

Phonology and Phonetics

12-2

Editor

Aditi Lahiri

Mouton de Gruyter
Berlin · New York

Tones and Tunes

Volume 2: Experimental Studies
in Word and Sentence Prosody

edited by

Carlos Gussenhoven
Tomas Riad

Mouton de Gruyter
Berlin · New York

2007

Neural substrates underlying the perception of linguistic prosody

Jackson T. Gandour

Abstract

Dichotic listening, lesion deficit, and brain imaging data are brought to bear on issues related to the neural substrates underlying the perception of prosodic units at the syllable, word, and sentence levels of representation. These data show that both hemispheres are recruited in the processing of suprasegmental features of speech. In addition to the left hemisphere, dichotic listening data also implicates the RH in processing phonemic stress and intonation. More brain-damaged patients, phonemic stress, sentence focus and intonation. More recent brain imaging data show that all prosodic units of representation investigated so far engage the RH. Brain imaging data moreover shows us that we have to go beyond linguistic units themselves to gain a full understanding of how linguistic prosody is processed in the brain. Tone, stress, focus, and intonation phenomena emerge from more general sensory-motor and cognitive processes in addition to those associated with language. There is no cognitive or linguistic function that is mediated by a single region or even a single hemisphere, and moreover, the networks are not necessarily unique to language. A complete account of language processing must allow for multiple dichotomies or scalar dimensions that either apply at different time intervals or interact within the same time interval at different cortical or subcortical levels of the brain.

1. Introduction

Speech prosody refers to melodic or suprasegmental features of an utterance (e.g., pitch) that may signal a variety of linguistic and paralinguistic functions (Lehiste 1996). There has been increasing interest in the functional neuroanatomy of speech prosody in the brain. This chapter provides a critical evaluation of dichotic listening, lesion deficit, and brain imaging data in regards to the perception of linguistic prosody. Prosodic units of interest include tone, phonemic stress, sentence focus, and sentence modality (declarative vs. interroga-

tive). By examining various aspects of speech prosody, we are able to compare neural processes that subserve syllable-, word-, and sentence-level prosodic frames. In so doing, we gain a more comprehensive view of linguistic prosody and the brain. We exclude production data since only perception data are available from all three sources. Data on affective prosody are also excluded so we can narrow the scope of inquiry to the linguistic domain (see Baum and Pell 1999; Gandour 1987, 1998b; Wong 2002, for earlier reviews of speech prosody and the brain).

2. Various approaches to hemispheric specialization of linguistic prosody

Hemispheric asymmetry underlying higher cortical functions is well established in the brain, but elucidating the neural mechanisms underlying the functional asymmetry of speech remains a controversial issue. Divergent views have emerged based on premises derived from studies of complex auditory signals ranging from nonspeech to speech. Some argue in favour of domain-specific circuits in the left hemisphere (LH) comparable to species-specific vocalizations as in animals (Benson et al. 2001), whereas others contend that LH lateralization for speech reflects cue-specific processing of temporal or spectral cues that happen to be crucially involved in speech perception (Ivry and Robertson 1998; Poeppel 2003; Tallal, Miller, and Fitch 1993). One view claims that all acoustic cues of prosody are processed in the right hemisphere (RH) (Klouda et al. 1988). These competing hypotheses are not necessarily mutually exclusive, and we will argue that both are essential in order to give a full account of the perception of linguistic prosody in the brain.

3. Windows on linguistic prosody from the perspective of tone languages

Languages that exploit phonologically relevant variations in pitch at the syllable level are called *tone languages* (Gandour 1994; Yip 2003). This chapter evaluates three major sources of evidence relevant to the neural substrates underlying the processing of lexical tone: dichotic listening; lesion deficits; brain imaging. Because pitch information is relevant in monosyllables, such languages are especially advantageous for comparing syllable-, word-, and sentence-level aspects of prosody. By comparing tone, phonemic stress, sentence focus and intonation, we can better conceptualize what cognitive and linguistic neural processes are shared across hierarchical levels of prosodic representation. Background phonetic information is given below on Mandarin

Chinese and Thai since the bulk of work on tonal processing in the brain to date comes from these two languages.

3.1. Mandarin Chinese

Mandarin Chinese is a tone language. In addition to consonants and vowels, Chinese has four tones: *ma*¹ 'mother', *ma*² 'hemp', *ma*³ 'horse', *ma*⁴ 'scold'. Tones 1 to 4 can be described phonetically as high level, high rising, low falling rising, and high falling, respectively (Howie 1976). Voice fundamental frequency (F_0) contours provide the dominant cue for tone recognition (Xu 1997); other acoustic cues include amplitude (Whalen and Xu 1992) and duration (Fu et al. 1998). Perceptual data on Mandarin tones indicate that variation in F_0 yields high levels of recognition for isolated tones (Howie 1976).

Acoustic analyses of stress location in Mandarin reveal systematic changes in F_0 and duration (Xu 1999). There are three distinct focus-related F_0 ranges that occur in both statements and questions: *expanded* in non-final stressed words; *suppressed* in words after the stress; and *neutral* in all other words. Syllable duration increases significantly under focus regardless of location. Sentence focus modulates the global shape of the F_0 trajectories, which, in turn, affects the height and even the shape of local F_0 contours.

