Tone and intonation processing: from ambiguous acoustic signal to linguistic representation

Liu, M.

Author: Liu, M.
Title: Tone and intonation processing: from ambiguous acoustic signal to linguistic representation
Issue Date: 2018-11-01

Chapter 6

General discussion

This dissertation investigated how ambiguous acoustic signals representing different prosodic information affect spoken language processing. Specifically, it investigated how pitch is processed within a linguistic system (i.e., Standard Chinese, Chapters 2 and 3) and across two linguistic systems (i.e., Standard Chinese and Xi'an Mandarin, Chapters 4 and 5) when the same pitch contour cues different linguistic functions (i.e., tone and intonation) or different categories of the same linguistic function (i.e., tone). Tone and intonation in Standard Chinese both adopt F0 as their primary acoustic correlate and therefore result in pitch processing difficulties. Chapter 2 tapped into the neural correlates of tone and intonation processing in Standard Chinese and presented ERP evidence for pitch processing costs due to the interaction of tone and intonation. Chapter 3 examined the role of semantic context in resolving pitch processing difficulties in tone and intonation processing in Standard Chinese. Chapters 4 and 5 went beyond Standard Chinese, and investigated how the tonal system of a closely related dialect of Standard Chinese (i.e., Xi'an Mandarin) affects tone processing (Chapter 4) and lexical access (Chapter 5) of bi-dialectal tonal language speakers. Together, these chapters revealed two of the most prominent pitch processing difficulties tonal language speakers encounter from within and across languages, advancing our current understanding of pitch processing from various aspects.

Chapter 2 investigated the online processing mechanisms of tone and intonation in Standard Chinese at the attentive stage using the event-related potential technique. Native Standard Chinese listeners were presented with semantically neutral Standard Chinese sentences, which contrast in final tones (rising T2 or falling T4) and intonations (Question or Statement). Their behavioral and electrophysiological responses were recorded. The behavioral results showed that, while the identification of tone was not hindered by intonation, the identification of intonation was greatly impeded due to the interference of tone. In the T4 conditions, question intonation was rather difficult to identify correctly, whereas the identification of statement intonation presented almost no difficulty. In the T2 conditions, question intonation was still difficult to identify, while identifying statement intonation also tended to be problematic. Regarding the ERP results, a clear P300 effect was observed for the question-statement contrast in sentences ending with T4, but no ERP effect was found for the question-statement contrast in sentences ending with T2. These results provide ERP evidence for the interaction of tone and intonation in Standard Chinese, confirming the findings from behavioral metalinguistic data that native Standard Chinese listeners can distinguish between question and statement intonation when the intonation is associated with a final T4, but fail to do so when the intonation is associated with a final T2.

The ERP results of tone and intonation processing at the attentive processing stage are highly consistent with findings for tone and intonation processing at the pre-attentive processing stage. Ren et al. (2009, 2013) reported that listeners are able to perceive the difference between question and statement intonation when the final tone is T4 (reflected in an MMN effect), but they cannot make a distinction between question and statement intonation when the final tone is T2 (reflected by the absence of an MMN effect) at the pre-attentive stage. These two studies used one-syllable sentences, while our study extended the length of the utterances from one syllable to five syllables. Chapter 2 therefore extends our understanding of online processing of tone and intonation 1) from the pre-attentive stage to the attentive stage; and 2) within a larger domain (i.e., multi-word utterances) than a single word utterance.

Chapter 3 tapped into the role of semantic context in resolving pitch processing difficulties in tone and intonation processing in Standard Chinese. In Standard Chinese, the F0 encodings of final lexical tone and sentence intonation can sometimes be in conflict (e.g., statements ending with a rising T2/questions ending with a falling T4) and sometimes be in congruency (e.g., statements ending with a falling T4/questions ending with a rising T2). Chapter 3 investigated how tone and intonation, with F0 encodings in conflict or in congruency, are processed and how semantic context may facilitate or inhibit their processing. Tone and intonation identification experiments were conducted in both semantically neutral and constraining contexts with the same group of native speakers of Standard Chinese. The resulting measurements included the commonly-reported response accuracy, as well as an additional measurement, reaction time.

Results showed that the overall performance of tone identification was better than that of intonation. Tone identification was seldom affected by intonation information, irrespective of semantic contexts. Participants were able to perceive tones accurately and quickly in both question and statement intonation in both contexts. Intonation identification, particularly question intonation, however, was susceptible to the final lexical tone identity and was greatly affected by the semantic context. Specifically, in the semantically neutral context, questions were difficult to identify, as evidenced in the lower response accuracy and longer reaction time, regardless of the lexical tone identity. In the semantically constraining context, both intonations took significantly less time to be identified than in the semantically neutral context. Moreover, questions ending with a falling tone were more accurately identified than questions ending with a rising tone. These results suggest that top-down information provided by the semantically constraining context can play a facilitating role for listeners to disentangle intonational information from tonal information, especially in sentences with a lexical falling tone in final position.

