L2 Acquisition and Processing of Mandarin Tone

Yue Wang

Dept. of Linguistics, Simon Fraser University, Burnaby, BC,

Canada V5A 1S6

Allard Jongman and Joan A. Sereno

Dept. of Linguistics, University of Kansas, Lawrence, KS 66044

Submitted to: *Handbook of Chinese Psycholinguistics*. Li, P., Tan, L., Bates, E. and O.J.L Tzeng (eds.). Cambridge University Press.

1. Introduction

Mandarin tones are manifested physically by different fundamental frequency (F0) values with F0 height and F0 contour as the primary acoustic parameters (Liu, 1924; Howie, 1976; Wu, 1986). In addition, amplitude and temporal properties such as overall duration and Turning Point are also effective phonetic correlates of the tones (Lin, 1965; Chuang, Hiki, Sone, and Nimura, 1971; Jongman and Moore, 1997). Studies in the perceptual domain have shown that the various acoustic cues are functionally integrated when Mandarin speakers identify the tones (Gandour, 1984; Massaro, Cohen, and Tseng, 1985; Gårding, Kratochvil, Svantesson, and Zhang, 1986; Blicher, Diehl, and Cohen, 1990; Shen and Lin, 1991; Moore and Jongman, 1997). The hemispheric processing of Mandarin tones reveals that, for native speakers, it is lateralized in the left hemisphere, suggesting that tones are processed as linguistic units, just like the segmental properties (Hsieh, Gandour, Wong, and Hutchins, 2001; Wang, Jongman, and Sereno, 2001).

These results about the processing of Mandarin tones for native speakers raise the question as to whether nonnative speakers process the tones auditorily (i.e., based on innate psychoacoustic mechanisms) or linguistically (i.e., resulting from language-specific experience). Indeed, for speakers whose native language is nontonal, tone has presented great difficulty, since the functional association between the F0 characteristics and the segmental structure is unfamiliar to them (e.g., Kiriloff, 1969; Bluhme and Burr, 1971; Shen, 1989). A number of cross-linguistic studies have been conducted to examine if and how nontonal speakers process the tones differently, how they improve the processing of the tones as a result of learning, and how this learning is instantiated in the brain.

2. L2 production of Mandarin tones

One of the general measures related to tone production is pitch range. Chen (1974) compared the pitch range between English and Chinese speakers, who were tested on stimuli involving words and sentences in both English and Chinese. The results show that the average pitch range of the native Chinese speakers speaking Chinese was 1.5 times wider than that of the English speakers when they spoke English. However, when the English speakers switched from English to Chinese, their pitch range increased substantially, although not to the extent of the Chinese speakers. The author claimed that native speakers of a nontonal language need to widen their pitch range to successfully acquire a tonal language.

Shen (1989) analyzed the tonal errors made by American learners who had studied Chinese for four months. The results indicated the error rate ranged from as high as 55.6% for Tone 4 to as low as 8.9% for Tone 2. The error rates for Tone 1 and Tone 3 were 16.7% and 9.4%, respectively. These results indicate that American learners have difficulty with all tones, but Tone 4 is especially difficult. Shen argued that Tone 4 was more likely to be subject to L1 interference, since it is prosodically less marked for English speakers.

Miracle (1989) also analyzed the tonal errors of second-year American learners of Chinese, and found an overall error rate of 42.9%. The errors were classified as either tonal register errors (too high or too low) or tonal contour errors. Miracle found that the errors were evenly divided between these two types, and were also evenly divided among the tones. In terms of the characteristics of the errors, the register errors for Tone 1 were characterized by realizing the high level tone too low in the tone space. The contour errors were characterized primarily by the Tone 1 level contour being replaced by a falling contour. For Tone 2, the register errors consisted primarily of beginning the tone too high in the tonal space. The contour errors came from substituting either a falling or level contour for the desired rising contour. For Tone 3, the register errors consisted exclusively of realizing the third tone too high in the tonal space. The contour errors consisted of realizing the third tone as a rising contour. For Tone 4, the register errors were entirely the result of placing Tone 4 in the mid-low register of the tonal space. The contour errors were primarily due to substituting a level contour for the expected falling contour.