In Mandarin *unmarked* yes-no questions (i.e., those involving no change in word order and no addition of a question word), F_0 contours appear to start higher than in their statement counterparts, and generally move upward throughout the duration of the sentence (Shen 1990; Xu 2006).

3.2. Thai

Thai is a tone language. It has five tones: *k'aa*^M 'stuck', *k'aa*^L 'galangal', *k'aa*^F 'kill', *k'aa*^H 'trade', *k'aa*^R 'leg'. The mid tone (M) can be described phonetically as mid level with a final drop, low tone (L) as low falling, falling tone (F) as high falling, high tone (H) as high rising, and rising tone (R) as low rising (Tingsabhadh and Abramson 1993). Perceptually, F_0 contours have been shown to be the primary acoustic correlate of Thai tones (Abramson 1962).

4. Sources of evidence on linguistic prosody and the brain

4.1. Dichotic listening

With the dichotic listening technique, two different auditory stimuli are presented at the same time, one in each ear (Hugdahl 1999). Dichotic presentation

of verbal auditory stimuli typically yields a right ear advantage (REA) when subjects are requested to report what they hear on each trial. The standard explanation for the REA is that the contralateral auditory pathways suppress the ipsilateral pathways at the level of the brain stem, thus favouring the right ear input to the language-dominant LH. Dichotic listening studies tell us which hemisphere predominates for a given function.

4.1.1. Tone

The seminal dichotic perception studies of tone focused on Thai (Van Lancker and Fromkin 1973, 1978). English speakers, either musically untrained or trained, served as controls. Stimuli consisted of a minimal set for tones in real Thai words, and a minimal set of hums with homologous pitch contours. A REA was elicited from tones in real words in the Thai group only. For the hums, no ear advantage was observed for either group. These findings are interpreted to suggest that pitch perception is lateralized to the LH when pitch variations signal language-specific functions. The absence of a REA in response to hums for the Thai group, however, may be confounded with lexical-semantic processing (Wong 2002) because the condition in which the Thai subjects listened to Thai words was the only one in which subjects heard meaningful words. Thus, results might be attributed to a word listening effect rather than a tone perception effect.

A LH superiority for dichotic perception of tone has also been demonstrated in Mandarin (Wang, Jongman, and Sereno 2001). A REA for tone in real Mandarin words was observed in Mandarin but not English listeners. In a companion study (Wang et al. 2004), four groups of listeners were examined: Mandarin, English-Mandarin bilinguals, Norwegian, and English. Results showed LH dominance of Mandarin tone by native and proficient bilingual listeners only. Despite their familiarity with tonal contrasts in Norwegian, no ear advantage was found in this group for processing Mandarin tones. Yet Norwegian listeners show a similar REA for Norwegian words differing minimally in tone (Moen 1993). These findings together reinforce the view that lateralization of pitch processing varies depending on language experience.

4.1.2. Word stress and sentence intonation

Three separate dichotic listening experiments were conducted to determine ear advantage for phonemic stress identification (Behrens 1985). Stimuli consisted of minimal pairs of real words (e.g., *hótlog* vs. *hot dóg*), nonsense words (e.g., *bógog* vs. *bot góg*), and low-pass filtered tokens which eliminated phonetic

and semantic information but preserved the stress pattern. A REA was obtained when both phonetic and semantic information are present (real words); no ear advantage when semantic content is absent (nonsense words); and a LEA when phonetic and semantic information is removed (filtered tokens). These results point to a dynamic interplay between the two hemispheres. The LH may be engaged for prosodic features when these features are linguistically relevant; otherwise, the RH predominates in the processing of prosody.

This notion of a dynamic interplay between the two hemispheres is supported by dichotic listening studies of the perception of intonation. A LEA was reported for both filtered and unfiltered intonation contours in nonsense utterances (e.g., *pa da ka*) (Blumstein and Cooper 1974). In the absence of syntactic or semantic information, prosody engaged the RH. This finding of a LEA was extended to affective and linguistic intonation contours in the context of natural speech (Shipley-Brown et al. 1988).

4.1.3. Summary

Dichotic listening yields a LH superiority for tone and phonemic stress in natural speech stimuli. Degrading the speech signal, i.e. removing semantic and/or phonetic information, results in no ear advantage or a shift in laterality to the RH. A RH superiority is seen for intonation regardless of whether the stimuli are natural or not. Because of potential confounds of lexical semantics and/or stimulus duration across studies, in addition to limitations of the experimental technique itself, it is difficult to isolate what linguistic, cognitive, or perceptual function is triggering the laterality effects. Nonetheless, the RH does appear to be engaged in processing prosody-specific cues.

4.2. Lesion deficits

Most of what we know about language and the brain has been drawn from studies of left-hemisphere-damaged (LHD) patients. Damage in specific areas may be responsible for specific language deficits, or *aphasias* (Dronkers, Pinker, and Damasio 2000). Lesion deficits tell us that the damaged area is crucially involved in performing a linguistic function. As in non-tone languages (Gandour 1998a; Packard 1993, 2006).

