Mandarin Tone Perception among Yanbian Korean Ethnic Students: Postprint

Authors: Lu Yao, Li Yinghao, Kong Jiangping

Date: 2018-11-29T00:00:00+00:00

Abstract

This study investigates the categorical perception of Mandarin tones by Yanbian Korean Chinese students through identification and discrimination experiments. The experimental results show that Yanbian Korean Chinese university students with a high degree of Korean-Chinese bilingualism approach or reach native speaker level in their perception of Mandarin tones. Except for the T2-T3 pair, the degree of perceptual categorization for other tone contrasts shows little difference from that of native speakers. Native language background and a high degree of bilingualism may be contributing factors for Yanbian Korean Chinese students' perception of Mandarin tones approaching or reaching native speaker level, while the low level of categorical perception for T2-T3 is consistent with previous research findings on Mandarin learners with different native language backgrounds, suggesting that the similarity in phonetic-acoustic features between these two tones may prevent even advanced learners from successfully establishing native-like perception patterns.

Full Text

Preamble

Research on second language (L2) tone perception has demonstrated that learners' first language (L1) prosodic systems significantly influence their acquisition of Mandarin tones. Studies conducted around 2010 revealed that L2 learners exhibit varying degrees of difficulty in perceiving Mandarin tones, with performance closely tied to their L1 background. Hao (2012) found that speakers of tonal and non-tonal languages follow distinct developmental trajectories when acquiring Mandarin tones, while So (2005) showed that Cantonese and Japanese listeners experience different patterns of tonal confusion due to their L1 prosodic systems. Minagawa-Kawai et al. (2005) further demonstrated that native and non-native speakers employ different neural strategies during categorical perception of phonemes, suggesting that L1 experience shapes the fundamental mechanisms of speech perception.

The concept of "fine-grain sensory encoding ability" has been proposed to account for individual differences in L2 tone perception among learners with similar L1 backgrounds. Wang (2013) investigated how L1 experience and training affect the perception of Mandarin tones, revealing that both factors significantly influence categorical perception development. Several studies (e.g., 2005; 2011; 2013) have examined the psychophysical mechanisms underlying categorical perception, highlighting the complex interplay between linguistic experience and perceptual processing.

The present study focuses on Yanbian Korean bilinguals, a unique population whose L1, Yanbian Korean, possesses a complex prosodic system. Yanbian Korean exhibits a pitch accent system where lexical items are distinguished by high (H) and low (L) tonal patterns, as documented in historical linguistic research (Ito 2008, 2014; Ito & Kenstowicz 2009). Unlike Mandarin' s lexical tone system, Yanbian Korean' s pitch accent operates at the word level, with patterns such as the final accent class (e.g., 어머니 [ə̀.mə̀.ní], where ù indicates low tone and í indicates high tone) and penultimate accent class (e.g., 개구리 [kè.dzù. ák.ts*ȉ]). This system coexists with segmental phonology and intonational phrasing, creating a multifaceted prosodic structure (Jun, 2005).

The phonological proximity between Yanbian Korean and Mandarin, particularly in their use of pitch contrasts, suggests that Yanbian Korean speakers may possess enhanced sensitivity to Mandarin tones. However, the systems differ fundamentally: Yanbian Korean employs a binary pitch accent contrast, while Mandarin utilizes four distinct lexical tones. This distinction raises important questions about transfer effects and the extent to which L1 prosodic experience facilitates L2 tone perception.

Categorical perception is typically assessed through two complementary tasks: identification and discrimination. Identification tasks measure listeners' ability to assign stimuli to discrete categories, while discrimination tasks assess their sensitivity to within- versus between-category differences. The categorical perception paradigm, established in classic studies (Liberman et al. 1957; Studdert-Kennedy et al. 1970; Xu et al. 2006), reveals that listeners show enhanced discrimination at category boundaries but reduced sensitivity within categories. Key metrics include the width of identification (Xcb) and the peakedness of discrimination (Wcb), which quantify the sharpness of categorical boundaries and the degree of perceptual clustering, respectively.

