

# The Handbook of East Asian Psycholinguistics

Volume 1: Chinese

---

*Edited by*

Ping Li

*University of Richmond*

Li Hai Tan

*University of Hong Kong*

Elizabeth Bates

*University of California, San Diego*

Ovid J. L. Tzeng

*Academia Sinica, Taiwan*

*General Editor:*

Ping Li

CAMBRIDGE UNIVERSITY PRESS  
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, Sao Paulo  
Cambridge University Press  
The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521833332](http://www.cambridge.org/9780521833332)

© Cambridge University Press 2006

This book is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2006

Printed in the United Kingdom at the University Press, Cambridge

*A catalogue record for this book is available from the British Library*

ISBN 10 0 521 83333 7 hardback

ISBN 13 978 0 521 83333 2

 CAMBRIDGE  
UNIVERSITY PRESS

---

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this book, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

# The Handbook of East Asian Psycholinguistics

Volume 1: Chinese

---

Edited by

Ping Li

University of Richmond

Li Hai Tan

University of Hong Kong

Elizabeth Bates

University of California, San Diego

Ovid J. L. Tzeng

Academia Sinica, Taiwan

General Editor:

Ping Li

CAMBRIDGE UNIVERSITY PRESS  
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, Sao Paulo  
Cambridge University Press  
The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521833332](http://www.cambridge.org/9780521833332)

© Cambridge University Press 2006

This book is in copyright. Subject to statutory exception  
and to the provisions of relevant collective licensing agreements,  
no reproduction of any part may take place without  
the written permission of Cambridge University Press.

First published 2006

Printed in the United Kingdom at the University Press, Cambridge

*A catalogue record for this book is available from the British Library*

ISBN 10 0 521 83333 7 hardback

ISBN 13 978 0 521 83333 2

 CAMBRIDGE  
UNIVERSITY PRESS

---

Cambridge University Press has no responsibility for the persistence or accuracy of  
URLs for external or third-party internet websites referred to in this book, and does not  
guarantee that any content on such websites is, or will remain, accurate or appropriate.

## 22 L2 acquisition and processing of Mandarin tones

*Yue Wang, Joan A. Sereno, and Allard Jongman*

### Introduction

Mandarin tones are manifested physically by different fundamental frequency (F0) values with F0 height and F0 contour as the primary acoustic parameters (Lin, 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 et al., 1971; Moore & Jongman, 1997). Studies in the perceptual domain have shown that various acoustic cues are functionally integrated when Mandarin speakers identify the tones (Gandour, 1984; Massaro, Cohen & Tseng, 1985; Gårding et al., 1986; Blicher, Diehl & Cohen, 1990; Shen & Lin, 1991; Moore & Jongman, 1997). The hemispheric processing of Mandarin tones for native speakers is lateralized in the left hemisphere, suggesting that tones are processed as linguistic units, just like segmental properties (Hsieh et al., 2001; Wang, Jongman & Sereno, 2001).

These results about the processing of Mandarin tones for native speakers raise the question as to whether non-native speakers process 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 & Burr, 1971; Shen, 1989). A number of crosslinguistic 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.

### 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 percent for tone 4 to as low as 8.9 percent for tone 2. The error rates for tone 1 and tone 3 were 16.7 percent and 9.4 percent, 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 percent. 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 tone 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.

### L2 perception of Mandarin tones

Crosslinguistic 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, Vakoeh, 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, Vakoeh, and Wurm (1996) found that tone language speakers were better at discriminating tones, in terms of speed and accuracy of their responses, than non-tone 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 non-native 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 non-native 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.

Non-native 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 nonfinal 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 nonfinal 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 non-native 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 ( $\Delta F0$ ), a temporal parameter (Turning Point), or both ( $\Delta F0/TP$ ). 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  $\Delta F0$  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.

#### **Training perception and production of tone in the laboratory**

The research described above reveals different patterns in the production and perception of Mandarin tones by non-native speakers of Mandarin Chinese.

An interesting question is whether and how Mandarin tones can be learned by non-native 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 et al., 1982; Janieson & Morosan, 1986; Logan, Lively & 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 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 a short period of perceptual tone training, non-native speakers of Mandarin improved both their perception and production of Mandarin tones (Leather, 1990; Wang et al., 1999; Wang, Jongman & Sereno, 2003). In Wang et al. (1999), American learners of Mandarin were trained over 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, Jongman & Sereno, 2003) 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 modification of the learners' perception and production of Mandarin tones.

### **Non-native processing of Mandarin tones**

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 et al., 2001; Klein et al., 2001; Wang, Jongman & Sereno, 2001; Wang et al., 2004). These

studies further reveal that the processing of Mandarin tones by non-native 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., 2001; Klein et al., 2001), these studies reveal 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 nonlinguistic stimulus. This is consistent with the behavioral results from a dichotic listening study (Wang, Jongman & Sereno, 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 et al. (2000, 2003) 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.

### **Summary and future directions**

The studies reviewed in this chapter reveal that native and non-native 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. A source of difficulty in tone acquisition has often been attributed to non-native 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 non-native speakers' lack of experience in processing the various phonetic features characterizing Mandarin tones. Non-native speakers may weigh the various acoustic cues differently (Gandour, 1983) or may have fewer perceptual resources left to attend to contextual information (Jongman & 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, Hurtig & Ringen, 1987).

Studies of hemispheric processing have consistently demonstrated that Mandarin tones are processed differently by native and non-native 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 non-native speakers of Mandarin (Hsieh et al., 2001; Klein et al., 2001; Wang, Jongman & Sereno, 2001; Wang et al., 2004).

Although native and non-native speakers process Mandarin tones differently, non-native learners' ability to identify the tones can be significantly improved after short perceptual training in the laboratory (Wang et al., 1999). This improvement appears to generalize to new contexts, to transfer to the production domain, and to be stored in learners' long-term memory (Wang, Jongman & Sereno, 2003). Furthermore, the improvement in Mandarin tone perception and production as the result of intensive training is accompanied by a change in cortical representations in the native-like direction (Wang et al., 2003). 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 native non-native 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 these 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 determinants of the cerebral functions in bilinguals (Kim et al., 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.