Perception and Production of Mandarin Chinese Tones

Allard Jongman

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

Wang Yue

Dept. of Linguistics, Simon Fraser University, Burnaby, BC, Canada V5A 1S6

Corinne B.Moore

ProQuest Media Solutions,, 3900 Kinross Lakes Parkway Richfield, OH 44286

Joan A. Sereno

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

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

1. Introduction

Lexical tones are pitch patterns that serve to provide contrasts in word meaning. They perform this function analogously to segments such as voiced and voiceless stops, except that tone may extend beyond one segment. Mandarin Chinese phonemically distinguishes four tones, with Tone 1 having high-level pitch, Tone 2 high-rising pitch, Tone 3 low-dipping pitch, and Tone 4 high-falling pitch (Chao, 1948). The same segmental context carries different meanings depending on the tone. For example, the meaning of Mandarin Chinese *ma* with Tone 1 is 'mother', the Tone 2 version means 'hemp', and the Tones 3 and 4 meanings are 'horse' and 'scold', respectively.

Pitch or tone is a function of the rate of vocal fold vibration (see Ohala, 1978, for a review of tone production). As stipulated in the myoelastic-aerodynamic theory of vocal fold vibration (e.g., Van den Berg 1958), vocal fold vibration is a cyclic process that is achieved by a complex combination of aerostatic and aerodynamic forces, operating together with air flow and laryngeal muscle forces. Changes in fundamental frequency (or in rate of vocal fold vibration) are made by manipulating tension in the vocal folds. This tension is increased or decreased by the laryngeal muscles, particularly the cricothyroid muscle and the thyroarytenoids. The cricothyroid has been linked to both rises and falls in pitch, while sternohyoid activity is associated with pitch lowering (Sagart, Halle, Boysson-Bardies, Arabia-Guidet 1986).

The rate of vocal fold vibration is quantified as the fundamental frequency (F0), expressed in Hertz (Hz). For example, when a speaker has a F0 of 300 Hz, her vocal folds are vibrating at a rate of 300 times per second. Rate of vocal fold vibration gives rise to the perceived pitch of speech sounds. Although the terms *pitch* and *fundamental frequency* are often used interchangeably, there is not a direct, one-to-one correspondence between perceived pitch and fundamental frequency. Psychophysical studies with pure tones suggest that while the perception of pitch as a function of fundamental frequency is approximately linear below 500 Hz, changes in pitch above that level are perceived more closely on a logarithmic scale relative to the physical changes in fundamental frequency.

2. Acoustic and perceptual characteristics of Mandarin tones

The four lexical tones in Mandarin Chinese are typical of many other tonal inventories in Chinese languages because F0 contour as well as F0 height is used to make phonemic distinctions. Phonetic analysis of these tones typically centers on measurement of fundamental frequency, although intensity and duration parameters have also been examined.

2.1 Fundamental frequency

Phonetic studies have examined the fundamental frequency contours of Mandarin Chinese tones (e.g., Liu 1924; Dreher and Lee 1966; Wang 1967; Dreher, Young and Lee 1969; Howie 1970; Chuang, Hiki, Sone and Nimura 1972; Rumjancev 1972; Moore and Jongman 1997). These studies indicate that F0 height and F0 contour are the primary acoustic parameters to characterize Mandarin tones. Wu (1986) measured F0 values of the four tones from narrow-band spectrograms. His measurements are shown in Table 1.

INSERT TABLE 1 ABOUT HERE

These acoustic characteristics are in good agreement with Chao's (1948) phonological description. Figure 1 shows an illustration of the F0 contours of the four Mandarin tones produced by a female speaker. In general, Tone 1 is high and relatively level over most of the tone's duration. Tone 2 exhibits a rise for much of its duration, and the onset of the rise occurs in the middle region of the F0 range and ends at a point approaching the F0 height of Tone 1. The Tone 3 contour occupies the lowest region of the F0 range overall, although extending at least to the midpoint of the range by the offset. Tone 3 onset is variable and can be close in frequency to that of Tone 2. Tone 4 begins high and falls to the bottom of the range.

