L2 Acquisition of Chinese Tone

L2 acquisition of Chinese tone involves coordinated encoding of pitch information across sensory-acoustic and cognitive domains, and is shaped by linguistic and non-linguistic experiences with pitch.

Chinese languages use tones to distinguish word meaning. Tones are acoustically manifested by changes primarily in fundamental frequency (F0, perceived as pitch) as well as amplitude and duration. Native Chinese speakers process these acoustic correlates as linguistic entities in their perception and production of tonal categories. For learners whose native language (L1) is non-tonal, tones may present great difficulty, since the functional association between the acoustic characteristics and the specific linguistic tonal domain is unfamiliar to them. Research on second language (L2) tone acquisition thus focuses on the extent to which L2 tone learning involves encoding of sensory-acoustic information and how it is shaped by linguistic experience. These issues can be addressed from three perspectives: how L2 Chinese tone processing differs from native patterns, what factors affect tone learning, and what the behavioral and neural manifestations are of increased proficiency and training on the tone learning process.

1. Native and Non-native Differences

Cross-linguistic research on Chinese tone perception generally reveals that while native listeners perceive tones using linguistically relevant categorical cues, non-native listeners without any L1 or L2 tone language background mostly rely on the acoustic properties of tone. Specifically, compared to native listeners, non-natives are more sensitive to within-category F0 differences, but less accurate in discriminating Chinese tones or classifying tonal exemplars into categories (Hallé et al. 2004; Peng et al. 2010; Y. Xu et al. 2006). Native perception and non-native perception also differ in the perceptual weighting of tonal features. While native Mandarin listeners attach more importance to F0 contour (the primary tonal cue in Mandarin), non-native listeners (e.g., English, Japanese) rely more on F0 height (assumed to involve calibrating acoustic input based on internally stored F0 templates) (Gandour 1983; Huang and K. Johnson 2011; Lee et al. 2014), or they rely more on general temporal information (Cabrera et al. 2014; L. Xu and Pfingst 2003). Likewise, research on Chinese tone production has shown that (English and Cantonese) elementary-level learners of Mandarin Chinese are less accurate in identifying and reading aloud Mandarin tones (based on written tonal labels) than passively repeating those tones following auditory targets, suggesting that non-native speakers have more
difficulty in associating F0 contours with tonal categories than hearing or articulating F0 contours (Hao 2012). These findings demonstrate that non-native speakers with little or no tone language background are inexperienced in processing Chinese tones as discrete phonological categories bearing contrastive linguistic significance, as their tone perception and production are auditory- and acoustic-based. Moreover, for the perception of tones in linguistic contexts, non-native listeners, as compared to natives, are also less sensitive to surrounding phonetic and prosodic environments. For example, while native Mandarin listeners show perceptual adjustments (normalization) to compensate for F0 variations in sentential contexts, due to speaker F0 range or speaking rate, English listeners cannot efficiently incorporate such extrinsic information but rather rely on intrinsic acoustic differences of tone (Jongman and Moore 2000; Luo and Ashmore 2014). Similarly, non-native listeners fail to accurately identify Mandarin tones when the tone-bearing components (vowels) are excised from consonant-vowel syllables, as they are unable to use the coarticulatory information from the remaining consonantal contexts that has been shown to provide cues to tone identification for native listeners (Gottfried and Suiter 1997; Lee et al. 2009, 2010a). These results suggest that context-dependent tone perception requires long-term linguistic experience integrating global and localized F0 cues.