3. L2 perception of Mandarin tones

Cross-linguistic studies have also been conducted to examine how speakers of tonal and nontonal languages differ in the perceptual processing of the tones. With respect to F0 height and contour, research has shown that the perceptual weights of these two dimensions were related to the linguistic experience of the listeners. Gandour (1983), using multidimensional scaling, examined the perception of tones by listeners of four tonal languages, including Mandarin, Cantonese, Taiwanese, and Thai, as well as by those of a nontonal language, English. He found that English listeners attached more importance to the height, and less to the contour dimension of F0 than did listeners of most tonal languages. Gandour claimed that since English has no contrastive tones, contour or otherwise, English listeners directed their attention almost exclusively to the F0 height of the stimuli. Lee, Vakock, and Wurm (1996) extended this research using a tone discrimination task. Both Cantonese and Mandarin tones were presented to Cantonese, Mandarin, and English listeners. Lee et al. (1996) found that tone language speakers were better discrimination tones, in terms of speed and accuracy of their responses, than were nontone language speakers. Tone language speakers seemed to acquire general tone discrimination abilities. Thus, it appears that listeners' strategy for tone perception depends to some extent on the linguistic function of pitch in their native language.

The perception of Mandarin tones has also been investigated in terms of the identification of the tonal categories. It was found that nonnative listeners' tone perception tended to be less "categorical" as compared to that of native listeners. Leather (1987) found that when labeling a synthetic tonal continuum from a Tone 1 at one end to a Tone 2 at the other, the shape of the labeling functions differed between the nonnative and the native listeners. Those of the Dutch and the English listeners tended to be less markedly sigmoid than those of the Chinese. He also found a greater spread in location of the category (Tone 1 or Tone 2) boundary among the Dutch and English listeners, suggesting that their labeling functions reflected unstable or linguistically inappropriate perceptual weighting. Similarly, Stagray and Downs (1993) found that Mandarin listeners had poorer differential sensitivity than English listeners because the Mandarin listeners had learned to categorize sounds of similar frequency together to facilitate their perception of tone phonemes. Their results suggested that listeners can differentially tune their auditory systems to certain physical properties of a sound as a function of their linguistic experience.

Nonnative tone perception has also been examined as a function of linguistic context and sentence position. Broselow, Hurtig, and Ringen (1987) tested American listeners' perception of Mandarin tones when the tones were presented in isolation as well as in the context of two and three syllables. One of the most important findings of this study was that the detectability of Tone 4 varied with its position in the context. That is, Tone 4 was the most easily identified tone when presented in isolation, and in the final position of doublets and triplets. However, its identification declined dramatically in non-final positions, and became the poorest among the four tones. The authors argued that the results reflected the interference from English intonation. Mandarin Tone 4 is acoustically similar to the unmarked pattern for declaratives at the end of an utterance in English, both involving a falling pitch. Therefore, it is conceivable that this tone is easy to perceive when it occurs sentence-finally, since it corresponds to a familiar pattern in English. However, when Tone 4 occurs in non-final position, it becomes unfamiliar to English listeners, as this is not a normal F0 pattern in English. Upon closer inspection, the authors found that when Tone 4 was misidentified, it was primarily identified as Tone 1, especially in utterance-final position. The authors also attributed this to native language interference. In English, high pitch is associated with a focused element in an utterance. Since both Tone 1 and Tone 4 start with a high pitch, English listeners are more used to paying attention to the high portion of the contour, as they do in a typical English declarative contour. English speakers apparently consider the fall at the end of the string as part of a typical sentence contour and therefore disregard it, rather than associate it with any particular syllable. These results suggest that English listeners' perception of Mandarin tones is influenced by their native intonation system. This is consistent with the findings of other studies showing the influence of English stress on the perception of Mandarin tone. For example, White (1981) observed that English listeners tended to perceive the Mandarin high tones as stressed and the low Tone 3 as unstressed, despite the fact that in Mandarin, the stress on a syllable is realized by duration and amplitude rather than F0.

Extrinsic factors such as the speaker's F0 and speaking rate also affected nonnative tone perception. Jongman and Moore (2000) studied the effects of speaker F0 and speaking rate on Chinese and American listeners' perception of a Tone 2-Tone 3 continuum that varied either along a spectral parameter (ΔF0), a temporal parameter (Turning Point), or both (ΔF0/TP). Their results showed that while both groups of listeners compensated for variations in F0 and speaking rate (i.e., showed normalization effects in perception), Chinese and American listeners did not weigh the acoustic cues in the same manner. Mandarin listeners showed the largest normalization effects for those continua which varied along a single acoustic dimension, with perception of Turning Point affected by speaking rate and perception of ΔF0 by speaker cues. For Mandarin listeners, language background aided in disambiguating phonemic contrasts. But English listeners, who do not use tones to make lexical distinctions, did not show such effects. Only when stimuli varied in both temporal and spectral acoustic dimensions did normalization occur. The normalization effects for the English listeners seemed to be a consequence of overall acoustic discriminability. Limitations on perceptual resources permitted the English listeners to attend to extrinsic information (F0 and speaking rate information in the precursor) only when intrinsic acoustic differences (both spectral and temporal parameters) became more perceptually salient.