INSERT FIGURE 1 ABOUT HERE

The importance of F0 as a primary cue to the perception of Mandarin tones is supported by a variety of studies, including Howie (1976) who performed perception experiments using synthetic speech. Three conditions were contrasted - synthetic speech modeled after natural F0 patterns, synthetic stimuli in which the F0 contours were made to sound monotone, and stimuli synthesized to sound like a whisper. Results showed that subjects were much better at identifying stimuli in which the pitch pattern was maintained; the monotone and whisper stimuli were identified just slightly above chance level.

Gandour (1984) explored whether F0 height or F0 contour serves as a more important perceptual cue. His results suggest that while both dimensions are important, Mandarin listeners seem to attach slightly more importance to F0 contour than F0 height.

However, in their investigation of only Tones 1 and 2, Massaro, Tseng and Cohen (1985) reported that neither F0 height nor F0 contour alone was sufficient for correct identification, but rather that both cues were used by subjects.

Some tones exhibit similar F0 contours. Acoustic properties for these particular tones have been studied in an effort to discover the extent to which F0 may distinguish the tones. Specifically, Tones 2 and 3 can be characterized in terms of two measures: Turning Point and ΔF0 (Moore and Jongman, 1997). Turning Point is the point in time at which the contour changes from falling to rising, and ΔF0 is the F0 change from the onset of the tone to the Turning Point. Tone 2 typically has an earlier Turning Point and smaller ΔF0 than Tone 3. Perception experiments have shown that these two measures are important cues to the distinction between Tones 2 and 3 (e.g., Blicher, Diehl, and Cohen, 1990; Shen and Lin, 1991; Shen, Lin and Yan, 1993; Moore and Jongman, 1997).

Similarly, acoustic analyses of Mandarin Tones 3 and 4, both of which have a falling contour over some portion of the tone, suggest several invariant properties of each tone: Tone 4 starts at a peak, or reaches one very quickly, then gradually falls to its low point, which may not be achieved until the next syllable; Tone 3, on the other hand, may either have a lower onset F0 than Tone 4, or it may reach its valley quickly, to maintain a low F0 level particularly over the second half of the vowel (Garding, Kratochvil, Svantesson and Zhang, 1986). Perception experiments in Garding et al. (1986) also showed that stimuli containing an early peak and a steep fall after the Turning Point were perceived as Tone 4, while stimuli with long durations at a low F0 level were perceived as Tone 3. These results concur with those of earlier studies indicating that F0 is the primary cue in tone perception.

2.2 Temporal properties of tones

Production data have shown that, in addition to F0 differences, tones may maintain consistent temporal differences. Mandarin tones differ in terms of overall duration (Dreher and Lee, 1966; Kratochvil, 1971; Ting, 1971; Chuang et al., 1972; Howie, 1976; Nordenhake and Svantesson, 1983). Tones 2 and 3 tend to be the longest, Tone 4 the shortest. The relative duration of any tone may change as a function of position in the sentence, however (Nordenhake and Svantesson, 1983). Blicher et al. (1990) showed that overall duration differences between Tones 2 and 3 in Mandarin are perceptually relevant in that stimuli ambiguous between Tones 2 and 3 were identified more often as Tone 3 when they were lengthened. In addition to overall duration, the temporal location of the Turning Point has been evaluated as a primary cue to the distinction between Tones 2 and 3. While Turning Point and $\Delta F0$ together typically specify tonal categories, it has been shown that, within a certain range of $\Delta F0$, Turning Point alone can change listeners' percepts from Tone 2 to Tone 3 (e.g., Shen and Lin, 1991; Shen, Lin and Yan, 1993; Moore and Jongman, 1997).

2.3 Amplitude

In addition to F0 and duration, studies have also investigated amplitude and its perceptual relevance in tone perception. Chuang et al. (1972) showed that Tone 4 has the highest overall amplitude while Tone 3 has the lowest. Lin (1988) attempted to categorize tone amplitude into five patterns: level, higher at onset, higher at offset, higher in the middle, and a double-peak amplitude contour. However, these patterns did not occur systematically with any one tone type, except for the double-peak pattern, which always occurred for Tone 3 tokens produced by male speakers. Lin's subsequent perceptual results suggest that F0 is the predominant cue, while duration and amplitude manipulations have little effect on tone perception.