Methodology

2.1 Participants

Thirty-four Yanbian Korean university students (15 male, 19 female) participated in this study. All were native speakers of Yanbian Korean who had acquired Mandarin Chinese as a second language through formal education. Participants reported nearly balanced bilingual competence, with daily exposure to both languages in academic and social contexts. Twenty-five native Mandarin speakers (10 male, 15 female) served as the control group. All participants had normal hearing and no history of speech or language disorders.

2.2 Stimulus Materials

The stimulus set comprised synthesized Mandarin tone continua generated using PSOLA (Pitch Synchronous Overlap and Add) resynthesis. Four Mandarin tones (T1, T2, T3, T4) were used to create six tone continua: T1-T2, T1-T3, T1-T4, T2-T3, T2-T4, and T3-T4. Each continuum consisted of 11 equidistant steps spanning the acoustic space between endpoint tones.

The synthesis process involved recording natural Mandarin productions and extracting pitch contours. For each continuum, the pitch values of the endpoint tones were interpolated linearly to generate intermediate steps. For example, in the T1-T2 continuum, Step 1 represented a canonical T1 (high-level tone) and Step 11 represented a canonical T2 (rising tone), with Steps 2-10 forming a graded transition. All stimuli were normalized to 500 ms duration and matched for intensity. The sampling rate was 22,050 Hz with 16-bit resolution.

Acoustic parameters for the synthesized stimuli are summarized in [Figure 2: see original paper] and [Figure 3: see original paper]. For the T1-T2 continuum, the pitch range extended from 172.72 Hz to 297.58 Hz. The T1-T3 continuum ranged from 126.65 Hz to 294.91 Hz, while the T1-T4 continuum spanned 102.58 Hz to 316.92 Hz. Similar acoustic specifications were established for the remaining continua.

2.3 Experimental Procedure

Participants completed two tasks in a sound-attenuated booth: an identification task and an AX discrimination task.

Identification Task: On each trial, participants heard a single stimulus and identified it as one of four Mandarin tones by pressing corresponding buttons labeled "T1,""T2,""T3,"and "T4."Each of the 132 stimuli (12 steps × 11 continua) was presented twice in random order, resulting in 264 trials. Responses were self-paced, with a 500 ms inter-trial interval.

Discrimination Task: The AX paradigm required participants to judge whether two sequentially presented stimuli were identical or different. Each trial consisted of two stimuli separated by a 500 ms inter-stimulus interval. The stimulus pairs could be identical (AA or BB) or different (AB or BA), with the second stimulus always differing from the first by 2-3 steps along the continuum. Participants responded using "same" or "different" buttons. Each continuum contributed 324 trials (18 same pairs + 9 different pairs × 12 steps), presented in random order across two sessions.

Stimuli were delivered binaurally through Sony MDR-7506 headphones at 72 dB SPL. The experiment was controlled by E-PRIME software.

2.4 Data Analysis

Three primary dependent measures were calculated:

(1) Identification Functions and Boundary Width (Xcb): For each participant and continuum, identification probabilities were fitted with binary logistic regression: log(P /(1-P)) = b + b x, where P represents the probability of identifying the stimulus as Tone 1, x denotes the step number along the continuum, and b and b are regression coefficients. The category boundary (Xcb) was defined as the step corresponding to a 50% identification rate: Xcb = -b /b . The boundary width (Wcb) quantified the steepness of the identification function, calculated as the difference between steps corresponding to 25% and 75% identification rates. Smaller Wcb values indicate sharper category boundaries and more robust categorical perception.