Research has also established that native and non-native Chinese tone processing involve different neural mechanisms at both the acoustic and linguistic level. Behavioral and neuroimaging studies show that such differences engage a large-scale brain network. While native tone processing primarily employs language-dominant left-hemisphere areas including prefrontal, frontal, temporal, and parietal regions, non-native processing involves more right-hemisphere or bilateral participation, similar to the processing of L1 prosodic or non-speech F0 information (Gandour et al. 2003, 2004; Hsieh et al. 2001, Klein et al. 2001; Y. Wang et al. 2001, 2004). Electro-physiological research with time-sensitive event-related brain potential (ERP) data corroborates that these differences may arise from multi-stage processing mechanisms at sensory-acoustic and linguistic levels. For instance, differences in early-evoked ERP responses to Mandarin tones primarily reflect sensitivity to F0 contour for native Mandarin listeners but to F0 height for English listeners (Chandrasekran et al. 2007), echoing the behavioral findings of experience-driven cue-weighting patterns (e.g., Gandour 1983). Further research agrees that the experience-dependent modulation of tonal information occurs at both early and late stages of processing. For Mandarin tonal contrasts and within-category F0 changes, English listeners exhibit smaller and more delayed ERP differences at the pre-attentive stage, but are more accurate than Mandarin listeners in discriminating the within-category F0 differences at the attentive stage (Chandrasekaran et al. 2009b). Thus, while native tone experience results in enhanced F0 sensitivity at an early processing stage and enhanced tonal categorization ability at a later stage, non-native tone processing appears to be
uniformly acoustically driven. Moreover, some recent research has observed that experience-dependent neural responses to F0 also occur at the sub-cortical level in the brainstem (Krishnan et al. 2005, 2012; Song et al. 2008; Wong, Skoe et al. 2007).

The native and non-native differences revealed in these behavioral and neural studies demonstrate language-specific processing of Chinese tone for native listeners, whereas non-native listeners without any tone language background primarily rely on acoustic properties in Chinese tone processing.

2. Factors Affecting Tone Learning

While non-native Chinese tone processing appears to be mediated by sensory-acoustic systems, experiential factors also affect how L2 tones are processed, particularly experience with pitch modulations both in the linguistic domain (e.g., L1 prosody) and in broader cognitive domains (e.g., musical experience).

The contribution of L1 prosody may involve enhanced sensitivity to F0 information based on the functional use of F0 in the L1. In the perception of Mandarin tones, While non-tonal L1 listeners (e.g., Dutch, Uyghur) have been shown to be more attentive to F0 movements signaling intonational variations, tonal L1 listeners (Cháng Shā 長沙 / Nán Tōng 南通 Chinese) are more sensitive to Mandarin tonal contrasts (Braun and E. Johnson 2011; Liang and Van Heuven 2007). Furthermore, tonal L1 listeners (Cantonese, Taiwanese) are better at differentiating Mandarin stimuli across categories than detecting within-category differences (Hallé et al. 2004; Peng et al. 2010). These results indicate that L1 experience with classifying F0 into categories at the phonemic level can be generalized to identifying linguistic tone boundaries in an L2. The influence of L1 prosody may also arise from acoustic similarities. Specifically, non-tonal L1 listeners categorize Mandarin tones into their L1 prosodic categories based on the perceived acoustic similarities in F0 information between Mandarin tones and their native tonal, stress, rhythmic, and intonational systems (Broselow et al. 1987; So and Best 2010, 2011; Yang and Chan 2010). While these findings imply that L1 prosodic experience may enhance tonal categorization and acoustic sensitivity, it should be noted that L1 prosodic experience might not always be facilitative. For instance, competing F0 adjustments in tone and intonation, resulting in tone-intonation clashes (e.g., a rising tone with falling intonation), pose great difficulty for native English listeners (Yang and Chan 2010). Additionally, L1 tone experience may also be inhibitory when L1 and L2 tonal properties do not match; for example, native Mandarin listeners tend to assimilate Cantonese or Thai mid/low falling tones (absent in Mandarin) into the Mandarin high falling tone category (So and Best 2010; Wu et al. 2014). Tone production research also reveals the interference of L1 prosody. Elementary-level English learners of Mandarin are shown to be less accurate in
producing the falling tone than the other tones in Mandarin, presumably because
the falling F0 contour is prosodically less marked for English speakers and is
thus more subject to incorrect substitution (Shen 1989). In sum, non-native
speakers may apply linguistic or acoustic-phonetic strategies in their Chinese
tone perception and production based on the functional use of prosody in their
native language.

In addition to linguistic background, experience with musical pitch provides an
advantage as well (Delogu et al. 2006, 2010; Gottfried 2007). Research reveals
that native English musicians outperform non-musicians in identifying Mandarin
tones, even in F0-degraded conditions, demonstrating musicians’ enhanced
sensitivity to pitch (Lee and Hung 2008). Furthermore, musical experience may
facilitate the perception of more linguistically-relevant acoustic dimensions of
tone. For example, While both American musicians and non-musicians perceive
tone height in Taiwanese equally well, musicians can better track the F0
contours (Lee et al. 2014). Consistent with this, musicality effects are associated
with the more efficient neural encoding of L2 Chinese tone. Musicians, relative to
non-musicians, exhibit larger ERP responses in discriminating tonal differences
and shorter latencies in categorizing tonal variations, as well as more robust F0
encoding in the auditory brainstem (Chandrasekaran et al. 2009a; Chobert and
Besson 2013; Marie et al. 2011; Wong, Skoe et al. 2007). These findings
implicate common behavioral and neural manifestations for music and linguistic
pitch functions, suggesting that long-term pitch exposure may shape
fundamental sensory circuitry in a domain-general manner.

