandarin listeners (henceforth, CN) and native Dutch listeners without knowledge of Mandarin (henceforth, NL).

In Mandarin, lexical tones are realized mainly by changes of F_0, and are used in a phonemic way to differentiate meaning. There are four lexical tones, namely high level tone (T1), low rising tone (T2), low dipping tone (T3), and high falling tone (T4). Of the Mandarin tone contrasts, that between T2 and T3 is claimed to be the most difficult one for L2 learners to acquire Huang [3]. In addition, in Mandarin, T2 and T3 are considered only partially contrastive Hume [4], as when two T3s are juxtaposed, the T3 sandhi rule requires the first T3 to change to a T2, and the sandhied T3 has the same surface form as an underlying T2. Specifically, the T3 sandhi rule is positionally asymmetrical: only a T2-T3 sequence causes lexical ambiguity, as the T2 could either be an underlying T2 or a sandhied T3, whereas a T3-T2 sequence is lexically unambiguous.

The Categorical Perception (henceforth, CP) paradigm is adopted in the current study. CP assumes that native listeners have formed phonologically contrastive categories and are able to ignore within-category phonetic variations while being sensitive to cross-boundary differences. Pitch change, the main phonetic correlate of lexical tone, is used in a phonemic way to differentiate meaning in Mandarin at a lexical level, while being used in non-tone languages for intonational or pragmatic purposes. Accordingly, only native listeners of tone languages perceive lexical tones categorically while non-native listeners process pitch change in a psycho-acoustical fashion in CP experiments Francis [1], Hallé [2]. In the current study, we test the perception of the T2-T3 contrast of AL, CN, and NL, using an eight-step continuum between end point T2 and end point T3. Listeners participate in three tasks: a forced choice identification task, an AX discrimination task and an AXB discrimination task. We adopted the two-step difference discrimination paradigm of [2] in AX and AXB. Crucially, we take into consideration effects of order of presentation of the stimuli on discrimination, which would reflect the influence of T3 Sandhi on processing, i.e. if participants had knowledge of T3 Sandhi and were able to use it in online processing, they would discriminate more accurately when presented with pairs in which a stimulus that is acoustically close to T3 precedes a stimulus close to T2, than pairs with the reversed order.

2. EXPERIMENTS

2.1. Forced choice identification experiment

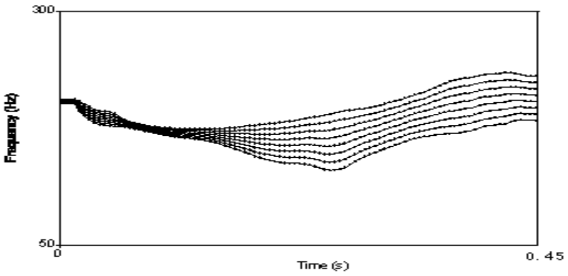
2.1.1. Participants

20 CN, 20 NL, and 17 AL were recruited for the experiment. CN are native Mandarin Chinese graduate students. NL are native Dutch graduate students without prior exposure to Mandarin or other tone languages. AL are native Dutch students in the third year of their BA majoring in Chinese language and culture. All AL have been to China for at least half a year. They also report themselves to have no problem using Mandarin in daily life, and to be able to understand their native Mandarin teacher easily during classes.

2.1.2. Stimuli

Endpoint T2 and T3 were realized on the carrier syllable /ma/, spoken by a female native speaker of Mandarin. First, pitch contours of these two tokens were extracted and time normalized; then six steps were interpolated between the normalized pitch contours. Finally six /ma/ syllables each carrying one in-between pitch contour were re-synthesized respectively. All together there were eight experimental stimuli whose pitch contours are given in Fig. 1, which are named step 1 to step 8 from up to bottom.

Figure 1: Stimuli continuum between /ma2/ and /ma3/.