Chapter 3 provides strong evidence for the role of semantic context in resolving pitch processing difficulties in Standard Chinese, particularly from the reaction time patterns, which have not been reported in earlier studies. The results reported in Chapter 3 also resolved the puzzle of the reversed patterns of question intonation identification in sentences ending with T2 and T4 in normal natural context (Xu & Mok, 2012a; Yuan, 2011) versus in low-pass filtered context (Xu & Mok, 2012b). In normal natural context, questions ending with T4 were more accurately identified than questions ending with T2 (Xu & Mok, 2012a; Yuan, 2011), whereas in low-pass filtered context, questions ending with T4 were less accurately identified than questions ending with T2 (Xu & Mok, 2012b). These reversed patterns could be due to different factors, and the results in Chapter 3 showed that context is a significant factor. It was found that the stronger and more informative the linguistic context is (semantically constraining context > semantically neutral context > low-pass filtered context), the better the identification of questions ending with T4. The opposite pattern holds for questions ending with T2.

Chapter 4 was concerned with cross-linguistic pitch processing. One long-neglected fact in linguistic research on Standard Chinese is that most speakers of Standard Chinese also speak a local dialect (Li & Lee, 2008; Wiener & Ito, 2014), which may share phonological features with Standard Chinese. Tonal information can be a determinant of the phonological similarities or differences between Standard Chinese and regional dialects, yet relatively little empirical research has been conducted on the tonal system of other language varieties spoken in China aside from Standard Chinese. Among these dialects, Xi'an Mandarin is particularly interesting for the seemingly simple, yet intricate mappings between its lexical tones with those in Standard Chinese (Li, 2001; Zhang, 2009).

In Chapter 4, the tonal systems of Xi'an Mandarin and Standard Chinese were compared empirically. Tones with similar contours from Xi'an Mandarin and Standard Chinese were paired, and both tone production and perception experiments were carried out on highly proficient bi-dialectal speakers of Xi'an Mandarin and Standard Chinese. Acoustic results showed that the F0 difference ranged from no F0 difference (level contour tone pair) through F0 curvature difference (rising contour tone pair) to F0 height difference (falling contour tone pair) and F0 contour difference (low contour tone pair). Except for the falling contour tone pair, all the other tone pairs also exhibited differences in tone duration. The varying acoustic differences in different tone pairs, together with the phonological rule, resulted in varying degrees of tonal similarity in tone perception. Specifically, the rating tendency of the tone pair of low contour was significantly different from that of the other three tone pairs. The low contour pair was judged as either different or the same, with slightly more same responses than different responses, whereas the latter three tone pairs were mostly judged as the same. That said, the tone pair of falling contour elicited more different responses than the tone pairs of level contour and rising contour. Overall, tones with similar contours between the two dialects were basically perceived to be the same. The two experiments together showed that there are indeed systematic mappings of tones between Xi'an Mandarin and Standard Chinese.

Chapter 4 compared the two systems in tone production with a more balanced design compared to the only previous acoustic study (Zhang, 2009). Moreover, it provides new empirical evidence for the mapping of the two tonal systems from a perceptual point of view. It also allows for an investigation of the relationship between tone production and perception in bi-dialectal tonal language speakers.

Chapter 5 further examined the effects of cross-dialect phonological similarity in segment and tone on bi-dialectal lexical access in spoken word recognition. The systematic mapping of tones between Standard Chinese and Xi'an Mandarin, as shown in Chapter 4, together with the large overlap of segmental features between the two dialects, makes cross-dialect homophones prevalent in the two languages. Cross-dialect minimal tone pairs (i.e., syllables sharing the segmental structure but not tonal contour) are also common in Standard Chinese and Xi'an Mandarin. Using an auditory-auditory priming paradigm, Chapter 5 investigated the effects of cross-dialect phonological similarity in segment and tone on auditory word recognition in a bi-dialectal context (i.e., Standard Chinese and Xi'an Mandarin).