4.2.1. Tone

Reports on speech perception deficits after brain damage consistently point to the LH as being dominant for the perception of lexical tones (cf. Wong 2002). Naeser and Chan (1980) found that the perception of Mandarin tones was impaired for their LHD Mandarin Chinese-Cantonese bilingual aphasic. No disturbances in Mandarin tonal identification were reported for right-hemisphere-damaged (RHD) patients (Hughes, Chan, and Su 1983). In Cantonese tone perception (Yiu and Fok 1995), LHD aphasics performed significantly worse than either normals or dysarthrics. Similar results have been reported for the perception of tones in Thai (Gandour and Dardarananda 1983), Toisanese (Eng et al. 1996), and Norwegian (Moen and Sundet 1996).

4.2.2. Word stress, sentence focus, and sentence intonation

Both hemispheres are implicated in the perception of word-level stress and sentence-level focus. Both LHD and RHD, LHD only, or RHD only were impaired relative to normals in the identification of phonemic and sentence focus (Brådvik et al. 1991; Bryan 1989; Emmorey 1987; Kimelman 1991; Pell 1998; Walker, Daigle, and Buzzard 2002; Weintraub, Mesulam, and Kramer 1981). Noteworthy is a study in which both LHD and RHD patients performed worse than normals in phonemic stress identification, but LHD patients' performance was worse than that for RHD (Baum 1998). In sentence focus identification, only LHD patients' performance was impaired relative to normals.

Intonation perception is also vulnerable to disruption following unilateral lesions to either hemisphere. In response to normal speech, RHD and LHD patients (Brådvik et al. 1991; Bryan 1989; Weintraub, Mesulam, and Kramer 1981) performed worse than normals in the identification and discrimination of sentence modality (statement vs. question). Other data suggest that LHD patients experience more difficulty than RHDs (Perkins, Baran, and Gandour 1996; Walker, Daigle, and Buzzard 2002).

The use of *prosodic speech* (low-pass filtered), renders sentences free of phonetic, semantic, and syntactic content, while preserving suprasegmental cues associated with the sentence melody. In response to prosodic speech, both RHD and LHD patients exhibited deficits in the identification of the linguistic modality of the utterance (declarative, interrogative, imperative) (Heilman, Scholes, and Watson 1984). Similar findings were found regardless of speech type (normal; syntactic/Jabberwocky; prosodic) (Pell and Baum 1997). In prosodic speech, RHD patients exhibited difficulty in discriminating linguistic modality relative to LHD and normal control groups (Perkins, Baran, and Gandour 1996).

In the identification of prosodic information used in syntactic disambiguation, LHD patients performed worse than RHD patients and normal controls (Perkins, Baran, and Gandour 1996; Walker, Daigle, and Buzzard 2002). Baum and Dwivedi (2003), however, found that RHD patients were largely insensitive to acoustic cues that span the sentence unit as a whole (cf. Walker, Fongemie, and Daigle 2001).

4.2.3. Summary

Whereas tonal deficits are found in LHD patients only, stress and intonation deficits occur in both LHD and RHD patients. All findings on tone perception have been based on identification tests only. To isolate whether the deficit is linguistic or perceptual in nature, tests of tonal discrimination are also necessary. Interpreting data from brain-damaged patients is difficult because multiple factors may have contributed to their relatively poor comprehension of tone, word stress, or sentence focus and intonation. Variation in site and extent of brain damage among patients increases the likelihood that these deficits most likely reflect a complex interaction of lower-level perceptual and higher-level linguistic and/or cognitive processes. While lesion-symptom data may potentially reveal information about a given region within a hemisphere, 'voxel-based symptom mapping' (Bates et al. 2003), a method for analyzing the relationship between tissue damage and behaviour on a voxel-by-voxel basis,¹ offers greater promise of revealing how multiple regions contribute to prosody processing in a widely distributed functional network.

4.3. Brain imaging

Brain imaging studies give us a window for investigating spatiotemporally distributed language processing in vivo in the normal human brain. Imaging techniques can provide answers to *where* processing occurs in the brain, as measured for instance with fMRI (functional magnetic resonance imaging) as well as when processing occurs in the brain, as measured for instance with EEG (electroencephalography), ERP (evoked response potentials), or MEG (magnetoencephalography) (see Rugg 1999, for review of imaging methods). Whereas imaging studies of linguistic prosody began to appear only recently (Gandour 1998b), we are now able to draw a more coherent picture of the neural substrates supporting the processing of tone, focus, and sentence modality. The studies to date have addressed, among others, issues related to hemispheric laterality underlying the processing of different prosodic units, the relative importance of acoustic and functional attributes in driving laterality effects, the

isolation of prosodic influences independent of lexical-semantics, the extent to which there are common neural substrates underlying suprasegmental and segmental processing, and the influence of language experience on linguistic prosody at subcortical as well as cortical levels of the brain.