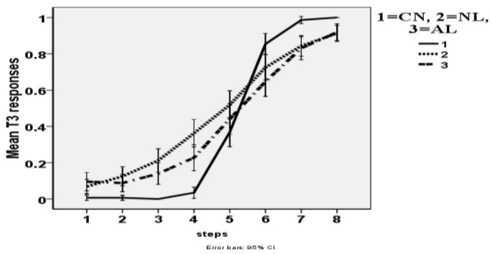
2.1.3. Procedure

CN and AL in each trial heard a single stimulus and were asked to choose between one of the two characters which represent /ma2/ and /ma3/ respectively. For NL, as they have no knowledge of Mandarin, we adopted the AXB identification paradigm, such that A is always an endpoint T2, B is always an endpoint T3, while X varies along the eight steps. NL were asked to indicate whether X is the same as the first syllable or same as the third syllable.

2.1.4. Results and discussion

A univariate ANOVA shows a significant difference among the responses of the three language groups: $F(2) = 5.505$, $p<0.005$. As shown in Fig. 2, for CN, there is an abrupt shift from T2 responses to T3 responses from step 4 to step 6; the closer to the endpoint, the more consistent the responses are, which is evidence of categorical perception of tones. For CN, steps 1-5 are identified as T2 while steps 6-8 are identified as T3. In contrast, NL yields a smooth identification curve along the continuum, and the variation among listeners' judgment is equal across steps, which serves as evidence that NL process tonal variation on a psycho-acoustical basis. AL approach the performance of CN in the way that their identification curve shows a steeper slope than NL, and same as for CN, their identification shifts from T2 to T3 at step 6. However, for the closer to endpoint steps, AL fail to show consistency as CN do. The performance of AL suggests that advanced learners do establish two categories of T2 and T3, however, the categorization is fuzzier than native listeners.

Figure 2: Mean T3 responses by CN, NL, and AL.



2.2. AX discrimination experiment

2.2.1. Participants

The same participants as in the forced choice identification experiment were recruited for the current task.

2.2.2. Stimuli

The same eight steps as in the forced choice identification experiment were used as stimuli for the current task.

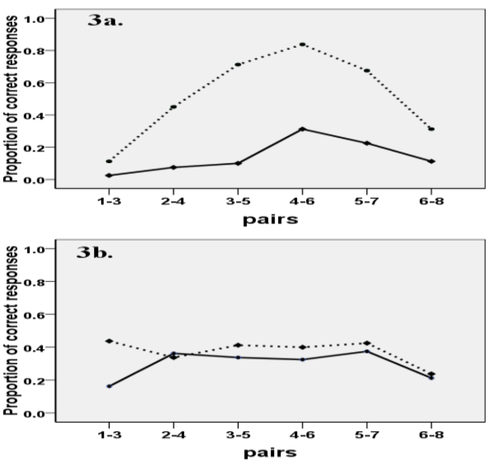
2.2.3. Procedure

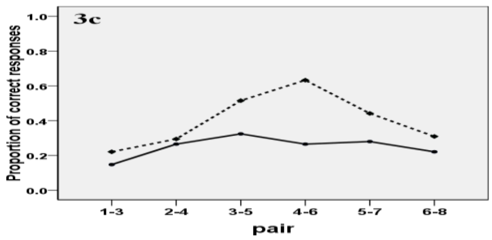
Participants listened to pairs of stimuli and were asked to tell whether the two stimuli are the same or different. Importantly, both ascending orders such as step 1-3 and step 2-4 and descending orders such as step 3-1 and step 4-2 occurred.