(2) Discrimination Accuracy and Peakedness (Ppk): Discrimination performance was computed using the formula P = P("S" |S) × P(S) + P("D" |D) × P(D), where P("S" |S) represents the hit rate for same trials, P("D" |D) represents the hit rate for different trials, and P(S) and P(D) denote the prior probabilities of same and different trials (1/3 and 2/3, respectively, in this design). Between-category discrimination accuracy (Pbc) was calculated for stimulus pairs crossing the identification boundary, while within-category accuracy (Pwc) was computed for pairs drawn from the same category. The discrimination peakedness (Ppk) indexed the difference between Pbc and Pwc, with higher values indicating stronger categorical perception.

Statistical analyses were conducted using Excel 2010 and IBM SPSS 20.0.

Results

3.1 Identification and Discrimination Functions

[Figure 4: see original paper] through [Figure 9: see original paper] display the identification and discrimination functions for each tone continuum. For native Mandarin speakers, all continua exhibited sharp identification boundaries and peaked discrimination functions, characteristic of robust categorical perception. The T1-T2 continuum showed a clear boundary at Step 4, with maximal discrimination accuracy at the category boundary. Similar patterns were observed for T1-T3, T1-T4, T2-T4, and T3-T4 continua.

For Yanbian Korean listeners, the results revealed a more nuanced pattern. While identification functions for T1-T2, T1-T3, T1-T4, and T2-T4 continua closely approximated those of native speakers, the T2-T3 continuum presented a notable exception. Identification boundaries were less sharply defined, and discrimination functions failed to show the characteristic peak at the category boundary, instead displaying relatively flat response profiles.

3.2 Category Boundaries and Boundary Width

Mean Xcb and Wcb values are presented in . For the T1-T2 continuum, native speakers' Xcb was located at Step 4.0, with Yanbian Korean listeners showing a comparable boundary (Xcb = 4.04). No significant group differences emerged for Xcb across most continua (all $p > 0.05$), except for T2-T3 ($F(1,1094) = 23.127$, $p < 0.005$) and T3-T4 ($F(1,742) = 38.25$, $p < 0.005$), where Yanbian Korean listeners exhibited significantly different boundary locations.

Analysis of Wcb revealed that Yanbian Korean listeners produced significantly wider boundaries for the T2-T3 continuum compared to native speakers, indicating less precise category differentiation. Specifically, the T2-T3 Wcb for Yanbian Korean listeners was 0.5 steps broader than that of native speakers ($p < 0.05$). For other continua, Wcb values did not differ significantly between groups.

3.3 Discrimination Accuracy and Peakedness

Discrimination performance metrics are summarized in . Native Mandarin speakers demonstrated significant discrimination peakedness (Ppk) across all continua, with Pbc substantially exceeding Pwc (all $p < 0.005$). The largest Ppk values were observed for T1-T2 (40.00%) and T1-T4 (35.92%) continua.

Yanbian Korean listeners showed comparable Ppk for most continua, with significant between-category advantages (Pbc > Pwc) for T1-T2, T1-T3, T1-T4, and T2-T4 (all $p < 0.005$). However, for the T2-T3 continuum, the difference between Pbc and Pwc failed to reach significance ($p = 0.17$), and Ppk was markedly reduced (8.47%) compared to native speakers. This pattern indicates impaired categorical perception specifically for the T2-T3 contrast.

Discussion

The present investigation reveals that Yanbian Korean bilinguals with extensive Mandarin exposure achieve near-native categorical perception of Mandarin tones across most contrasts. Their performance on T1-T2, T1-T3, T1-T4, and T2-T4 continua demonstrates sharp identification boundaries and robust discrimination peakedness, comparable to native Mandarin speakers. These findings align with previous research indicating that L1 prosodic experience can facilitate L2 tone acquisition when the systems share phonetic dimensions.

The exceptional difficulty with T2-T3 perception, however, mirrors findings reported for diverse L2 learner populations, including speakers of non-tonal languages and even advanced Mandarin learners. The phonetic similarity between T2 (rising) and T3 (dipping) contours appears to create a persistent perceptual challenge that resists amelioration through L1 transfer or extended L2 experience. This difficulty may stem from the overlapping pitch contours and dynamic characteristics of these tones, which require fine-grained sensory discrimination beyond what the Yanbian Korean pitch accent system provides.