4.3.1. Tone

Tone languages give us an opportunity to examine how language experience influences the way pitch is processed in the brain. PET (positron emission tomography) and fMRI studies of tone perception in Mandarin (Hsieh et al. 2001; Klein et al. 2001) and Thai (Gandour et al. 2000; Gandour et al. 2002) have consistently shown that the LH is engaged for tonal processing. In cross-language comparisons, native speakers of tone languages relative to non-tone language controls (English) showed activation in left posterior frontal regions. For English listeners, these pitch variations are not linguistically relevant. Similarly, when judging Thai tones, Chinese listeners failed to show activation of LH regions (Gandour et al. 2000; Gandour et al. 2002) even though they are familiar with lexical tones from their own language. Thus, pitch processing engages the LH only when the pitch patterns are phonologically significant to the listener; otherwise, they are lateralized to the RH. A direct comparison of normal speech and prosodic speech yielded no significant differences in brain activity for the Chinese group, suggesting that Chinese listeners may be extracting linguistically-relevant pitch features at a prelexical stage of processing (cf. Wong et al. 2004). The perception of Mandarin onsets and rhymes is similarly lateralized to the LH (Hsieh et al. 2001). We therefore infer that LH mechanisms mediate processing of linguistic information irrespective of acoustic cues or type of phonological unit (i.e., segmental or suprasegmental).

Another point of view is that functional asymmetries can be accounted for by computational demands evoked by complex auditory signals that contain critical spectral (RH) or timing (LH) cues. In a cross-language study (Gandour et al. 2002), Thai and Chinese subjects made same-different judgments of pitch (e.g., $p^h in^R p^h in^R$ – same) and timing (e.g., $p^h in^R p^h in^R$ – different) patterns presented in the same auditory stimuli in two contexts: linguistic (Thai speech) and nonlinguistic (nonspeech hums). For Thai listeners, the left inferior prefrontal cortex was activated in speech minus nonspeech contrasts for both tone and vowel length. No matter whether the phonological contrast is signaled primarily by spectral (Thai tones) or temporal cues (Thai vowel length: e.g., /bat^h/ ‘card’ vs. /baat^h/ ‘monetary unit’), the LH appears to be dominant in processing phonologically relevant distinctions in a listener’s native language.

Because tones necessarily co-occur with real words, it has been argued that the aforementioned cross-language differences in hemispheric laterality reflect nothing more than a lexical effect (Wong 2002). This issue was addressed directly in a cross-language (Mandarin, Thai) fMRI study by creating hybrid stimuli in which Thai tones were superimposed onto Mandarin syllables (*tonal chimeras*: e.g., $zhu^R xiao^R$ – same) and Mandarin tones onto the same syllables (real words: e.g., $zhu^3 xiao^2$ – different) (Xu et al. 2006). In a comparison of native vs. non-native tones, overlapping activity was identified in the left planum temporale (Fig. 1). In this area, a double dissociation between language experience and neural representation of pitch occurred such that stronger activity was elicited in response to native as compared to non-native tones. In the absence of a lexical-semantic confound, we argue that this neural activity is related to the prelexical processing of tones. It now appears that relatively early stages of acoustic-phonetic processing in auditory cortex can be modulated by stimulus features that are phonologically relevant in particular languages (Griffiths and Warren 2002).

Because tones are suprasegmental in nature, another question that naturally arises is whether they are processed similarly to segmental units in the brain? In a cross-language (Chinese, English) study (Gandour et al. 2003b), same-different judgments of Mandarin onsets, rhymes, and tones were compared in non-minimal vs. minimal pairs. For the Chinese group, comparisons of non-minimal and minimal pairs showed that frontal and parietal activity varies across phonological units. Moreover, bilateral activation of the posterior inferior frontal gyrus was seen only in response to tones. Rhymes induced greater activation in the left posterior middle frontal gyrus when compared to onsets and tones. These findings point to differentiated cortical networks underlying the processing of segmental and suprasegmental units.

While it is important to identify neural activity at the cortical level, a complete understanding of the neural organization of language can only be achieved by viewing language processes as a set of computations or mappings between representations at different stages of processing (Hickok and Poeppel 2004). The question arises whether early, pre-attentive stages of pitch processing at subcortical levels may also be shaped by language experience. At the level of the brainstem, electrophysiological responses were recorded from native speakers of Mandarin and English while they listened passively to the four Mandarin tones (Krishnan et al. 2004, 2005). Pitch strength and accuracy of pitch tracking revealed that the Chinese group exhibits stronger pitch representation and smoother pitch tracking than the English group. We conclude that this neural plasticity at the brainstem level is induced by language experience and may be enhancing or priming linguistically-relevant features of pitch con-

tours. The broader significance of this finding is that language-specific or even speech-specific operations may begin before the signal reaches the cerebral cortex.

Whether tonal processing engages a general-purpose, spatially distributed network that subserves other linguistic and cognitive functions is another question that can advance our understanding of how prosodic units are processed in the brain. One general-purpose cognitive component is attention. In a direct comparison of tone vs. syllable recognition, activation was observed in left frontal and parietal areas (Li et al. 2003). Interestingly, these same regions have been implicated in attention-related processing independent of domain or modality (Corbetta et al. 2000; Shaywitz et al. 2001). It is therefore likely that tonal extraction involves the use of attention-related executive functions that are differentially sensitive to internal dimensions of a whole stimulus. Another general-purpose cognitive component is immediate memory. In a comparison of tone recognition at the end vs. the beginning of a memory list, increased activations were found in the left inferior frontal gyrus, right lateral cerebellum, and medial frontal gyrus for the target tone in first position (Li et al. 2004). These regions are believed to be part of a neural network that subserves articulatory encoding of auditory input for immediate memory irrespective of the phonetic class (segmental or suprasegmental) of the linguistic unit.