There is some evidence, however, that amplitude contours alone can be utilized as perceptual cues. Using signal-correlated noise stimuli, which removes F0 and formant structure of natural speech but retains amplitude information, Whalen and Xu (1992) showed that subjects were able to identify all but Tone 1 tokens from the amplitude contours alone.

In sum, phonetic studies on acoustic characteristics of tones have found that fundamental frequency, ΔF0, Turning Point, duration, and amplitude all constitute phonetic correlates and perceptual cues for tone, with fundamental frequency typically being the most important. The following sections discuss phonetic studies that examine production and perception of tones as they interact with surrounding context and with the broader prosodic system of intonation and stress.

3. Tonal coarticulation

In the first comprehensive study of tonal coarticulation in Mandarin, Shen (1990) showed that coarticulation is bidirectional, and that coarticulation affects F0 height of the entire tone, rather than only onset and offset values. Most affected were tones following Tone 2 and then Tone 1, whose high offset F0 values had the effect of raising the onset F0 of the following tones; also affected was Tone 4, whose high onset raised the offset of the preceding tone. Not affected were lexical tones preceding Tones 2 and 3 whose onset falls in the middle of the frequency range. Shen's other findings were that tone contour was not affected by coarticulation, and that tonal coarticulation did not extend beyond one syllable.

Xu (1993; 1994) also finds evidence for both anticipatory and carryover coarticulation effects in tone production. Examining both production and perception of tones, Xu showed that coarticulation may alter the canonical F0 contour of the tone to the point that tones are not identifiable. However, when semantic information was removed from the tones through waveform editing, listeners compensate for the coarticulation as if the coarticulation effects were actually present in the signal, although these altered tones are unidentifiable when presented in isolation. While Xu's results confirm those from Shen (1990) regarding bidirectional coarticulation, he also finds that carryover effects are greater than anticipatory effects. Furthermore, his findings contradict Shen's in that he observes carryover effects not just in adjacent syllables, but extending beyond neighboring syllables into the second or third syllable as well. Another interesting result captured by Xu is that anticipatory coarticulation in Mandarin tones is dissimilatory for all but the low tone. This research reveals that variation in tone production and perception is systematic, and that this variation may be utilized by listeners in tone identification.

4. Interactions of tones, stress and intonation

Tones interact with other prosodic domains including stress and intonation. The interaction between tones and intonation has generated studies examining tone patterns in sentential environments. In general, the properties of tonal contour, relative duration, and peak amplitude seem to be maintained across different intonation types (Ho, 1976, 1977; Shen, 1989). Indeed, Connell, Hogan, and Roszypal (1983) showed that the perception of tones is stable across a variety of intonation types, a finding that made the authors conclude that this stability allows intonation to function in Mandarin.

As in many other languages, F0, duration, and amplitude are acoustic correlates of stress in Mandarin (Kratochvil, 1969). Using discriminant analysis on a corpus of spontaneous speech produced by a single speaker, Coster and Kratochvil (1984) showed that the F0 dimension most successfully corresponded with the actual tone and stress level identifications for the syllables. The amplitude parameter corresponded with the actual categories at better than chance level, while duration was not found to be successful at determining stress and tone categorization.

Moore (1993) compared F0 and duration properties of stressed and unstressed tones with underlyingly atonic syllables. Atonic syllables consists of a small group of lexical items that include grammatical markers or lexicalized syllables (Chao, 1968; Duanmu, 1990). Atonic syllables carry the neutral tone (Tone 0) whose F0 pattern is predictable depending on the preceding tone. The neutral tone is also known to be short in duration. Results showed that unstressed syllables remain distinct from toneless syllables in that their durations are longer, their canonical F0 contours are not lost, and unlike toneless syllables, onset F0 values were distinct from preceding tonal offsets. Stress was observed to modify the F0 peak of tones, and relative durations between the tones were preserved in the unstressed tones as compared to the stressed tones. Thus, the phonological influence of stress alters acoustic characteristics to make some properties more prominent with respect to others in the phrase, but stress alone may not be responsible for unrecoverable changes to the intrinsic characteristics of the tone.

The importance of duration as a perceptual cue to stress was also examined by Shen (1993), who found that listeners identified tones with longer durations more often as stressed and concludes that duration, rather than F0, is the primary acoustic correlate of stress in Mandarin tones.