4. Training perception and production of tone in the laboratory

The research described above revealed different patterns in the production and perception of Mandarin tones by nonnative speakers of Mandarin Chinese. An interesting question is whether and how Mandarin tones can be learned by nonnative speakers so that their tone perception and production become more native-like.

One approach assessing speech learning is training in laboratory settings. Previous research has shown that the adult human perceptual system can be modified with auditory training of L2 segmental properties (Pisoni, Aslin, Perey, and Hennessy, 1982; Jamieson and Morosan, 1986; Logan, Lively, and Pisoni, 1991).

Leather (1990) examined the effect of production training on perception in a group of Dutch speakers. The participants were trained to produce four Mandarin words (with the same syllable "*yu*") differing in tone. He found that the Dutch speakers were able to perceive the differences in tone after the production training. The author also observed an effect of perceptual training on production. After the Dutch listeners were perceptually trained to identify the four words differing in tone, they displayed improved ability to produce the words with the correct tones even though they had not been trained to produce the words. Leather concluded that training in one modality tended to be sufficient to improve the learners' performance in the other.

Other studies have also consistently shown that, after short perceptual tone training, nonnative speakers of Mandarin improved both their perception and production of Mandarin tones (Leather, 1990; Wang, Spence, Jongman and Sereno, 1999; Wang, Jongman, and Sereno, 2003a). In Wang et al. (1999), American learners of Mandarin were trained during the course of two weeks to identify the four tones in 100 natural words. Results showed that the perception of Mandarin tones improved significantly after training. Moreover, this improvement generalized to new stimuli and new voices, and was retained when probed six months after training. Further studies (Wang et al., 2003a) also showed that the tone contrasts gained perceptually transferred to production, as judged by native

Mandarin listeners. Acoustic analyses of the pre-and post-training productions revealed the nature of the improvement, showing that post-training tone contours approximated native norms to a greater degree than pre-training tone contours. Interestingly, pitch height and pitch contour were not mastered in parallel, with the former being more resistant to improvement than the latter. These results suggest that training produces highly generalized learning that yields long-term modifications of the learners' perception and production of Mandarin tones.

5. Non-native processing of Mandarin tone

As discussed in Jongman et al. [this volume], behavioral and neuroimaging research both show that native Chinese speakers process Mandarin tones as a linguistic property predominantly in the left hemisphere (Hsieh, Gandour, Wong, and Hutchins, 2001; Klein, Zatorre, Milner, and Zhao, 2001; Wang, Jongman, and Sereno, 2001). These studies further reveal that the processing of Mandarin tone by nonnative speakers differs as a function of the linguistic role of tone to these speakers.

While the results of PET (positron emission tomography) studies have shown a left hemisphere frontal activation for Mandarin Chinese tones by native Chinese speakers (Hsieh et al., 2000; Klein et al., 2001), these studies revealed that, for speakers of a nontone language such as American English, the processing of Mandarin tone lies in the homologous right hemisphere frontal regions, similar to the processing of pitch as a non-linguistic stimulus. This is consistent with the behavioral results from the dichotic listening study (Wang et al., 2001), which showed a lack of left-hemisphere dominance in the processing of Mandarin tone by American speakers who had no prior experience with a tone language.

Recent neuroimaging research further investigated how cortical involvements in tone processing could be modified in low proficiency learners as proficiency improves. Using fMRI (functional magnetic resonance imaging), Wang, Sereno, Jongman and Hirsch (2000, 2003b) examined the acquisition of Mandarin Chinese tone contrasts by adult American beginning learners, by comparing cortical activation during a tone identification task before and after a two-week training procedure. They found that improvements in performance were associated with an increase in activation in Wernicke's area (left superior temporal gyrus, Brodmann's Area 22) and emergence of additional activity within adjacent regions (left superior temporal gyrus, Brodmann's Area 42). Measures included a comparison of changes in cortical activity observed before and after training and an analysis of the magnitude of post-training activated cortex. The findings indicate that the early cortical effects of learning a second language involve both the expansion of preexisting language-related areas and the recruitment of additional cortical regions, suggesting the plasticity of the adult human brain in the acquisition of Mandarin tone.