Because lexical tones do not occur in English, the question arises as to how learning a tone-based second language (Mandarin) affects cortical processing of pitch (Wang et al. 2003). Neural changes associated with Mandarin tone training were found in the left superior temporal gyrus and right inferior frontal gyrus. Recruitment of the former enhances a preexisting language-related region, the latter, a pitch-related region (Zatorre, Mondor, and Evans 1999). The RH activation suggests that during the early stages of learning Mandarin tones, native speakers of English acquire a novel function by enhancing an existing system originally specialized for the processing of nonlinguistic tone.

To date, one electrophysiological study has examined the temporal dynamics of tonal processing in the cerebral cortex (Schirmer et al. 2005). Results indicate that tone and segmental information have a similar impact on semantic processing and that the processing of both information types has a comparable time course.

4.3.2. Tone and intonation

The question now arises whether the LH dominance for tonal processing extends to other prosodic units. A cross-language (Chinese, English) fMRI study (Gandour et al. 2003a) investigated discrimination of tone and intonation in

paired 3-syllable nonsense utterances. Results showed that pitch contours associated with syllable-based tones are processed in the LH by Chinese listeners only, whereas pitch contours associated with sentence-based intonation are processed predominantly in the RH. In a direct comparison of intonation and tone, only the former is found to specifically engage RH mechanisms in fronto-temporal regions. These findings argue against the view that all aspects of speech prosody are lateralized to any single hemisphere.

Tone and intonation, however, span over shorter and longer temporal domains, respectively. In another cross-language (Chinese, English) fMRI study (Gandour et al. 2004), the aim was to determine whether the temporal interval in which prosodic units occur has an influence on how speech prosody is processed in the brain. Subjects were asked to make same-different judgments of Mandarin intonation and tone presented in both 1-syllable and 3-syllable utterance pairs. Regardless of prosodic unit or temporal interval, Chinese listeners exhibited greater activity than English listeners in a left inferior parietal region. Only Chinese listeners exhibited a leftward asymmetry in inferior parietal and superior temporal regions. Both language groups, however, shared a rightward asymmetry in the mid portions of the superior temporal sulcus and middle frontal gyrus (Fig. 2). Thus, Mandarin tone and intonation are best thought of as a mosaic of multiple local asymmetries that allows for the possibility that different regions may be differentially weighted in laterality depending on language-, modality- and task-related features. Not all aspects of speech prosody are lateralized to the RH. Cross-language differences in laterality of particular brain regions depend on a listener's implicit knowledge of the relation between external stimulus features (auditory) and internal conceptual representations (linguistic). All regions in the frontal, temporal, and parietal lobes that are lateralized to the LH in response to all tasks or subsets thereof are found in the Chinese group only. LH laterality is construed to reflect higher-order processing of internal representations of Chinese tone and intonation, RH laterality lower-order processing of complex auditory stimuli.

These findings from a tone language converge with fMRI data on sentence intonation perception in non-tone languages (German, English). In prosodic speech, RH regions predominate; whereas in normal speech, lexical and syntactic processing elicits activity in LH areas (Meyer, Alter, and Friederici 2003; Meyer et al. 2002). In high memory load tasks that result in recruitment of frontal lobe regions, a rightward asymmetry is found for prosodic stimuli, a leftward asymmetry for sentence processing (Plante, Creusere, and Sabin 2002).

4.3.3. Focus and intonation

Only a couple of brain imaging studies have been conducted to compare brain activity between *linguistic* focus and intonation perception (Dronkers et al. 2004; Wildgruber et al. 2004). In (Tong et al. 2005), Chinese and English listeners were asked to make same-different judgments of focus and intonation in pairs of 3-word sentences. Between-group comparisons revealed that the Chinese group exhibited significantly greater activity in the left supramarginal gyrus and posterior middle temporal gyrus relative to the English group for both focus and intonation (Fig. 3). These same two regions showed a leftward asymmetry for focus in the Chinese group only. Converging evidence implicates the left supramarginal gyrus in phonological processing (Jacquemot et al. 2003), the posterior middle temporal gyrus in word level comprehension (Dronkers et al. 2004). Activation in the left posterior middle temporal gyrus may occur because focus requires attention to individual words of the sentence, whereas intonation directs attention to the whole sentence (cf. Dogil et al. 2002). The rightward asymmetry in the middle frontal gyrus, irrespective of the type of sentence-level prosody or language experience, more likely implicates general auditory attention and working memory processes associated with pitch perception (Zatorre, Mondor, and Evans 1999). This area is not domain-specific since it is similarly recruited for extraction and maintenance of pitch information in processing music (Koelsch et al. 2002; Zatorre, Evans, and Meyer 1994). These findings are consistent with the emerging view that speech prosody perception more generally involves a dynamic interplay among widely distributed regions not only within a single hemisphere but also between the two hemispheres.