The consensus of these studies is that stress affects both F0 and duration properties of tones. F0 highs become higher, and to a lesser extent, lows become lower (Shen, 1990; Shih, 1988). Underlying tone contours are resilient to the influence of stress, however. Durations are longer for stressed tones; tones short enough in duration to make identification difficult may be perceived as completely unstressed and toneless (Lin, 1985). These general findings have contributed to an understanding of how acoustic parameters vary as a function of stress level. The intonational level only appears to impose constraints on F0 while stress employs both F0 and duration.

The phonological pattern of tone sandhi has also inspired several phonetic studies. Tone sandhi refers to the phonological phenomenon whereby a tone takes on features of a neighboring tone in a particular environment. An example that has gained widespread attention is Mandarin Tone 3 sandhi, in which a Tone 3 preceding another Tone 3 becomes a Tone 2. Wang and Li (1967; see also Peng, 2000) showed that listeners were not able to distinguish the Tone 3 that had undergone tone sandhi from a Tone 2.

5. The influence of speaker F0 range

As suprasegmentals, tones are perceived relative to other tones, although they are also distinguished by intrinsic acoustic properties. For tones that contrast in both of these dimensions, intrinsic F0 information may be sufficient for correct identification. To identify tones differing only in F0 height, however, listeners must refer to their knowledge of the speaker's F0 range, and where tones occur within that range. For example, a low tone produced by a high-pitched speaker and a high tone produced by a low-pitched speaker may be acoustically very similar. The process by which listeners adjust perception according to speaker-specific acoustic information is referred to as speaker normalization. While a few previous studies have investigated the role of extrinsic F0 (e.g., Leather, 1983; Fox and Qi, 1990), Moore and Jongman (1997) have provided the most convincing evidence for speaker normalization in tone perception.

Moore and Jongman (1997) investigated speaker normalization in the perception of Tones 2 and 3 by examining listeners' use of F0 range as a cue to speaker identity. Two speakers were selected such that Tone 2 of the low-pitched speaker and Tone 3 of the high-pitched speaker occurred at equivalent F0 heights. Three tone continua varying in either Turning Point, ?F0, or both acoustic dimensions, were then appended to a natural precursor phrase from each of the two speakers. Results showed identification shifts such that identical stimuli were identified as low tones for the high precursor condition, but as high tones for the low precursor condition. Stimuli varying in Turning Point showed no significant shift, suggesting that listeners normalize only when the precursor varies in the same dimension as the stimuli. Overall, these results suggest that tone identification is influenced by changes in F0 range, and that this information is used as a frame of reference according to which ambiguous tones may be interpreted.

6. Hemispheric specialization for tone

Previous research indicates that the left hemisphere is more adept at phonemic processing, including phonemes, syllables, and words (Kimura, 1961; Shankweiler and Studdert-Kennedy, 1967; Studdert-Kennedy and Shankweiler, 1970) while the right hemisphere is better at melodic and prosodic processing, including music, pitch contours, and affective prosody (Kimura, 1964; Curry, 1967; Bryden, 1982). The processing of lexical tone therefore presents an interesting case. The fact that lexical tones are used to make phonemic contrasts would suggest that they are primarily processed by the left hemisphere. However, since lexical tones are instantiated by modulations of F0, one could instead expect the right hemisphere to be dominant in their processing.

Using a task in which listeners were asked to identify dichotically presented tone pairs by stating which tone they heard in which ear, Wang, Jongman, and Sereno (2001) showed that Mandarin tones are predominantly processed in the left hemisphere by native Mandarin speakers. Moreover, native speakers of American English with no prior experience with a tone language did not show a dominance of either hemisphere, suggesting that left hemisphere dominance for the Mandarin speakers arose from the intrinsic linguistic significance of the F0 modulations. These results are consistent with recent neuro-imaging studies that show that prefrontal cortex in the right hemisphere is involved in pitch judgment tasks (Zatorre, Evans, Meyer, and Gjedde, 1992), whereas left hemisphere prefrontal cortex is involved in the processing of lexical tone in native speakers of tone languages when the presented tones have linguistic relevance (Gandour, Wong, Hsieh, Weinzapfel, Van Lancker, and Hutchins, 2000; Klein, Zatorre, Milner, and Zhao, 2001; Wang, Sereno, Jongman, and Hirsch, in press).