6. Summary and future directions

The studies reviewed in this chapter revealed that native and nonnative speakers of Mandarin show different patterns in the perception and production of Mandarin tones. For native speakers acquiring Mandarin as L1, tonal pattern is an integral part of each word they learn, but such functional association between segmental structure and F0 contour is nonexistent in nontonal speakers' linguistic behavior. Therefore, the source of difficulty in tone acquisition has often been attributed to nonnative speakers' lack of sensitivity to tonal categories. For example, Dutch and English speakers, as opposed to Chinese, showed greater spread in location of the perceptual category crossover (Leather, 1987).

A second factor is the nonnative speakers' lack of experience in processing the various phonetic features characterizing Mandarin tones. Nonnative speakers may weigh the various acoustic cues differently (Gandour, 1983) or may have fewer perceptual resources left to attend to contextual information (Jongman and Moore, 2000). Moreover, the source of difficulty in learning tones has generally been attributed to interference from L1 features, with knowledge of the function of pitch in the English stress and intonation systems found to highly influence American listeners' perception of Mandarin tones (White, 1981; Broselow et al., 1987).

Studies of hemispheric processing have consistently demonstrated that Mandarin tones are processed differently by native and nonnative speakers. While for native speakers, the neural substrate underlying the ability to identify Mandarin tone is predominantly lateralized in the left hemisphere, this hemispheric specialization for lexical tone is not characteristic of nonnative speakers of Mandarin (Hsieh et al., 2001; Klein et al., 2001, Wang et al., 2001).

Although native and nonnative speakers process Mandarin tones differently, nonnative learners' ability to identify the tones can be significantly improved after a short perceptual training in the laboratory (Wang et al., 1999). This improvement appears to generalize to new contexts, transfer to the production domain, and is stored in learners' long-term memory (Wang et al., 2003a). Furthermore, the improvement in Mandarin tone perception and production as the result of intensive training was accompanied by a change in cortical representations in the native-like direction (Wang et al., 2003b). These results suggest that the adult production and perceptual system still demonstrates plasticity, and that cortical representations might be continuously modified as learners gain more experience with Mandarin. Most of the research investigated either naïve nonnative speakers of Mandarin or beginning learners. This raises the question as to whether the processing of Mandarin tone for advanced learners can be authentically native-like as the learners achieve high proficiency in Mandarin. Studies at the segmental domain have consistently shown that age of L2 exposure and fluency are important determinants in L2 acquisition and processing (Flege, 1997; Kuhl, 2000). Studies of the cortical representations also demonstrate these two factors as the determinant of the cerebral functions in bilinguals (Kim, 1997; Dehaene et al., 1997; Perani et al., 1998). Therefore, one of the important directions for future research is to systematically investigate learning of Mandarin tone as a function of proficiency and age of acquisition for a better understanding of the mechanisms underlying the dynamic process of language learning at the suprasegmental level.

References

Blicher, D.L., Diehl, R.L., and Cohen, L.B. (1990). Effects of syllable duration on the perception of the Mandarin tone2/tone3 distinction: Evidence of auditory enhancement. *Journal of Phonetics* 18, 37-49.

Bluhme, H. and Burr, R. (1971). An audio-visual display of pitch for teaching Chinese tone. *Studies in Linguistics* 22, 51-57.

Broselow, E., Hurtig, R.R., and Ringen, C. (1987). The perception of second language prosody. In G. Ioup and S.H. Weinberger (eds.), *Inter-language Phonology, The*

Acquisition of Second Language Sound System. Cambridge: Newbury House Publishers. pp. 350-361.

Chen, G.T. (1974). The pitch range of English and Chinese speakers. *Journal of Chinese Linguistics* 2, 159-171.

Chuang, C.K., Hiki, S., Sone, T., and Nimura, T. (1971). The acoustical features and perceptual cues of the four tones of standard colloquial Chinese. *Proceedings of the Seventh International Congress on Acoustics (Budapest),* 297-300.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, I., Paulesu, E., Perani, D., van de Moortele, P.F., Lehericy, S., and Le Bihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport*, 8, 3809-3815.