4.3.4. Summary

Brain imaging data reveal that linguistic prosody is processed in neither a single region nor a specific hemisphere, but engages multiple areas comprising a large-scale spatially distributed network in both hemispheres (Friederici and Alter 2004). Cross-language comparisons provide unique insights into the functional roles of different areas of this cortical network that are recruited for processing different aspects of linguistic prosody (e.g., auditory, phonological). By comparing tone, focus, and intonation tasks, we are able to distinguish hemispheric roles of areas sensitive to selected aspects of linguistic processing (LH) from those engaged in auditory processing (RH). Phonetic extraction of tones reveals a cortical network in the LH that recruits mediational processes of attention and working memory. These mediational components apply to language as well as to other domains. Lateralization is clearly influenced by

language experience which shapes the internal prosodic representation of an external auditory signal. Indeed, we now have evidence that neural plasticity induced by language experience not only occurs at the cortical level but also at the level of the auditory brainstem. A complete understanding of the neural organization of language can only be achieved by adopting the view that language processing reflects a set of computations or mappings between representations at different stages of processing (Hickok and Poeppel 2004). In the perception of linguistic prosody, early stages of processing on the input side may perform computations on the acoustic data that are relevant to *linguistic* as well as non-linguistic auditory perception. The picture of linguistic prosody that emerges is a mosaic of multiple regional asymmetries that allows for different regions being differentially weighted depending on language experience, stimulus properties, and cognitive processes evoked by task demands. Prosodic analysis of complex sounds is mediated primarily by RH regions, but laterality shifts to task-dependent regions in the LH when language processing is required.

5. Overall summary and concluding remarks

These findings together illustrate how far we have progressed in our understanding of hemispheric specialization for linguistic prosody (Table 1). The major advance lies in our assessment of the role of the RH. From dichotic listening, the RH is implicated for processing phonemic stress and intonation; from lesion deficits, phonemic stress, sentence focus and intonation; and from brain imaging, all prosodic units of representation. Brain imaging data, moreover, shows us that we have to go beyond suprasegmental units themselves to gain a full understanding of how linguistic prosody is manifested in the brain. Any attempt to map high-level discrete linguistic units onto the brain is a misguided heuristic strategy. From a physiological perspective, the brain is an analog device. Tone, stress, focus, and intonation phenomena emerge from more general sensory-motor and cognitive processes, in addition to those associated with language. For example, activation of frontal and temporal areas in the RH appear to stem from their roles in mediating pitch processing irrespective of domain. Moreover, brain imaging draws our attention to networks instead of isolated regions. There is no cognitive or linguistic function that is subserved by a single region or even a single hemisphere, and moreover, the networks are not necessarily unique to language. A conceptualization of the brain as being driven by just one simple dichotomy is inadequate. A complete account of language processing must allow for multiple dichotomies or scalar

dimensions that either apply at different time intervals or interact within the same time interval at difference cortical or subcortical levels of the brain. We have also seen how various executive functions related to attention and working memory mediate prosody perception, in addition to stimulus characteristics, task demands, and listeners' experience. By applying different brain imaging techniques, we can look forward to significant advances in our understanding of the temporal and spatially distributed functional neuroanatomy of language.

Table 1. Convergence of findings on hemispheric contributions to the perception of linguistic prosody

| | <u>Dichotic Listening</u> | | <u>Lesion Deficits</u> | | <u>Brain Imaging</u> | |
|-----------------|---------------------------|----|------------------------|----|----------------------|----|
| | LH | RH | LH | RH | LH | RH |
| Tone | ● | ○ | ● | ○ | ● | ● |
| Phonemic Stress | ● | ● | ● | ● | ● | ● |
| Focus | ● | ● | ● | ● | ● | ● |
| Intonation | ● | ● | ● | ● | ● | ● |

Note. A filled circle (●) indicates that the data source (dichotic listening, lesion deficits, brain imaging) yields evidence in support of a hemispheric contribution (LH, RH) in response to the perception of a given prosodic unit (tone, phonemic stress, sentence focus, sentence modality). An unfilled circle (○) indicates that no contribution from a given hemisphere is evident for that prosodic unit and data source. A blank means that no data are available in the extant literature for that prosodic unit and data source. LH = left hemisphere; RH = right hemisphere.



Figure 1. This fMRI activation map shows cortical areas in the left temporal lobe that are activated in response to discrimination of Mandarin and Thai tones by Chinese and Thai listeners. For the Chinese group, red indicates those non-overlapping areas in which Mandarin tones elicited stronger activation than Thai tones. For the Thai group, blue indicates those non-overlapping areas in which Thai tones elicited stronger activation than Mandarin tones. Across groups, yellow represents an overlapping area of activation in the left planum temporale in which native tones elicited stronger activation as compared to non-native tones.

