
Tone assimilation by Mandarin and Thai listeners with and without L2 experience

ARTICLE INFO

ABSTRACT

Article history:
Received 29 August 2013
Received in revised form
17 June 2014
Accepted 25 June 2014

Keywords:
Tone assimilation
Linguistic experience
Mandarin
Thai
Allotone
Category assimilation task

Evidence indicates that perceptual assimilation of segments is tied to L1 and L2 contrasts at a lower phonetic level for listeners without L2 experience, but at both a phonetic and a higher phonological level for those with L2 experience. It is less clear, however, that the same is true for suprasegmental features. In this examination of perceptual assimilation of lexical tones, 40 listeners, whose L1 and L2 are one of the two tone languages, Mandarin and Thai, along with another 40 native listeners of the two tone languages without L2 experience performed a mapping-rating assimilation task in which they first identified which L1 tone sounded most similar to the L2 tone they heard, and then rated the goodness of match on a 5-point scale. The inexperienced listeners assimilated L2 tones to L1 tones with the most similar acoustic properties, i.e., F0 height and contour. The experienced listeners were additionally influenced by phonological tone changes in Mandarin. In particular, falling rising tones were assimilated to the rising tone or low falling tone in Mandarin or Thai and vice versa. These findings are discussed in relation to current conceptions of perceptual assimilation.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Considerable research on the perception of segmental information has examined the factors influencing the assimilation of L2 phonetic and phonological categories to their L1 counterparts (Levy, 2009; Nishi, Strange, Akahane-Yamada, Kubo, & Trent-Brown, 2008; Schmidt, 1996; Strange, 1999; Strange, Bohn, Trent, & Nishi, 2004). In contrast with phonetic assimilation, in which L2 phonetic categories are perceived to be similar to L1 phonetic categories based on acoustic or gestural properties, phonological assimilation occurs because L1 and L2 phonetic categories have the same phonemic status. However, few studies have focused on perceptual assimilation at the suprasegmental level with respect to matters such as tone. Segmental assimilation may apply to tone, as tone functions phonemically in languages such as Mandarin and Thai; thus it is reasonable to expect that tone assimilation is affected by the acoustic properties and phonemic status of L1 and L2 tone categories. The current study explores tone assimilation by native speakers of two tone languages within the framework of recent models of cross-language speech perception.

1.1. Issues related to perceptual assimilation

Three central issues in recent models of cross-language perceptual assimilation are of particular interest in this investigation. One concerns the existence of assimilation processes at both the phonetic and phonological levels, a distinction that re

regard to the assimilation patterns involving categorized or assimilable categories, phonetic or phonological assimilation may be characterized as Two Category (two L2 categories are assimilated to two different L1 categories), Single Category (two L2 categories are assimilated to one L1 category) and Category Goodness (two L2 categories are assimilated to two L1 categories, but one is better). A second issue of concern here is the effect of language experience on perceptual assimilation. The PAM predicts assimilation exclusively at the phonetic level for listeners without L2 experience but very likely also at the phonological level for those with L2 experience (Best & Tyler, 2007). In contrast, the SLM postulates variable phonetic assimilation for listeners with varying L2 experience. The third issue concerns the extent to which phonetic similarity can predict assimilation at the phonetic and phonological levels. According to the PAM, phonetic similarity between L1 and L2 phones may be predictive of phonetic assimilation but not necessarily phonological assimilation (Best, 1995; Best & Tyler, 2007; Bundgaard-Nielsen, Best, & Tyler, 2011; Guion, Flege, Akahane-Yamada, & Pruitt, 2000; Levy, 2009; Levy & Strange, 2008). The SLM, however, posits that despite the importance of acoustic phonetic similarity, assimilation of phonetic categories can occur not only between acoustically similar phones in L1 and L2, but also between dissimilar ones.

1.2. Phonetic assimilation

Phonetic assimilation refers to a process in which listeners attend to the acoustic or gestural similarities between incoming L2 speech sounds and existing L1 phonetic categories (Best et al., 1988; Flege, 1995, 2007). Research has shown that both inexperienced and experienced listeners demonstrate perceptual assimilation at the phonetic level in the case of vowels (e.g. Levy, 2009; Nishi et al., 2008), consonants (e.g., Best et al., 2001), and lexical tones (e.g., Leung, 2008; So, 2012; So & Best, 2010). For example, in the perception of German/French vowels, both inexperienced and experienced American English (AmE) listeners assimilated German/French vowel categories /i, e, a, u/ to the AmE vowels /i, e, a, u/, respectively, based on the closer acoustic space between these vowels in the two languages (Levy, 2009; Strange et al., 2004). At the suprasegmental level, studies have investigated perception of Mandarin tones by Cantonese speakers with Mandarin experience (Leung, 2008) and those without (So, 2012; So & Best, 2010). Results indicate that Mandarin tones were identified as their acoustic equivalents in Cantonese, such as the high level tones and the rising tones, confirming the effects of acoustic similarities on tone perception.