7. Summary and future directions

The research on production and perception of Mandarin tones reviewed in this chapter demonstrates that fundamental frequency, amplitude, and temporal properties such as overall duration and Turning Point are effective phonetic correlates of tone. Evidence was also presented to illustrate how tones interact with other suprasegmental processes such as intonation and stress. In particular, research suggests that the surface

F0 pattern in sentences is driven by the tones. Stress and intonation may stretch, contract or shift F0 range, but the tone is left distinct.

These findings suggest that tones in Mandarin Chinese are affected by some of the same phonetic processes as segments. Coarticulation, intrinsic pitch, and context effects demonstrate that tones are sensitive to the surrounding phonetic environment. A complex array of acoustic properties may be used to distinguish tones, as in segments.

One distinction between suprasegmentals and segments resides in their relationship to context information. Descriptions of segments may be stated in terms of formant frequencies in the case of vowels, for example, or manner and place of articulation cues for consonants. These descriptions only include properties that are intrinsic to the segment. Tone descriptions, on the other hand, must always include a relative property, such as "high" or "low". Thus, as these terms imply, tone identity is assumed to be intimately linked to the surrounding context (e.g., Lehiste, 1970). More research is required to understand exactly how these acoustic properties are affected by context.

The increased availability of neuroimaging techniques will result in a more detailed understanding of the processing of both segmental and suprasegmental information, including the extent to which speakers of tone languages process F0 fluctuations in a way that is fundamentally different from speakers of non-tonal languages.

References

Berg, J.W. van den (1958). Myoelastic-aerodynamic theory of voice production. *Journal of Speech and Hearing Research* 1, 227-244.

Blicher, D. L., Diehl, R. and Cohen, L. B. (1990). Effects of syllable duration on the perception of the Mandarin Tone 2/Tone 3 distinction: evidence of auditory enhancement. *Journal of Phonetics* 18, 37-49.

Bryden, M.P. 1963. Ear preference in auditory perception. *Journal of Experimental Psychology* , **65**, 103-105.

Chao, Y-R. (1948). *Mandarin primer* (Harvard University Press, Cambridge).

Chao, Y-R. (1968). *A grammar of spoken Chinese* (University of California Press, Berkeley).

Chuang, C. K., Hiki, S., Sone, T., and Nimura, T. (1972). The acoustical features and perceptual cues of the four tones of Standard Colloquial Chinese. *Proceedings of the Seventh International Congress on Acoustics* (Adadémial Kiado, Budapest), 297-300.

Connell, B. A., Hogan, J.T., and Rozsypal, A.J. (1983). Experimental evidence of interaction between tone and intonation in Mandarin Chinese. *Journal of Phonetics* 11, 337-351.

Coster, D. C., and Kratochvil, P. (1984). Tone and stress discrimination in normal Beijing dialect speech. *New Papers on Chinese Language Use: Contemporary China Papers* 18 (Australian National University, Canberra), 119-132.

Curry, F.K.W. 1967. A comparison of left-handed and right-handed subjects on verbal and non-verbal dichotic listening tasks. *Cortex* , **3**, 343-352.

Dreher, J. and Lee, P.C. (1966). Instrumental investigation of single and paired Mandarin tonemes. *Research Communication 13*, Douglas Advanced Research Laboratories.

Dreher, J.J., Young, E.L., and Lee, P.C. (1969). Mandarin Triplet Contours. *Research Communication* 107, Douglas Advanced Research Laboratories.

Duanmu, S. (1990). *A formal study of syllable, tone, stress and domain in Chinese languages*. Ph.D. dissertation, Massachusetts Institute of Technology.

Fox, R., and Qi, Y. Y. (1990). Context effects in the perception of lexical tone. *Journal of Chinese Linguistics* 18, 261-283.

Gandour, J. (1978). The perception of tone. In *Tone: A Linguistic Survey*, V.A. Fromkin (ed.) (Academic Press, NY), 41-76.

Gandour, J. (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.

Ho, A. T. (1976). The acoustic variation of Mandarin tones. *Phonetica* 33, 353-367.