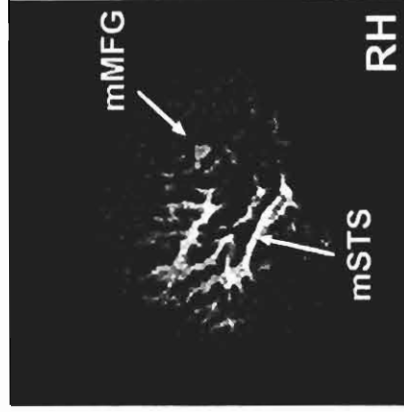


Figure 2. Two regions are found in the frontal and temporal lobes that are lateralized to the RH across tone and intonation tasks in both Chinese and English groups. This fMRI activation map, for example, shows a comparison of intonation in 3-syllable utterances relative to passive listening for the Chinese group. This right-sided frontotemporal network serves to maintain pitch information regardless of its linguistic relevance. mMFG, middle portion of middle frontal gyrus; mSTS, middle portion of superior temporal sulcus. The red-yellow colours indicate those areas activated significantly ($p < 0.005$) more in discrimination of tone and intonation as compared to passive listening.

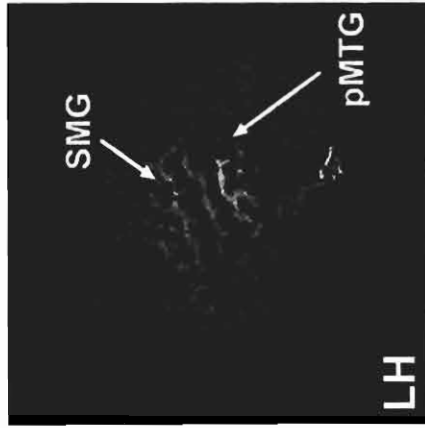


Figure 3. This fMRI activation map shows that the Chinese group exhibited significantly greater activity in the left SMG and pMTG relative to the English group for both tasks. According to Hickok and Poeppel (2004), the SMG serves as part of an auditory-motor interface and supports phonological processing; the pMTG as part of an auditory-meaning interface, and supports conceptual-semantic processing. SMG, supramarginal gyrus; pMTG, posterior middle temporal gyrus. The red-yellow colours indicate those areas activated significantly ($p < 0.005$) more by the Chinese group relative to the English group.

Notes

1. A voxel is a unit of volume corresponding to the smallest element depicted in a three-dimensional brain image.

References

- Abramson, Arthur S.
1962 *The vowels and tones of standard Thai: Acoustical measurements and experiments*. Bloomington: Indiana University Research Center in Anthropology, Folklore, and Linguistics.
- Bates, Elizabeth, Stephen Wilson, Ayşe P. Saygin, Fred Dick, Marty Sereno, Robert T. Knight, and Nina F. Dronkers
2003 Voxel-based lesion-symptom mapping. *Nature Neuroscience* 6: 448–450.
- Baum, Shari R.
1998 The role of fundamental frequency and duration in the perception of linguistic stress by individuals with brain damage. *Journal Speech Language Hearing Research* 41: 31–40.
- Baum, Shari R., and Marc D. Pell
1999 The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology* 13: 581–608.
- Baum, Shari R., and Veena D. Dwivedi
2003 Sensitivity to prosodic structure in left- and right-hemisphere-damaged individuals. *Brain and Language* 87: 278–289.
- Behrens, Susan
1985 The perception of stress and lateralization of prosody. *Brain and Language* 26: 332–348.
- Benson, R., Doug Whalen, Matthew Richardson, Brook Swainson, Vincent Clark, Song Lai, and Alvin M. Liberman
2001 Parametrically dissociating speech and nonspeech perception in the brain using fMRI. *Brain and Language* 78: 364–396.
- Blumstein, Sheila, and William E. Cooper
1974 Hemispheric processing of intonation contours. *Cortex* 10: 146–158.
- Brådvik, Björn, C. Dravins, Stig Holtås, Ingmar Rosén, Erik Rydning, and David Ingvar
1991 Disturbances of speech prosody following right hemisphere infarcts. *Acta Neurologica Scandinavica* 84: 114–126.
- Bryan, Karen
1989 Language prosody and the right hemisphere. *Aphasiology* 3: 285–299.
- Corbetta, Maurizio, J. Michelle Kincade, John Ollinger, Marc P. McAvoy, and Gordon L. Shulman
2000 Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience* 3: 292–297.
- Dogil, Grzegorz, Hermann Ackermann, Wolfgang Grodd, Hubert Haider, Hans Kamp, Jörg Mayer, Axel Riecker, and Dirk Wildgruber
2002 The speaking brain: A tutorial introduction to fMRI experiments in the production of speech, prosody and syntax. *Journal of Neurolinguistics* 15: 59–90.
- Dronkers, Nina F.
1996 A new brain region for coordinating speech articulation. *Nature* 384: 159–161.
- Dronkers, Nina F., Steven Pinker, and Anthony Damasio
2000 Language and the aphasias. In *Principles of neural science*, 4th ed., E. R. Kandel, J. H. Schwartz, and T. M. Jessell (eds.), 1169–1187. New York: McGraw-Hill.
- Dronkers, Nina F., David P. Wilkins, Robert D. Van Valin, Brenda B. Redfern, and Jeri J. Jaeger
2004 Lesion analysis of the brain areas involved in language comprehension. *Cognition* 92: 145–177.
- Emmorey, Karen D.
1987 The neurological substrates for prosodic aspects of speech. *Brain and Language* 30: 305–320.