Furthermore, pitch experiences from linguistic and musical backgrounds
integrate to affect L2 tone processing. While musical training provides a greater
advantage than L1 tone background for L2 tone identification, both experiential
factors equally facilitate learning tone-words in semantic contexts (that is,
learning to associate tones with word meanings rather than just identifying the
tones). However, the combination of these factors does not lead to cumulative
facilitatory effects for musicians with an L1 tone background (Cooper and Y.
Wang 2012; Mok and Zuo 2012). These results reveal that general auditory
processing mechanisms such as pitch experience from musical training
positively transfer to aid L2 tone identification. Meanwhile, while L1-specific tone
features may interfere with L2 tone identification, the general L1 tone experience
of utilizing F0 lexically to map phonetic features to meaning appears to facilitate
tone-word learning.

3. Learning Patterns

Given the contributions of these pre-existing experiential factors, the subsequent
question concerns the acquisition of Chinese as a function of L2 experience.
Research involving non-natives with different levels of Chinese proficiency
reveals that learners shift their attention to more linguistically-relevant tonal aspects as they gain experience in Chinese. For example, experienced but not naïve non-native Mandarin listeners can efficiently use the critical F0 contour cue (rather than F0 height) in Mandarin tone discrimination (Guion and Pederson 2007). Likewise, while naïve non-native listeners exhibit phonetic-level assimilation based on acoustic similarities between the L1 and L2 (e.g., assimilating the L2 Mandarin falling-rising tone to the L1 Cantonese low falling tone based on the initial falling F0 contour in Mandarin) (So 2012), more advanced learners are additionally influenced by phonological tone changes (e.g., assimilating the L2 Thai high rising tone to the L1 Mandarin falling-rising tone, influenced by their experience with a phonological change in Mandarin from falling-rising to rising tone) (Wu et al. 2014). These findings point to a progressive process of L2 tone learning, shifting from acoustic to phonemic representations in establishing new tonal categories.

Research to empirically assess how such learning occurs employs the well-established high-variability phonetic training paradigm aimed at assisting learners to establish L2 phonetic categories by exposing learners to a high variety of exemplars in a category (Logan et al. 1991). Tone perception and production training studies show that non-native tone speakers (e.g., Dutch, English, Hmong, Japanese) can better perceive and produce Chinese tones after training, with the improvements extended to new linguistic contexts and retained in long-term memory (Leather 1990; Y. Wang et al. 1999, 2003a; X. Wang 2012). Additionally, these training studies show a positive relationship between perception and production in Chinese tone learning, as training in one domain leads to improvements in the other domain (Leather 1990; Y. Wang et al. 2003a). Together, the findings indicate that training yields long-term modifications of L2 tone perception and production. Further training research demonstrates that successful learning involves redistribution of acoustic cue weighting, focusing on the primary F0 cues (e.g., contour) used in the L2 rather than the properties (e.g., height) that function more prominently in the L1 (Chandrasekran et al. 2010; Francis et al. 2008). Similarly, acoustic analyses of trainees’ productions show a greater degree of improvement in F0 contour than F0 height (Y. Wang et al. 2003a). In terms of extending phonetic-level training effects to higher linguistic domains, findings show that (Native English or Thai) learners who are better at identifying Mandarin or Cantonese tones can more successfully learn to associate the tones with word meanings in these languages suggesting a phonetic-phonological-lexical continuity for tone learning in that linguistic learning may be mediated by auditory-acoustic processing (Cooper and Y. Wang 2012; Wong and Perrachione 2007). Furthermore, training enhances tone identification when target tones appear in sentential contexts (X. Wang 2012, 2013). These data indicate that training-induced learning involves reweighting of acoustic cues as well as integrating linguistic information to develop L2 tonal categories.