Theoretical accounts of L2 speech perception (Best 1995; Flege 1992, 1999) suggest that L1-L2 phonetic similarity can either facilitate or hinder perception depending on the degree of overlap. In this case, the binary H/L contrast in Yanbian Korean may provide a foundation for perceiving Mandarin' s tonal inventory, but the subtle distinction between T2 and T3 exceeds the granularity of the L1 system. This interpretation is consistent with the "Perceptual Assimilation Model" and "Speech Learning Model," which predict that L2 sounds sharing some but not all features with L1 categories pose particular learning challenges.

From a pedagogical perspective, the T2-T3 contrast represents a critical challenge for Mandarin instruction, even for learners with tone-language backgrounds. The persistence of perceptual difficulty among advanced Yanbian Korean bilinguals suggests that explicit training targeting this specific contrast may be necessary. Techniques such as perceptual fading, adaptive training, and enhanced acoustic cue weighting could potentially improve discrimination sensitivity for this challenging pair.

The present findings also contribute to our understanding of bilingual processing in tone languages. Yanbian Korean speakers' balanced bilingualism appears to support the development of distinct phonological systems, allowing them to maintain native-like perception in both languages. This capacity for phonological compartmentalization may be enhanced by the functional separation of their languages in daily use, with Korean dominating domestic contexts and Mandarin prevailing in academic settings.

Conclusion

This study demonstrates that Yanbian Korean bilinguals achieve near-native categorical perception of Mandarin tones, with the notable exception of the T2-T3 contrast. Their success across most tone pairs underscores the facilitative role of L1 pitch accent experience, while the persistent T2-T3 difficulty highlights the limits of positive transfer. These results have important implications for theories of L2 speech perception and for the design of targeted tone training programs. Future research should explore whether focused intervention can remediate the T2-T3 perceptual deficit and investigate the neural correlates underlying this specific difficulty.

References

1. Wang, Y. (2005). *A Study on the Acquisition of Mandarin Tones by L2 Learners and Its Relationship with Tone Perception* [D]. Beijing: Beijing Language and Culture University.
2. Zhang, L., Wang, M., & Liu, X. (2011). Categorical perception of Mandarin tones by L2 learners with different L1 backgrounds [J]. *Applied Linguistics*, (4): 94-99.

3. Liu, X. (2014). An experimental study on Mandarin tone perception by Korean learners [J]. *Phonetic Research*, (2): 97-104.
4. Chen, Y., Li, M., & Zhou, H. (2013). Cross-language perception of Mandarin tones by English and Japanese speakers [J]. *Foreign Language Teaching and Research*, (6): 823-827.
5. Zhou, J. (2013). *The Perception of Mandarin Tones by L2 Learners* [D]. Shanghai: Shanghai International Studies University.
6. Li, M., & Zhang, Y. (2010). The relationship between musical training and tone perception [J]. *Acta Psychologica Sinica*, 42(9): 899-908.
7. Li, M., & Zhang, Y. (2011). The acquisition of Mandarin tones by Korean learners [J]. *Language Teaching and Research*, (1): 8-15.
8. Yang, L. (2009). *Categorical Perception of Mandarin Tones by Native and Non-native Speakers* [D]. Nanjing: Nanjing University.
9. Xu, M. (2016). The effect of L1 background on L2 tone perception [J]. *Nanjing: Nanjing University.*
10. Huang, L. (2010). The relationship between L1 prosodic system and L2 tone acquisition [J]. *Linguistic Research*, (2): 10-18.
11. Huang, L. (2010). The acquisition of Mandarin tones by speakers of tone languages [J]. *Language Teaching and Research*, (3): 9-15.
12. Zhang, X., & Wang, L. (2014). The effect of musical experience on tone perception [J]. *Music Research*, 3: 206-215.
13. Best, C. T. (1995). A Direct Realist View of Cross-Language Speech Perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 171-204). Timonium, MD: York Press.
14. Chen, A. (2013). *Universal Biases in the Perception of Mandarin Tones: From Infancy to Adulthood* [D]. Utrecht University.
15. Chen, F., Peng, G., Yan, N., et al. (2016). The development of categorical perception of Mandarin tones in four- to seven-year-old children [J]. *Journal of Child Language*, 44(6): 1-15.
16. Deutsch, D., Henthorn, T., Marvin, E., et al. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences and evidence for a speech-related critical period [J]. *Journal of the Acoustical Society of America*, 119(2): 719-722.
17. Flege, J. E. (1992). Speech Learning in a Second Language. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, and Implications* (pp. 565-580). Timonium, MD: York Press.