2.2.4. Results and discussion

In order to shed light on the processing of T3 Sandhi, we look at the responses for ascending order and descending order separately. Within each language group, a repeated measures ANOVA shows that both order and pair have a significant effect for the proportion of correct responses (CN: F_{pair} (5, 75) = 57.901, p<0.001, F_{order} (2,78) = 2.448, p<0.001; NL: F_{pair} (5, 75) = 3.471, p = 0.007, F_{order} (2, 78) = 6.311, p= 0.014). Fig. 3 gives the proportion of correct responses of the three language groups for each pair with orders separated. Unsurprisingly, Mandarin listeners discriminate much better if the first stimulus in the test pair is closer to T3 and second one is closer to T2, which could be explained by the asymmetry of T3 Sandhi. NL performed poorly (below chance accuracy) along the continuum, which implies that discrimination of T2 and T3 is indeed difficult. However, unexpectedly, without any knowledge of Mandarin, NL shows the same significant effect of order. Although the order effect mainly comes from a single pair (1-3 v.s. 3-1), NL yielded higher accuracy for descending orders across the pairs consistently. Moreover, pair and order only significantly interact for CN (F(10, 70) = 14.667, p<0.001), but not for NL (F(10, 70) = 2.209, p>0.05). In other words, CN benefit most in discrimination when both a tonal contrast is present and the order of presentation is unconfusing. NL, on the other hand, are only facilitated by order information as they lack separate representations of T2 and T3.

Figure 3: Proportion of correct responses by CN, NL and AL in AX task. Panel 3a represents that of CN, panel 3b represents that of NL, and panel 3c represents that of AL. Dotted line represents responses of descending order, and solid line represents responses of ascending order.





For AL, as for CN, both pair and order have a significant effect for the proportion of correct responses, and pair and order interact significantly (F_{pair}(5, 63) = 6.432, p=0.000; F_{order}(1, 67) = 19.18, p=0.000, $F_{pair*order}$(5, 63) = 3.2, p=0.012). Moreover, AL perform fairly similar to CN: A univariate ANOVA shows no significant difference between these two groups (F(11)=0.004, p>0.5). AL discriminate better if first token in the stimuli pair is closer to T3 and the order effect is also exaggerated when the stimuli pair straddle tonal category boundary. These results indicate that advanced L2 learners are able to build up representations of lexical tones, and to use the phonological grammar in a native like way for processing.

2.3. AXB discrimination experiment

2.3.1. Participants

The same participants as in the previous experiments took part in the current task.

2.3.2. Stimuli

The same eight steps as in previous experiments are used as stimuli in the current task.

2.3.3. Procedure

Participants listened to AXB triplets in which A and B were always two-step apart, and they were asked to indicate whether X is the same as A or same as B. Importantly, all four combinations, namely AAB (e.g. step 1-1-3), ABB (e.g. step 1-3-3), BAA (e.g. step 3-1-1) and BBA (e.g. step 3-3-1) occurred in the experiment.

2.3.4. Results and discussion

Similarly to the AX discrimination experiment, we first look at the responses of the four combinations separately within each language group. Fig. 4 gives the proportion of correct responses of the three language groups for each pair with combinations separated. For CN, a repeated measures ANOVA shows that both pair and combination have a significant effect for the proportion of correct responses (F_{pair} (5, 75) = 15.690, p<0.001, $F_{combination}$(3, 77) = 27.841, p<0.001). In contrast, for NL, only combination is a significant factor ($F_{combination}$(3, 77) = 11.212, p>0.001; F_{pair}(5, 75) = 0.577, p>0.05).