Segmental and suprasegmental studies have also indicated that the assimilated phonetic categories may differ in their degree of acoustic phonetic similarity to L1 counterparts. For example, Strange et al. (2004) noted that the German vowel /ɪ/ was assimilated to the American English /i/ by native listeners of American English, despite the fact that the phonetic dissimilarities are greater between these vowels than those between the vowels /i, e, /a/, or /u/ in L1 and L2. At the suprasegmental level, Cantonese listeners have been found to identify Mandarin high falling tone as Cantonese high falling tone despite the fact that the Mandarin tone ends with a much lower fundamental frequency (F0) than the Cantonese tone, suggesting that phonetic assimilation may be induced by partially common features, in this case a falling contour (Leung, 2008; So & Best, 2010). Tone assimilation research by So (2012) also indicated that Mandarin falling rising tone was assimilated to both Cantonese low falling tone and rising tone because the former represents the initial part of falling rising tone and the latter the final part.

In some cases, it has been suggested that use of phonetic cues may affect assimilation results (Levy, 2009;

phonetic or phonological equivalent of an L1 tone category by listeners from different L1 tone language backgrounds but with L2 learning experience of the other language (So, 2012). To address this issue, the current study adopted the category assimilation method (e.g., Levy, 2009) to explore the assimilation patterns between Thai and Mandarin tone categories by inexperienced and experienced listeners. The rationale for selecting these two tone systems lay in the fact that the Thai tone system is similar to the Mandarin tone system in terms of the size of tone inventory and in the nature of acoustic correlates (Chao, 1948; Ladefoged, 2001). There are five tone categories in Thai (mid falling, low falling, high falling, high rising and falling rising) and four in Mandarin (high level, rising, falling rising and falling), both having level and contour pitch patterns. Moreover, Mandarin falling rising tone has two conditional allophonic variants: rising tone and low falling tone. As discussed earlier, falling rising tone is realized as rising tone under the condition where it precedes another falling rising tone at the word level. In most other cases, falling rising tone is realized as low falling tone when it is followed by the other three tones, i.e., high level, rising and high falling tone in a word (Shih, 1997). Thus low falling tone also functions as an allotone of falling rising tone in Mandarin (cf., Hallé, Chang, & Best, 2004).

Furthermore, as discussed previously, research findings have not been consistent with respect to phonetic and phonological assimilation patterns as a function of L2 experience. Thus the current study investigates tone assimilation at the phonetic and phonological levels by native speakers of Mandarin and Thai who differed not only in L1, but also in L2 tone experience. Compared to the inexperienced listeners who had no Mandarin experience, the experienced listeners had 0.5-2 years of exposure to Mandarin (for the experienced Thai listeners) or Thai (for the experienced Mandarin listeners).

The experiments employ a cross-language perceptual assimilation task, also known as a category assimilation task, which has been used widely to probe segmental assimilation (Levy, 2009; Nishi et al., 2008; Schmidt, 1996; Strange, 2007; Strange et al., 2004) as well as suprasegmental assimilation (So, 2012; So & Best, 2011). The task consists of a mapping portion and a rating portion,

acoustic equivalents in the other language. Nevertheless, the high F0 of the Mandarin level tone may be associated with the Thai high rising tone (45) and the falling contour of the Thai mid falling tone (32) may be related to the Mandarin falling tone (51).

In order to examine the assimilation patterns of Thai and Mandarin tones, it is crucial to investigate the actual acoustic properties of the two tonal systems. Conventional descriptions indicating categorical distinctions of the tones may not reflect the acoustic properties of tone categories naturally produced by different speakers in different circumstances (Gandour & Potisuk, 1991). Acoustically, a tone category can be produced with varying pitch trajectories in citation form. For instance, Abramson (1978) and Gandour (1978) described Thai mid tone as a mid (level) tone while Ladefoged (2001) referenced it as a mid falling tone. In Mandarin, rising tone and falling rising tone can be produced from highly distinguishable to very confusable depending on the realization of their acoustical parameters (Shen & Lin, 1991). Thus, in the current research, prior to the tone assimilation study, F0 correlates of Thai and Mandarin tones were measured acoustically. Following previous studies (e.g., Bent, 2005; Ciocca, Whitehill, & Ng, 2002; Wang, Jongman, & Sereno, 2003), the current research used the primary acoustic correlates of lexical tone, i.e., F0 height and contour, to compare and categorize Mandarin and Thai tones in that F0 height and contour have been verified as the most important perceptual cues contributing solely or jointly to tone identification (Thai, Abramson, 1997; Mandarin, Guion & Pederson, 2007; Lin, 1988). Other acoustic features of tone, such as tonal duration or voice quality may affect tone perception to some extent, but are not considered in the current study due to their relatively small effect on tone perception (e.g., Taiwanese, Lin & Repp, 1989).

~~2.1.1~~

~~2.1.2~~

A gender-balanced set of recordings of Thai and Mandarin tones was collected from two speakers each of Mandarin and of Thai for acoustic analysis. The Mandarin speakers were born and raised in Beijing and were in their early 20s at the time of recording. The male speaker had 1.5 years of residence in North America and the female speaker, 2.5 years. The Thai speakers, who had a mean age of 22 years, were born and raised in Bangkok. The male speaker had resided in North America for 4 years and the female speaker 2 years. The four speakers all reported normal speaking and hearing ability. All were undergraduate students of Simon Fraser University in Canada, and none had experience with other tone languages.