18. Flege, J. E. (1999). The relation between L2 production and perception [R]. In *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1273-1276). San Francisco.

19. Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners [J]. *Journal of Phonetics*, 32(3): 395-421.

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

21. Hao, Y. C. (2017). Second language perception of Mandarin vowels and tones [J]. *Language and Speech*, 60(3): 238-259.

22. Ito, C. (2008). Historical development and analogical change in Yanbian Korean accent. In *Harvard Studies in Korean Linguistics XII* (pp. 165-178).

23. Ito, C., & Kenstowicz, M. (2009). Mandarin loanwords in Yanbian Korean II: Tones [J]. *Language Research*, 45(1): 85-109.

24. Ito, C. (2014). Loanword accentuation in Yanbian Korean: A weighted-constraint analysis [J]. *Natural Language & Linguistic Theory*, 32: 537-592.

25. Jiang, C., Hamm, J. P., Lim, V. K., et al. (2012). Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics [J]. *Memory & Cognition*, 40(7): 1109-1121.

26. Jun, S. A. (2005). Korean intonational phonology and prosodic transcription. In *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 201-229).

27. Liberman, A. M. (1957). Some results of research on speech perception [J]. *Journal of the Acoustical Society of America*, 29(1): 117-123.

28. Luo, H., Boemio, A., Gordon, M., et al. (2007). The perception of FM sweeps by Chinese and English listeners [J]. *Hearing Research*, 224(1): 75-83.

29. Minagawa-Kawai, Y., Mori, K., & Sato, Y. (2005). Different brain strategies underlie the categorical perception of foreign and native phonemes [J]. *Journal of Cognitive Neuroscience*, 17(9): 1376-1385.

30. Peng, G., Zheng, H. Y., Gong, T., et al. (2010). The influence of language experience on categorical perception of pitch contours [J]. *Journal of Phonetics*, 38(4): 616-630.

31. Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels [J]. *Attention, Perception, & Psychophysics*, 13(2): 253-260.

32. Schouten, M. E. H., & Van Hessen, A. J. (1992). Modeling phoneme perception. I: Categorical perception [J]. *Journal of the Acoustical Society of America*, 92(4): 1841-1855.

33. So, C. K. (2005). The influence of L1 prosodic background on the learning of Mandarin tones: Patterns of tonal confusion by Cantonese and Japanese naïve listeners. In *Proceedings of the 2005 Annual Conference of the Canadian Linguistic Association*.

34. Studdert-Kennedy, M., Liberman, A. M., Harris, K. S., et al. (1970). Motor theory of speech perception: A reply to Lane' s critical review [J]. *Psychological Review*, 77(3): 234-249.

35. Wang, W. S. Y. (1976). Language change [J]. *Annals of the New York Academy of Sciences*, 280(1): 61-72.

36. Wang, X. C. (2013). Perception of Mandarin tones: The effect of L1 background and training [J]. *The Modern Language Journal*, 97(1): 144-160.

37. Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction [J]. *Journal of the Acoustical Society of America*, 120(2): 1063-1074.

Note: Figure translations are in progress. See original paper for figures.