Ho, A. T. (1977). Intonation variations in a Mandarin sentence for three expressions: interrogative, exclamatory, and declarative. *Phonetica* 34, 446-456.

Howie, J. M. (1970). *The vowels and tones of Mandarin Chinese: Acoustical measurements and experiments*. Ph.D. dissertation, Indiana University.

Howie, J. M. (1976). *Acoustical studies of Mandarin vowels and tones* (Cambridge University Press, Cambridge).

Kimura, D. 1961. Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology* , **15**, 166-171.

Kimura, D. 1964. Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology* , **16**, 335-358.

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.

Kratochvil, P. (1969). Syllabic volume as acoustic correlate of perceptual prominence in Peking dialect. *Unicorn* 5, 1-28.

Kratochvil, P. (1971). An experiment in the perception of Peking dialect tones. *A Symposium on Chinese Grammar: Scandinavian Institute of Asian Studies Monograph Series 6*, Scandinavian Institute of Asian Studies, (Lund, Sweden), 7-31.

Leather, J. (1983). Speaker normalization in perception of lexical tone. *Journal of Phonetics* 11, 373-382.

Lehiste, I. (1970). *Suprasegmentals.* (MIT Press, Cambridge).

Li, C. N., and Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, 4, 185-199.

Lin, M. C. (1988). Putong hua sheng diao de sheng xue texing he zhi jue zhengzhao [Standard Mandarin tone characteristics and percepts]. *Zhongguo Yuyan* 3, 182-193.

Liu, F. (1924). *Szu Sheng Shih Yen Lu* [Experimental studies of tone]. Ch'un Yi Shanghai.

Massaro, D. W., Cohen, M. M., and Tseng, C. C. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics* 13, 267-289.

Moore, C.B. (1993). Some observations on tones and stress in Mandarin Chinese. *Working Papers of the Cornell Phonetics Laboratory* 8, 82-117.

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.

Nordenhake, M., and Svantesson, J.-O. (1983). Duration of Standard Chinese word tones in different sentence environments. *Working Papers* 25, (Lund, Sweden), 105-111.

Ohala, J. J. (1978). Production of tone. In *Tone: A Linguistic Survey*, V.A. Fromkin (ed.) (Academic Press, NY), pp. 5-32.

Peng, S.-H. (2000). Lexical versus 'phonological' representations of Mandarin sandhi tones. In M.B. Broe and J.B. Pierrehumbert (eds.) *Acquisition and the lexicon: Papers in Laboratory Phonology V*. Cambridge: Cambridge University Press, pp. 152-167.

Rumjancev, M. K. (1972). Ton i Intonacija v Sovremennom Kitajskom Jazyke [Tone and intonation in Modern Chinese] (Izdatel'stvo Moskovskogo Universiteta, Moscow). Reviewed by A. V. Lyovin (1978). *Journal of Chinese Linguistics* 6, 120-168.

Sagart, L., Halle, P., de Boysson-Bardies, B., and Arabia-Guidet, C. (1986). Tone production in Modern Standard Chinese: An electromygraphic investigation. *Cahiers Linguistique Asie-Orientale* 15, 205-221.

Shankweiler, D., and Studdert-Kennedy, M. 1967. Identification of consonants and vowels presented to left and right ears. *Quarterly Journal of Experimental Psychology*, **19**, 59-63.

Shen, X. (1990). Tonal coarticulation in Mandarin. *Journal of Phonetics* 18, 281-285.

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

Shen, X. S. (1993). Relative duration as a perceptual cue to stress in Mandarin. *Language and Speech* 36, 415-433.

Shen, X., Lin, M., and Yan, J. (1993). F0 turning point as an F0 cue to tonal contrast: A case study of Mandarin tones 2 and 3. *Journal of the Acoustical Society of America* 93, 2241-2243.

Shih, C. (1988). Tone and intonation in Mandarin. *Working Papers of the Cornell Phonetics Laboratory* 3, 83-109.

Studdert-Kennedy, M. and Shankweiler, D. 1970. Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America,* **48**, 579-594.

Wang, W.S-Y., and Li, K-P. (1967). Tone 3 in Pekinese. *Journal of Speech and Hearing Research* 10, 629-636.