The stimulus set consisted of 12 syllables, each produced with the four Mandarin and five Thai tones, as well as hums which contain only tonal (but no segmental) information (Lee, Vukoch, & Wurm, 1996). As indicated in Table 2, 10 of the targets were real words in Mandarin or Thai varying in syllabic structure [CV(V), CVC, CCV(V)], and two were nonsense syllables in both languages.

of 44.1 kHz and 16-bit resolution. The Mandarin words were read in the carrier sentence“我 _____

mid level tone carried a much smaller falling range (0.3) compared to Thai low falling tone (2.6). Therefore, the former was grouped with level tones and the latter with falling tones. Nevertheless, the Thai mid tone having the features of a level tone is consistent with [Abramson \(1978\)](#) and [Gandour \(1978\)](#).

In summary, the results of the acoustic analysis illustrate acoustic correspondences between the Mandarin and Thai tone inventories. The results are consistent with the numerical descriptions by [Ladefoged \(2001\)](#), except for T-ML which he describes as a mid falling tone. These results will be used to interpret the assimilation data described below.

46

461

46

distance between L1 and L2 categories. Fit index is calculated by multiplying the mean percentage of responses and mean goodness rating in order to express the assimilation of an L2 category to L1 categories.

The degree of response diversity was calculated to assess the consistency in responses for each L2 tone category being mapped to every L1 tone category. Frequencies of responses in each L1 tone category were tallied, and the degree of response diversity (K') was calculated using the following formula (Koopman, Personal communication; [Simpson, 1949](#)):

$$K' = \frac{1}{\sum_{i=1}^R P_i^2}, \quad (2)$$

where R is the total number of L1 tone categories and P_i is the percentage of responses in which an L2 category (i) was assimilated to a particular L1 tone category. Minimum diversity ($K'=1$) indicates that an L2 tone category was consistently mapped to a single L1 tone

A two-factor mixed-design ANOVA was then computed using K' as the dependent variable, experience group (IT and ET) as a between-groups factor, and input tone (M-HL, M-R, M-FR and M-HF) as a repeated measure. Main effects of group, $F(1, 34)=7.2$, $p<.011$] and tone, $[F(3, 102)=11.6$, $p<.0001$], as well as the interaction of group and tone, $F(3, 102)=8.2$, $p<.0001$], were significant. To evaluate the interaction effect, follow-up one-factor ANOVAs were performed using tone as the independent variable for each group. A significant effect of tone was detected for the ET group, $F(3, 45)=25.1$, $p<.0001$], but not for the IT group $F(3, 57)=1.1$, $p=.35$]. Bonferroni-adjusted post hoc analyses for the ET group indicated significant differences in K' between M-HL and M-R ($p<.035$) and M-HL and M-FR ($p<.03$).

in terms of modal responses. Furthermore, L2 experience was found to affect the mappings of the Thai tones to the Mandarin tones, in particular the contour tones, as opposed to the level tones.

~~49~~ The K' scores were obtained for the mappings of each Thai input tone, with the results displayed in [Fig. 3](#). A mixed-design ANOVA was computed, using K' as the dependent variable, experience group (IM and EM) as the between-group factor, and input tone (T-ML, T-LF, T-HF, T-R and T-FR) as the repeated measure. A significant main effect of tone was observed, [

the Mandarin listeners appeared to perceive L2 tones to be more similar to L1 tone categories than the Thai listeners, especially

similarities. Furthermore, the discrepancy may be explained by the fact that the Cantonese tone system encompasses more tone categories and more complex phonological tone changes relative to many other tone systems, such as Thai and Mandarin. As suggested by [Zhang et al. \(2010\)](#), a 'denser'

๖A

See [Table A1](#) and [A2](#).

๖

Abramson, A. S. (1975). The tones of central Thai: Some perceptual experiments. In: J. G. Harris, & J. R. Chamberlain (Eds.), [พจนานุกรมภาษามลายู ฉบับจุฬาลงกรณ์ ร้อยคำ](#)
Bangkok: Central Institute of English Language

(pp. 1–16).

- Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America*, 107, 2711–2724.
- Hallé, P. A., Best, C., & Levitt, A. (1999). Phonetic versus phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics*, 27, 281–306.
- Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395–421.
- Hao, Y.-C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, 40, 269–279.
- Hashimoto, O. K. Y. (1972). *Phonology of Cantonese*. Cambridge: Cambridge University Press.
- Hattori, K. (2009). Perception and production of English /r/-/l/ by adult Japanese speakers (Doctoral dissertation). University College London.
- Huang, T. (2001). The interplay of perception and phonology in tone 3 sandhi in Chinese putonghua. *Ohio State University working papers in linguistics*, 23–42.
- Huang, T., & Johnson, K. (2010). Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners.