Categorizing Mandarin Tones into Japanese Pitch-accent Categories: The Role of Phonetic Properties

Connie K. So

MARCS Auditory Laboratories, University of Western Sydney, Australia

c.so@uws.edu.au

Abstract

This study examined how native Japanese speakers, who were naïve to Mandarin, categorized Mandarin tones (in citation form) into their native pitch–accent categories. Results showed that Japanese listeners categorized the nonnative Mandarin tones into their native pitch accent categories, in ways that were consistent with the phonetic features of listeners' native language. The findings support the new assumption of PAM for suprasegmentals [14] that non-native prosodic categories (e.g., lexical tones) will be assimilated to the categories of listeners' native prosodic system.

Index Terms: lexical tone perception, Perceptual Assimilation Model (PAM), phonological and phonetic influences

1. Introduction

It has been found that naïve adult listeners can perceptually assimilate non-native tones to the categories of their native prosodic systems, such as tone, pitch-accent, and intonation [13, 16], in ways that appear consistent with the assumptions of the Perceptual Assimilation Model (PAM) [3]. This raises an important question as to how adults perceive non-native lexical tones. Do they perceive non-native tones according to the some properties of the categories in their native prosodic system (e.g., phonetic properties: *pitch patterns*, *duration, voice quality*, *rhythmic /amplitude variations, etc*)? For example, will non-native rising tones be assimilated to the listeners' prosodic categories that have a rising pitch contour? While recent studies suggest that native speakers of (Australian) English and French were able to categorize Mandarin tones in terms of their intonational categories [14, 15], it will be also interesting to know how native speakers of pitch accent languages perceptually categorize Mandarin tones into their existing prosodic system (i.e., pitch accent system). To bridge the gap, the present study examined the categorization of Mandarin tones by Japanese speakers, who had not learned Mandarin and musical instrument at the time they participated the experiment.

Studies have shown that one's native language exert significant impact on perception of non-native segments [e.g., 17, 18]. Typically, it is constrained by the *phonological and phonetic properties* of listeners' native language. An example of a native *phonological* constraint is that Japanese speakers have difficulty in discriminating and categorizing English /r/ and /l/, as these phonemes are not contrastive in the Japanese phonological system [1, 12]. A case of native *phonetic* constraints was seen when American English listeners categorized the Zulu aspirated voiceless velar stop [k^h] and ejective [k']. They perceived both as the voiceless stop [k^h], and the non-native sound [k'] was perceived as a deviant English /k/. In other words, they did perceive the non-contrastive gradient difference but did not hear it as any English phonological contrast, because English has the voiceless stop [k^h] but no ejectives [4]. While studies generally suggested that listeners have difficulty distinguishing non-native segments that do not exist and contrast in their native language, English adults quite easily discriminated Zulu clicks, which are not contrastive in English [4]. The authors suggested that the clicks were nonspeech, rather than phonological elements in their native language.

Listeners' native languages also guide their perception of non-native tones substantially [6, 9, 10, 11, and 19]. For example, for studies examining the perception of Mandarin tones [9, 10, 11], it has been found that non-native speakers perceive Mandarin tonal categories (Tone 1 [55] - high level, Tone 2 [35] - mid rising, Tone 3 [214] - falling rising, and Tone 4 [51] - high falling) differently from native speakers, who can identify subtle differences between tones. For studies investigating the perception of tonal features (or dimensions) by listeners from different language backgrounds, Gandour [6, 7] found that native English listeners tended to focus on pitch height even though English is a non-tone language, while listeners from Cantonese and Mandarin focused on both pitch height and pitch direction when perceiving tones. Researchers have also found that native speakers of another tone language tended to outperform those of a non-tone language on the same non-native tone contrasts. For example, Cantonese speakers discriminated Mandarin tones better than did English speakers [11]. Similarly, Mandarin speakers discriminated two Thai tones (mid vs. low) better than English speakers did before the listener groups received training on the Thai tones [19].

In contrast, recent studies have demonstrated that listeners' tonal experience gained from their native language does not necessarily facilitate their perception and/or learning of non-native tones. Three typological-different language groups were tested to investigate how linguistic experience with tones might affect listeners' perception and/or learning of Mandarin lexical tones [13, 16]. It was found that native speakers of Hong Kong Cantonese, Japanese, and Canadian English had more perceptual difficulty in perceiving three phonetically similar pairs of Mandarin Tones (Tones 1&4, Tones 2&3, and Tones 1&2 (Note, not much difficult for the Cantonese speakers)) than the other three phonetically dissimilar pairs, (Tones 1&3, Tones 2&4, and Tones 3&4). Further, Hong Kong Cantonese speakers tended to make more perceptual errors in perceiving the Tone 1&4 and Tone 2&3 pairs than did Japanese and English speakers, suggesting that both the *phonological and phonetic properties* of the Cantonese tone system affect their perception of Mandarin tone. Phonemically, although high level [55] and high falling tones [53] occur in Cantonese, they are not contrastive in the language (both are allotones of Cantonese Tone 1; see [2, 20]. Therefore, it is likely that Cantonese listeners who have not learned Mandarin may perceive Mandarin Tone 1[55] and Tone 4[51] as variants of their Cantonese Tone1. Phonetically, because the phonetic properties (i.e., pitch contours) of Mandarin Tone 2[35] and Tone 3[214] are substantially similar, as both involve subtle rising pitch patterns, Cantonese

speakers are likely to perceive the two Mandarin tones as variants of Cantonese rising tones, such as Tone 2 [25] (Note: some describe this Cantonese tone as [35]). Similarly, for the Japanese and English listeners, the *phonetic properties* of the Japanese pitch accent and English intonational systems affect their perception of Mandarin tones. In particular, Mandarin Tone 2 and Tone 4 seemed to be similar to Japanese *LH* and *HL*, and English *Question* and *Statement* intonational categories, respectively. As for the phonological influences, it should also be mentioned that the phonemic vowel length contrast in Japanese also help Japanese listeners perceive the differences between Mandarin Tone 1 and Tone 4 better, as Tone 4 is well known for its short duration among the four tones. Taken together, both the phonological and phonetic properties of Japanese affect their native speakers to perceive non-native tones. The authors further suggested that the Perceptual Assimilation Model (PAM) [3] for suprasegmentals seemed feasible because the results were consistent with the PAM assumption and assimilation predictions. In other words, the listeners of the three language groups might assimilate non-native tones (instead of segments) to their native prosodic systems, and the assimilation predictions seemed to be applicable to the findings. For example, for Cantonese listeners, the case for Mandarin Tones 1&4 perceived as Cantonese Tone 1 can be interpreted as the *Single-Category* (SC) assimilation pair. Mandarin Tone 2 and Tone 3 can be considered as a *Category Goodness* (CG) assimilation pair, and Mandarin Tone 1 and Tone 2 can be considered as a *Two Category* (TC) assimilation pair. For Japanese listeners, the case for Mandarin Tones 2&4 perceived as Japanese *LH*, and *HL* could be interpreted as a *Two Category* (TC) assimilation pair. Similarly, the case for Mandarin Tones 2&4 perceived as English *Question* and *Statement* can be also considered as a *Two Category* (TC) assimilation pair. (see [14,16] for more examples for the listener groups).

A more recent study [14] has demonstrated that native English listeners can perceive non-native tones (in citation form) in terms of their intonation categories. The results indicated that Mandarin Tone 1 (High level) is perceived as *Flat Pitch*, Tone 2 (mid-rising) as *Question*, Tone 3 (falling-rising) as *Uncertainty* (Some English listeners perceived it as *Question*), and Tone 4 (high falling) as *Statement.* The findings supported the assumption that non-native lexical tones will be assimilated to the categories of listeners' native prosodic system (i.e., the intonation system in the study). The authors also suggested that native English listeners assimilated the phonetic properties of Mandarin tones (e.g., pitch patterns) to those of English intonation categories, when both substantially share similar phonetic features. Further, a new assumption of PAM for suprasegmentals was proposed – *listeners will assimilate non-native prosodic categories into their native prosodic categories when perceiving non-native tones* [14,15].

While native speakers of a non-tone language (English) are able to categorize the non-native (Mandarin) tones into their intonation categories according to the phonetic similarities between the categories of the two languages, it will be also important to test the new assumption of PAM for suprasegmentals with native speakers of a pitch accent language, such as Japanese. The results will reveal if native speakers of a pitch accent language would categorize Mandarin tones into their pitch accent categories, and if the phonetic properties of listeners' native language influence their tonal categorizations.

Therefore, this study examined how native Japanese speakers perceived Mandarin tones (in citation form) according to their own native prosodic categories, in this case their native pitch accent system. It was assumed that categorization would occur when the phonetic and/or the phonological characteristics/properties were compatible between the Mandarin tones and the listeners' native prosodic categories. Accordingly, based on the results of previous studies [13, 16], it was predicted that Japanese speakers would assimilate Mandarin Tone 1 [55] to their Japanese *HH*, Tone 2 [35] and Tone 3 [214] as the Japanese *LH*, and Tone 4 [51] as their *HL*. However, if they did not assimilate Mandarin tones to their pitch-accent categories when perceiving the tones, their responses should show a random selection pattern (i.e., no specific selections).

2. Method

2.1. Participants

Thirty Japanese participants were recruited in this study (mean age: 27.2 years). They were all either undergraduate students at the University of Western Sydney, who received course credits after they completed the experiment, or residents living in Sydney at the time of the experiment who received AUD $40 for their participation. This study included only participants who had never received any formal musical training, as previous studies have shown that listeners with musical training outperformed those without such training in both production and perception tasks with non-native tones [1, 8]. Before they performed the experiment, they all passed a pure-tone hearing screening (250- 8000 Hz at 25 dB HL).

2.2. Stimuli

The stimuli for this study were produced by three native Mandarin speakers (mean age: 24 years). They were asked to produce the four Mandarin tones on the syllable /fu/ in citation form. The syllable /fu/ was selected because its pronunciation is similar to Japanese *fuu* words that can carry each of the three Japanese pitch accents (see Table 1 for the target Japanese words with its associated pitch-accents).

Five tokens of each target word (/fu/ with each of the four Mandarin tones) were produced by each Mandarin speaker. Among them, three samples per tone-word per speaker were verified perceptually by another three native Mandarin speakers (mean age: 27.7 years) to ensure the selected stimuli were intelligible to native Mandarin speakers. All of the perceptual stimuli were correctly identified by the native speakers.

Table 1. *The three Japanese pitch accent categories, corresponding words, and glossaries.*

Pitch accent	Target words	Glossary
HH	**封**(筒) / **fuu** (too)	Envelope
HL	**夫**(婦) / **fuu** (fu)	Husband (and wife)
LH	(今)**風** / (ima) **fuu?**	Modern style?

Note: The bolded words are the target words. The bracketed words are used for helping listeners perceive the pitch-accent for the target words.

2.3. Procedure

Participants were asked to categorize randomized individual presentations of 72 trials of the individual stimuli (3 speakers x 4 tones x 3 tokens per tone x 2 repetitions) into their native Japanese prosodic categories: the three pitch accent categories, *HH*, *LH*, and *HL*. The experiment was given to the participants via a laptop. The target stimuli were presented individually from a separate PC screen. On each trial, three buttons were provided on the screen. Each button was labeled with one of the three Japanese words bearing a target pitch accent (see Table 1); the fourth button was labeled *Unknown* (確認不可能). Listeners were instructed to select the *Unknown* button only when they could not identify a tone into any of their native pitch accent categories.

3. Results

Native Japanese participants' assimilation percent (%) of Japanese Pitch-accent Categories (PA-Categories): *HH*, *LH*, and *HL*, for each tone was presented in Figure 1. (Note that: unlike other previous studies, in which the Unknown data was only 5% or less [14, 15], in this study, the analysis for the Japanese listeners' data was performed with the *Unknown* data, as it contributed 13.3 % (287 counts) of the total responses, 2160 counts).

Individual t-tests were carried out to test each PA-Category mean for each target tone, against chance of 25%. The results confirmed that the means of the following Japanese PA-Categories for the target Mandarin tones were all significantly above the chance level (25%): *HH* for Mandarin Tone 1 [$t(29) = 7.256$, $p < 0.001$], *LH* for Tone 2 [$t(29) = 10.590$, $p < 0.001$], both *LH* and *Unknown* for Tone 3 [$t(28) = 3.270$, and 5.868, $ps < 0.001$, respectively], and *HL* for Tone 4 [$t(28) = 7.940$, $p < 0.001$].

A Chi-square test revealed a significant association between the two factors, Tone (x4) and PA-Categories (x4), χ^2 (9) = 1391.269, $p < 0.001$. A further mixed design 2-way ANOVA (Tone x PA-Category) found no significant effect of Tone (*n.s.*), but a significant effect of PA-Category [$F(3, 305) = 2.653$, $p < 0.05$] on listeners' mean assimilation percentage (%). Their interaction was also significant [$F(9,305) = 31.858$, $p < 0.001$].

Individual 1-way ANOVAs for the four Mandarin tones were carried out to investigate the PA-Category effect for each tone target. It was found that the PA-Category effect was significant for each tone: Tone 1 [$F(3,75) = 22.462$, $p < 0.0001$], Tone 2 [$F(3, 86) = 61.360$, $p < 0.0001$], Tone 3 [$F(3,74) = 10.830$, $p < 0.0001$], and Tone 4 [$F(3, 70) = 24.296$, $p < 0.0001$]. Post-hoc HSD Tukey tests further indicated the following results for each tone. For Tone 1, the mean percentage of *HH* assimilations (59%) was significantly greater than each of the other three counterparts, *LH* assimilations (15%), *HL* assimilations (23%), and *Unknown* (4%; $ps < 0.001$), and the mean of *HL* assimilations was significantly greater than that of *LH* assimilations, $p < 0.001$. For Tone 2, the mean percentage of *LH* assimilations (65%) was significantly greater than each of the other three counterparts, *HH* assimilations (18%), *HL* assimilations (8%), and *Unknown* (9%; $ps < .001$). For Tone 3, both the mean percentages of *LH* assimilations (44%) and *Unknown* (37%) were significantly greater than each of *HH* assimilations (14%) and *HL* assimilations (5%; $ps < 0.001$). For Tone 4, the mean percentage of *HL* assimilations (65%) was significantly greater than each of the other three counterparts, *HH* assimilations (23%), *LH* assimilations (8%), and *Unknown* (4%; $ps < 0.001$).

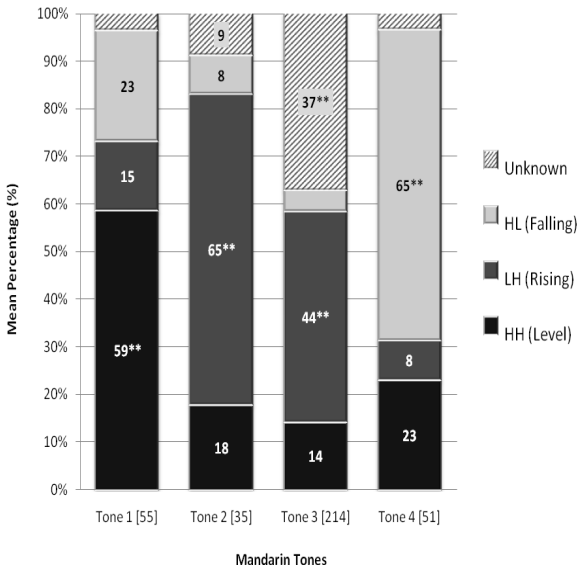

Figure 1: Native Japanese listeners' tonal categorizations for each tone (in %). The total number of responses for each tone category was 540. Categories that were selected 5% or less are not labeled. The symbol ** ($p < 0.01$) shows that the mean percentage of choice of the PA-Category is significantly above the chance level (25%).

4. Discussion

The results, as expected, clearly showed that Japanese listeners were able to perceptually assimilate Mandarin tones into their native pitch accent categories. The predictions, stated in the earlier section, were generally supported by the results, except that the categorization of Mandarin Tone 3 was not simply categorized into a single Japanese pitch accent category. Clearly, their responses for each target tone were not randomly selected, as they chose the pitch-accent categories that are phonetically similar to the Mandarin tones.

Specifically, the results showed that Mandarin Tone 1 was categorized as Japanese *HH* (level) pitch accent (59%); Tone 2 was categorized as the *LH* pitch accent (65%); and Tone 4 was categorized as their *HL* pitch accent (65%). For Tone 3, although they selected the *Unknown* option in a significant amount (37%), Tone 3 was categorized mainly as the *LH* pitch accent (44%). The categorization pattern for Mandarin Tone 3 implied that sometimes listeners were unable to make their selections from the category options, while sometimes they could perceive the phonetic similarities between their *LH* and Mandarin Tone 3[214]. Although both choice options were significantly above the chance level (25%), it was still reasonable to say that categorizing Tone 3 to Japanese *LH* was still weaker, compared to the categorizations of the other three Mandarin tones. Overall, the results evidence that the phonetic properties of the Japanese pitch accent system (e.g., pitch patterns) did play a significant role in Japanese speakers' perceptual categorizations of Mandarin tones.

In addition, the results of this study generally support the interpretations of tonal assimilations in PAM in previous studies [13, 16]. For example, Japanese speakers assimilate both Mandarin Tone 2 and Tone 4 as their Japanese *LH* and *HL*, respectively. Thus, these two tones can be considered as a *Two Category* (TC) assimilation pair. However, for Tone 1 and Tone 3, it appeared that only Tone 3 could be considered as uncategorized, since its categorization to *LH* was relatively weak.

The findings of the present study will also be useful to Mandarin teaching for Japanese learners at the initial stage. Since the Japanese listeners in this study were naïve to Mandarin (i.e., no prior knowledge of Mandarin and musical instruments), their perceptual performance are compatible with that of Japanese beginners of Mandarin at the initial stage of learning. In particular, the findings of the present study, as well as those of the previous studies [13, 16], suggest that Japanese speakers generally do not have too much problem in perceiving the Mandarin tones. This may due to the reasons that (1) their pitch-accent patterns of Japanese bi-morae words are phonetically similar to the three Mandarin tones – Tones 1, 2, and 4 (out of four), and that (2) Japanese vowel length contrast (a phonemic feature) may help naïve Japanese to perceive the Tone1[55]-Tone 4[51] contrast easier, since Mandarin Tone 4 are the shortest tones among the four [e.g., 13]. However, similar to the findings of previous studies [13, 16], Japanese speakers do have difficulty in perceiving Mandarin Tone 3[214]. Therefore, more focus on helping Japanese learners perceive Tone 3 should be made at the initial stage. For example, teachers may help them perceive the differences between Tone 2[35] and Tone 3[214], by pointing out the major phonetic differences in the production between the two tones, such as the durational and pitch pattern differences [13]. Once they master the Mandarin tones, other elements of the target language, such as phonological, semantic (lexical), syntactic, and pragmatic information, should be introduced in the teaching gradually, as they are closely interacted with the phonetic properties.

5. Conclusions

Overall, this study tested the new assumptions of PAM for suprasegmentals by examining the foreign (Mandarin) tonal categorizations by native Japanese speakers, who had no prior Mandarin knowledge. This study clearly demonstrated that Japanese speakers were able to categorize the Mandarin tones into their native pitch-accent categories. They perceived Mandarin Tone 1[55] as their *HH*, Tone 2[35] as Japanese *LH*, Tone 3[214] more as Japanese *LH* (but also 37% of time they selected the *Unknown* option), and Tone 4[51] as their *HL*. In addition, their selections depended on the phonetic similarities between Mandarin tone categories and Japanese pitch accent categories. Thus, the results support the new assumption of PAM at the suprasegmental level – listeners will assimilate non-native prosodic categories into their native prosodic categories when perceiving non-native tones [14, 15].

6. Acknowledgements

This work is supported by the Australian Research Council. The author thanks Hiroko Umeno for her assistance with data collection.

7. References

[1] Alexander, J. A., *et al.*, "Lexical tone perception by musicians and non-musicians," in *Proceedings of Interspeech 2005*, Lisbon, Portugal, 2005.

[2] Bauer, R. S. and Benedict, P. K., *Modern Cantonese Phonology*. Berlin and New York: Mouton de Gruyter, 1997.

[3] Best, C., "A direct realist view of cross-language speech perception," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, W. Strange, Ed., Timonium, Maryland: York Press, 171-232. 1995

[4] Best, C. T., *et al.*, "Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system," *The Journal of the Acoustical Society of America*, vol. 109, 775-794, 2001.

[5] Brown, C., "The Interrelation between Speech Perception and Phonological Acquisition from Infant to Adult," in *Second Language Acquisition and Linguistic Theory*, J. Archibald, Ed., Oxford: Blackwell, 4-63, 2000.

[6] Gandour, J. T., "Tone perception in Far Eastern languages," *Journal of Phonetics*, vol. 11, 149-175, 1983.

[7] Gandour, J. T., "Tone dissimilarity judgments by Chinese listeners," *Journal of. Chinese Linguistics*, vol. 12, 235-261, 1984.

[8] Gottfried, T. L. and Riester, D. N., "Preliminary results from study of pitch perception and Mandarin tone identification," *Journal of the Acoustical Society of America*, vol. 108, 2604, 2000.

[9] Hallé, P. A., *et al.*, "Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners," *Journal of Phonetics*, vol. 32, 395-421, 2004.

[10] Leather, J., "Speaker normalization in perception of lexical tones," *Journal of Phonetics*, vol. 11, 373-382, 1983.

[11] Lee, Y.-S., *et al.*, "Tone perception in Cantonese and Mandarin: A cross-linguistic comparison," *Journal of Psycholinguistic Research*, vol. 25, 527-542, 1996.

[12] Miyawaki, K., *et al.*, "An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English," *Perception and Psychophysics*, vol. 18, 331-340, 1975.

[13] So, C. K., "Effects of L1 prosodic background and AV training on learning Mandarin tones by speakers of Cantonese, Japanese, and English," PhD Thesis PhD, Department of Linguistics Simon Fraser University, Vancouver, 2006.

[14] So, C. K. and Best, C. T., "Do English Speakers Assimilate Mandarin Tones to English Prosodic Categories?," in *Proceedings of Interspeech 2008 incorporating SST 2008*, Brisbane, Australia, 2008.

[15] So, C. K. and Best, C. T., "Categorizing Mandarin Tones into Listeners' Native Prosodic Categories: The Role of Phonetic Properties," in *New Sounds 2010*, Poznan, 2010.

[16] So, C. K. and Best, C. T., "Cross-Language Perception of Non-native Tonal Contrasts: Effects of Native Phonological and Phonetic Influences," *Language and Speech*, vol. 53, 2010.

[17] Polka, L., "Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions," *The Journal of Acoustical Society of America*, vol. 89, 2961-2977, 1991.

[18] Werker, J. F. and Tees, R. C., "Phonemic and phonetic factors in adult cross-language speech perception," *The Journal of Acoustical Society of America*, vol. 75, 1866-1878, 1984.

[19] Wayland, R. P. and Guion, S. G., "Training English and Chinese listeners to perceive Thai tones: A preliminary report," *Language Learning*, vol. 54, 681-712, 2004.

[20] Yip, M., *Tone*. Cambridge: Cambridge University Press, 2002.