Balanced bi-dialectal speakers of Xi'an Mandarin and Standard Chinese took part in an auditory-auditory priming experiment with a generalized lexical decision task in Chapter 5. The primes were monosyllabic homophones from either Xi'an Mandarin or Standard Chinese, while the targets were disyllabic Xi'an Mandarin or Standard Chinese words. Primes and the first syllable of the target words had five configurations. They either overlapped in both segment and tone within a dialect (identical) or across two dialects (interdialectal homophones), or they overlapped in segment only within a dialect or across two dialects. The baseline condition was that they overlapped neither in tone nor segment within a dialect. Results showed that Standard Chinese primes did not yield significant priming effects for within- or cross-dialect segment-only overlap targets. Standard Chinese primes did not produce significant priming effects for within-dialect identical targets either. However, they did yield significant inhibitory priming effects for cross-dialect homophone targets. This overall pattern was reversed for Xi'an Mandarin primes because these primes were not treated differently from their interdialectal homophonous primes in the current mixed dialect setting. These results suggest that cross-dialect phonological similarity in segment alone does not affect lexical access in bi-dialectal auditory word recognition, while cross-dialect phonological similarity in both segment and tone poses a threat to the recognition system of bi-dialectal listeners. It is clear that tonal information plays a significant role in constraining word activation in bi-dialectal auditory word recognition.

The results reported in Chapter 5 extend our understanding of the role of segment and tone in auditory word recognition in tonal languages from the monolingual context to the bi-dialectal context, and reveal a non-selective processing mechanism in bi-dialectal lexical access during auditory word recognition, as has been demonstrated for bilingual lexical access.

In summary, this dissertation has demonstrated that pitch processing in Standard Chinese is subject to both within- and cross-linguistic influences. The ambiguous acoustic signals due to dual functions of the F0 channel in signalling tone and intonation in Standard Chinese cause pitch processing difficulty at the sentential level. This pitch processing difficulty has a neural correlate and can be resolved via top-down information provided by a constraining semantic context. Acoustic ambiguities in Standard Chinese can also arise from a closely related Chinese dialect that shares tonal similarities with Standard Chinese, here Xi'an Mandarin. The cross-dialect tonal similarities affect tone processing and further interfere in lexical access during spoken word recognition in bi-dialectal tonal language speakers.

The present research on tone and intonation processing in Standard Chinese may contribute to the potential typology of the interaction between tone and intonation in tonal languages. Simply comparing the results of the present study with those Cantonese studies has demonstrated different mechanisms of tone and intonation interaction. As shown from the neural correlates, in Standard Chinese, the interaction of tone and intonation leads to difficulties in intonation processing (Liu et al., 2016b), whereas in Cantonese, it is tone processing rather than intonation processing that is problematic for native listeners (Kung et al., 2014). It seems that in tonal languages when tone and intonation interact, whether tone or intonation causes processing difficulties can be language-dependent. Nevertheless, the pitch processing difficulties in both languages (tone or intonation) can be resolved via top-down information provided by a constraining semantic context.

The pitch influence of a closely related dialect on Standard Chinese has received little attention in previous studies. In this study, it was found that bi-dialectals of Standard Chinese and Xi'an Mandarin activate lexical candidates from both dialects when presented with a cross-dialect homophone, but with different activation levels. The results suggest that future studies on Standard Chinese tones may have to control for participants' dialect background to avoid potential tonal influence from dialects. Moreover, although the bi-dialectals we recruited are comparable in their language competence of the two dialects, they tend to be Standard Chinese dominant after all. Since it is almost unlikely to find more balanced bi-dialectals than the participants investigated here, the result suggests that truly balanced bilinguals or bi-dialectals may never exist. Language users always have a preference for one or the other language or dialect. Overall, this dissertation investigated two of the most prominent pitch processing difficulties that tonal language speakers encounter from within and across languages. Several implications can be made for future research.

First, Chapters 2 and 3 exclusively investigated the interaction of tone and intonation in Standard Chinese when the final tone is T2 or T4. It would be instructive to include other tones to gain a fuller picture of the interaction of tone and intonation in Standard Chinese. Some research questions worth investigation include how intonation modulates acoustically similar tones (e.g., T2 and T3) in sentence-final position, how such modulation affects tone and intonation processing, and what the underlying neural mechanisms are.

Second, apart from tone and intonation, F0 also cues other linguistic functions, such as focus, in Standard Chinese. Future research might consider investigating how other linguistic functions of F0 affect the interaction of tone and intonation in Standard Chinese, and how the different information is encoded in the minds of native speakers. Such investigations could enable us to better understand the F0 capacity Standard Chinese employs for each of these prosodic functions, which, in turn, may shed light on the acoustic modelling of Standard Chinese and neurobiological studies of language prosody in general.

Third, Chapter 4 and Chapter 5 investigated cross-dialect tone processing and bi-dialectal word recognition, which is still an understudied research area. More parallel studies from other tonal dialects should be conducted to verify and deepen current understanding. Future research should also examine bi-dialectal word recognition in more detail. Comparative studies of monolingual, bi-dialectal and bilingual word recognition should be carried out to reveal their similarities and differences. Moreover, models of spoken word recognition should take the bi-dialectal situation into account.