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., Sereno, J.A., Jongman, A., and Hirsch, J. (in press). fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience.*

Whalen, D.H., and Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, 49, 25-47.

Xu, Y. (1993). *Contextual tonal variation in Mandarin Chinese*. Unpublished Ph.D. dissertation, University of Connecticut.

Xu, Y. (1994). Production and perception of coarticulated tones. *Journal of the Acoustical Society of America*, 95, 2240-2253.

Wu, Z.J. (1986). *The spectrographic album of mono-syllables of Standard Chinese* (Social Science Press, Beijing).

Zatorre, R., Evans, A., Meyer, E., and Gjedde, A. (1992). Lateralization of phonetic and pitch processing in speech perception. *Science*, 256, 846-849.

Table 1. Average F0 values (in Hz) of the four Mandarin tones produced by one male and one female speaker (Wu, 1986). The first and last values represent the onset and offset values, respectively, while the middle value for Tone 3 represents the fundamental frequency at the Turning Point.

	Tone 1	Tone 2	Tone 3	Tone 4
Male	190-177	124-178	124-68-141	223-80
Female	307-305	222-318	221-165-242	352-166

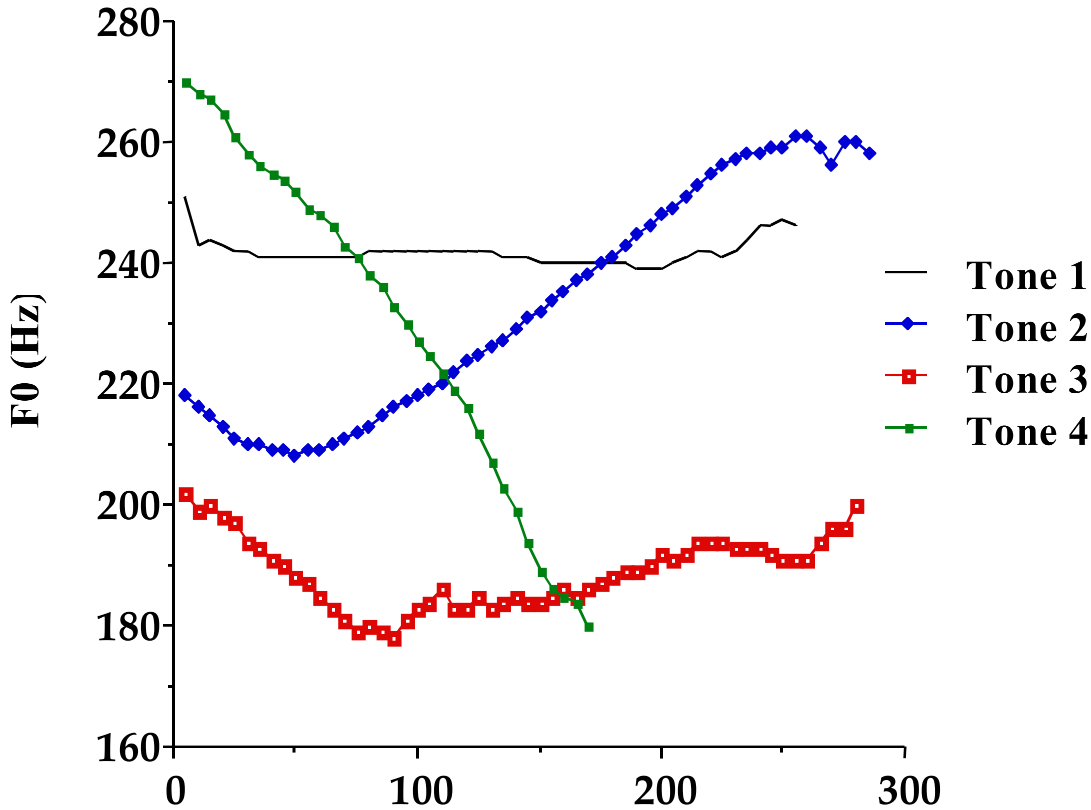

Figure 1. F0 contours for each of the four Mandarin Chinese tones for the segmental context *ma* spoken in isolation by a female speaker (from Moore and Jongman, 1997).