The effects of multi-level experiential factors may also affect how Chinese tones
are learned during training. At the auditory level, learners’ perceptual aptitude appears to influence their tone learning abilities. For example, the high-variability training paradigm turns out to be effective only for learners with strong perceptual abilities (assessed through a pre-training non-speech pitch contour perception test), whereas those with weaker perceptual abilities benefit more from low-variability training (Perrachione et al. 2011), indicating that perceptual aptitudes may predict successful category learning. Similarly, musical training and non-linguistic-context pitch training both lead to improved word learning proficiency, suggesting that enhanced perception of acoustic-level tonal information may contribute to success in a linguistic-level tonal task (Cooper and Y. Wang 2013; Wong and Perrachione 2007). In the linguistic domain, learners’ L1 prosodic experience influences tone learning. Learners with tonal (Mandarin) and non-tonal (English) L1s reveal different training-induced changes in their perception of Cantonese tones depending on their L1 experience where Mandarin learners are more influenced by the Mandarin tonal categories and English learners are more influenced by the English intonational system (Francis et al. 2008). Moreover, tonal-L1 (Thai) learners exhibit higher post-training proficiency in Cantonese tone-words than do non-tonal (English) learners, indicating that Thai learners’ L1 experience with pitch-to-meaning associations beneficially transfers to their L2 (Cooper and Y. Wang 2012).

Brain research further reveals neuroplasticity associated with tone learning, involving recruitment of new brain resources and a shift to more efficient, native-like neural circuitry. For example, cortical effects of Mandarin tone learning by American learners involve both the expansion of language-related areas (Brodmann’s area 22) and employment of adjacent cortical regions (Brodmann’s area 42) in the left superior temporal gyrus (Y. Wang et al. 2003b). Further research also identifies the participation of the primary auditory cortex and sub-cortical brainstem in encoding pitch cues for linguistic-level learning, with increased volume in the left Heschl’s gyrus (transverse temporal gyrus) and enhanced brainstem pitch responses post-training (Song et al. 2008; Wong et al. 2008). Moreover, effective training is associated with a shift to native-like cortical representation of tones. In particular, successful tone learners exhibit increased activation in the left-hemisphere regions involved in L1 pitch processing, whereas learners with limited improvement show increased activation in the right auditory cortex responsible for non-linguistic pitch processing (Wong, Perrachione et al. 2007).

These learning patterns indicate progressive changes in behavioral performance and neural representations with training and increased proficiency, suggesting that L2 tone processing is continuously shaped by experience and learning.

4. SUMMARY AND FUTURE DIRECTIONS
Taken together, research cumulatively points to a multi-facet model of L2 tone acquisition in Chinese, encapsulating experience-dependent and domain-general factors. Specifically, Chinese tone learning appears to be modulated by multiple levels of processing across sensory-motor and cognitive domains, under the influence of linguistic, pitch, and musical experience. Furthermore, learning involves a dynamic process, as different aspects of these experiential factors are utilized at different stages of learning. The long-term experiences with pitch across domains result in neuroplasticity with respect to tone learning, reflected by changes in the temporal dynamics and cortical/sub-cortical organization of L2 tone in the brain. These findings provide support for a general theoretical account of pitch processing with integrated bottom-up and top-down processes (Zatorre and Gandour 2008).

However, it is still unclear how behavioral and neural representations are continuously shaped as learning progresses. Longitudinal research tracking tone learning trajectories is needed to define the agents of neuroplasticity at different learning stages. Additionally, while a substantial body of work has focused on the perceptual patterns of Chinese tone learning, there is little research on tone production learning. Further research must examine non-native tone production as well as the relationship between the production and perception of tone (evidenced by the training studies) in order to determine if the coordination of the articulatory and auditory-acoustic neural pathways observed in L2 segmental learning (Callan et al. 2004) is also employed in L2 tone learning. A related future direction is to also explore how visual speech input, which has been shown to facilitate L2 segmental learning (Hazan et al. 2005) and L1 tone perception (Chen and Massaro 2008), contributes to L2 tone learning. Collectively, research along these avenues will present a more complete understanding of the way in which multiple sensory-motor and cognitive systems cooperate functionally in L2 tone learning.


L2 tone acquisition; linguistic experience; music experience; pitch processing; neural plasticity

YUE WANG

ALLARD JONGMAN

JOAN A. SERENO