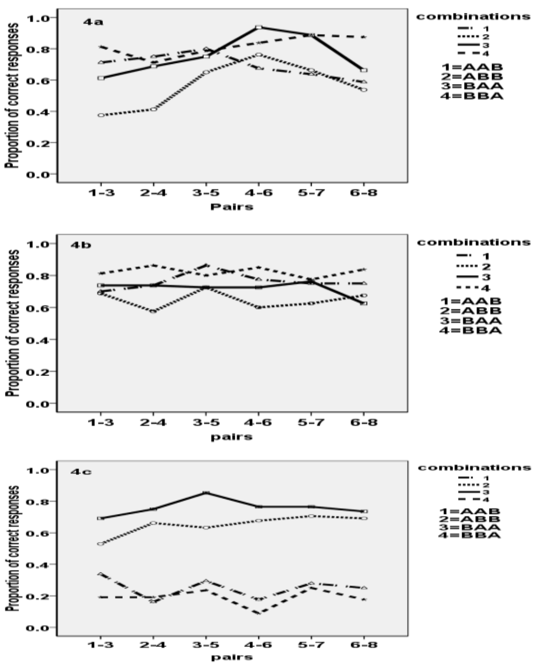


Figure 4: Proportion of correct responses by CN, NL and AL in AXB task. Panel 4a represents that of CN, panel 4b represents that of NL, and panel 4c represents that of AL.

For CN, the best discrimination is reached at pair 4-6 in the BAA combination, in which the triplet straddles the tonal boundary and the order is unconfusing. Meanwhile, the response curves of BAA and ABB are fairly parallel; a higher accuracy rate is obtained for BAA, as ABB could be misperceived as BBB due to influence of T3 Sandhi. The interference of T3 Sandhi and the lack of boundary information make the discrimination of the first two pairs in the ABB combination the worst. Also, the discrimination for BBA is fairly successful (above 70% correct) as two consecutive T3's are unnatural for native listeners and may cause extra attention. In comparison, for NL, whose performance is mainly psycho-acoustically based, the accuracy rate of responses could be ranked as BBA>AAB>BAA>ABB. We interpret this result as follows. First, due to the lower demand of memory load, NL discriminate better if the first two rather than the last two stimuli in the triplets are identical. Second, when memory load is the same, discrimination is easier if B precedes A than vice versa. NL discriminated much better in AXB than in AX, possibly due to a learning effect, but consistently with the results of the AX task, the order bias in discrimination by NL suggests that there might be universal bias in the perception of pitch change, such that a T2-T3 sequence tends to be naturally misperceived as identical.

For AL, as for CN, both pair and combination have significant effects on the proportion of correct responses (F_{pair}(5, 59) = 2.691, p=0.029; $F_{combinations}$(3, 61) = 97.503, p=0.000). As can be seen in Fig. 3c, on the one hand, the discrimination of AAB and BBA are below chance, regardless of which pair they are, and a post-hoc analysis reveals no significant difference between these two combinations (p>0.05). On the other hand, the discrimination of ABB and BAA is much more successful, and among these, the accuracy rate of BAA is significantly better than that of ABB (post-hoc p=0.000). AL and CN behave similarly in that discrimination is more accurate if a closer to T3 stimulus precedes a closer to T2 stimulus. However, compared to CN, AL fail to show any discrimination peak for across boundary triplets. Moreover, compared to NL, the accuracy rate for ABB and BAA of AL is not higher while that for AAB and BBA is even much lower than NL. It seems that L2 learners, on their way towards mastering the tonal grammar, are partly able to inhibit their psycho-acoustical driven processing, but at the same time, when the task becomes more complicated, fail to approach the native way of processing tonal information categorically.

3. REFERENCES

[1] Francis, A.L., Ciocca, V., Ng, B.K.C. 2003. On the (non) categorical perception of lexical tones. *Attention, Perception, & Psychophysics* 65, 1029-1044.

[2] Hallé, P.A., Chang, Y., Best, C.T. 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics* 32, 395-421.

[3] Huang, T. 2004 Language-specificity in auditory perception of Chinese tones. *http://rave.ohiolink.edu/etdc/view?acc_num=osu1092856661*

[4] Hume, E., Johnson, K. 2003. The impact of partial phonological contrast on speech perception. *Proc. 15th ICPhS* Barcelona, 2385-2388.

[5] Kiriloff, C. 1969. On the auditory discrimination of tones in Mandarin. *Phonetica* 20, 63-67.

[6] Shen, X. 1989. Toward a register approach in teaching Mandarin tones. *Journal of Chinese Language Teachers Association* 24, 27-47.