Flege, J.E. (1997). The role of phonetic category formation in second-language speech learning. In J. Leather and A. James (eds.): *New Sounds 97. Proceedings of the Third International Symposium on the Acquisition of Second-Language Speech*, Amsterdam: University of Amsterdam, pp. 79-88.

Gandour, J.T. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics,* 11, 149-175.

Gandour, J.T. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics,* 12, 235-261.

Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., and Hutchins, G.D. (2000). A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience*, 12, 207-222.

Gårding, E., Kratochvil, P., Svantesson, J.O., and Zhang, J. (1986). Tone 4 and Tone 3 discrimination in modern standard Chinese. *Language and Speech,* 29, 281-293.

Hsieh, L., Gandour, J., Wong, D., and Hutchins, G.D. (2001). Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain and Language*, 76, 227-252.

Howie, J.M. (1976). *Acoustical Studies of Mandarin Vowels and Tones*. Cambridge University Press, Cambridge.

Jamieson, D.G., and Morosan, D.E. (1986). Training non-native speech contrasts in adults: Acquisition of the English /S/-/C/ contrast by francophones. *Perception and Psychophysics* 40, 205-215.

Jongman, A., and Moore, C. (2000). The role of language experience in speaker and rate normalization processes. *Proceedings of the 6th International conference on Spoken Language Processing*, I, 62-65.

Kim, K.H.S., Relkin, N.R., Lee, K.M., and Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, 388, 171-174.

Kiriloff, C. (1969). On the auditory discrimination of tones in Mandarin. *Phonetica* 20, 63-67.

Klein, D., Zatorre, R.J., Milner, B., and Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImage* 13, 646-653.

Kuhl, P.K. (2000). A new view of language acquisition, *Proceedings of the National Academy of Science* 97, 11850-11857.

Leather, J. (1987). F0 pattern inference in the perceptual acquisition of second language tone. In A. James and J. Leather (eds.), *Sound Patterns in Second Language*

Acquisition. Dordrecht: Foris Publications, pp. 59-81.

Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather and A. James (eds.): *New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second Language Speech*. Amsterdam: University of Amsterdam, pp.305-341.

Lee, Y., Vakoch, D. and Wurm, L. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, 25(5), 527544.

Lin, M.C. (1965). The pitch indicator and the pitch characteristics of tones in Standard Chinese. *Acta Acoustica (China),* 2, 8-15.

Liu, F. (1924). *Szu Sheng Shih Yen Lu*. Ch'un Yi Shanghai.

Logan, J.S., Lively, S.E.,and Pisoni, D.B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America* 89, 874-886.

Massaro, D.W., Cohen, M.M., and Tseng, C. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics,* 13, 267-290.

Miracle, W.C. (1989). Tone production of American students of Chinese: A preliminary acoustic study. *Journal of Chinese Language Teachers Association,* 24, 49-65.

Moore, C.B., and Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *Journal of the Acoustical Society of America* 102, 1864-1877.

Perani, D., Paulesu, E., Sebastian Galles, N., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S.F., Fazio, F., and Mehler, J. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain*, 121, 1841-1852.

Pisoni, D.B., Aslin, R.N., Perey, A.J. and Hennessy, B.L (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance* 8, 297-314.

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

Shen, X.S. and Lin, M.C. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech,* 34, 145-156.

Stagray, J.R., and Downs, D. (1993). Differential sensitivity for frequency among speakers of a tone and a nontone language. *Journal of Chinese Linguistics,* 21, 143-163.

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

Wang, Y., Sereno, J.A., Jongman, A., and Hirsch, J. (2000). Cortical reorganization associated with the acquisition of Mandarin tones by American learners: An fMRI study. *Proceedings of the 6th International conference on Spoken Language Processing* II, 511-514.

Wang, Y., Jongman, A., and Sereno, J.A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language* 78, 332-348.

Wang, Y., Jongman, A., and Sereno, J.A. (2003a). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America,* 113, 1033-1044.

Wang, Y., Sereno, J.A., Jongman, A., and Hirsch, J. (2003b). fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience*, 15:7, 1-9

White, C.M. (1981). Tonal perception errors and interference from English intonation. *Journal of Chinese Language Teachers Association* 16, 27-56.

Wu, Z.J. (1986), *The Spectrographic Album of Mono-syllables of Standard Chinese*. Beijing: Social Science Press.