- Eng, Nancy, Lorraine Obler, Katherine Harris, and Arthur S. Abramson
1996 Tone perception deficits in Chinese-speaking Broca's aphasics. *Aphasiology* 10: 649–656.
- Friederici, Angela D., and Kai Alter
2004 Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language* 89: 267–276.
- Fu, Qian-Jie, Fan-Gang Zeng, Robert V. Shannon, and Sigfrid D. Soli
1998 Importance of tonal envelope cues in Chinese speech recognition. *Journal of the Acoustical Society of America* 104: 505–510.
- Gandour, Jackson T.
1987 Tone production in aphasia. In *Phonetic approaches to speech production in aphasia and related disorders*, J. Ryalls (ed.), 45–57. Boston, MA: College-Hill.
1994 Phonetics of tone. In *The Encyclopedia of Language and Linguistics*, Vol. 6, R. Asher, and J. Simpson (eds.), 3116–3123. New York: Pergamon Press.
1998a Aphasia in tone languages. In *Aphasia in Atypical Populations*, P. Coppens, A. Basso, and Y. Lebrun (eds.), 117–141. Hillsdale, NJ: Lawrence Erlbaum.
1998b Phonetics and phonology. In *Handbook of Neurolinguistics*, B. Stemmer, and H. A. Whitaker (eds.), 207–219. San Diego, CA: Academic Press.
- Gandour, Jackson T., and Rochana Dardarananda
1983 Identification of tonal contrasts in Thai aphasic patients. *Brain and Language* 18: 98–114.
- Gandour, Jackson T., Donald Wong, Li Hsieh, Bret Weinzapfel, Diana Van Lancker, and Gary D. Hutchins
2000 A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience* 12: 207–222.
- Gandour, Jackson T., Donald Wong, Mark Lowe, Mario Dzemidzic, Nakarin Satthammuwong, Yunxia Tong, and Xiaojian Li
2002 A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. *Journal of Cognitive Neuroscience* 14: 1076–1087.
- Gandour, Jackson T., Mario Dzemidzic, Donald Wong, Mark Lowe, Yunxia Tong, Li Hsieh, Nakarin Satthammuwong, and Joseph Lurito
2003a Temporal integration of speech prosody is shaped by language experience: An fMRI study. *Brain and Language* 84: 318–336.
- Gandour, Jackson T., Yi Xu, Donald Wong, Mario Dzemidzic, Mark Lowe, Xiaojian Li, and Yunxia Tong
2003b Neural correlates of segmental and tonal information in speech perception. *Human Brain Mapping* 20: 185–200.
- Gandour, Jackson T., Yunxia Tong, Donald Wong, Thomas Talavage, Mario Dzemidzic, Yisheng Xu, Xiaojian Li, and Mark Lowe
2004 Hemispheric roles in the perception of speech prosody. *Neuroimage* 23: 344–357.
- Griffiths, Timothy D., and Jason D. Warren
2002 The planum temporale as a computational hub. *Trends in Neurosciences* 25: 348–353.
- Heilman, Kenneth, Robert Scholes, and Robert Watson
1984 Comprehension of affective and nonaffective speech. *Neurology* 34: 917–921.
- Hickok, Gregory, and David Poeppel
2004 Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition* 92: 67–99.
- Howie, John M.
1976 *Acoustical studies of Mandarin vowels and tones*. New York: Cambridge University Press.
- Hsieh, Li, Jackson T. Gandour, Donald Wong, and Gary D. Hutchins
2001 Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain and Language* 76: 227–252.
- Hugdahl, Kenneth
1999 Brain lateralization: Dichotic listening studies. In *Encyclopedia of neuroscience*, 2nd ed, G. Adelman, and B. H. Smith (eds.), 276–279. Amsterdam: Elsevier.
- Hughes, C. P., Jin Lieh Chan, and Ming Shung Su
1983 Aprosodia in Chinese patients with right cerebral hemisphere lesions. *Archives of Neurology* 40: 732–736.
- Ivry, Richard, and Lynn Robertson
1998 *The two sides of perception*. Cambridge, MA: MIT Press.
- Jacquemot, Charlotte, Christophe Pallier, Denis LeBihan, Stanislaus Dehaene, and Emmanuel Dupoux
2003 Phonological grammar shapes the auditory cortex: A functional magnetic resonance imaging study. *Journal of Neuroscience* 23: 9541–9546.
- Kimelman, Mikael D. Z.
1991 The role of target word stress in auditory comprehension by aphasic listeners. *Journal of Speech and Hearing Research* 34: 334–339.
- Klein, Denise, Robert Zatorre, Brenda Milner, and Viviane Zhao
2001 A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *Neuroimage* 13: 646–653.
- Klouda, Gail V., Donald A. Robin, Nancy R. Graff-Radford, and William E. Cooper
1988 The role of callosal connections in speech prosody. *Brain and Language* 35: 154–171.
- Koelsch, Stefan, Thomas C. Gunter, D. Yves von Cramon, Stefan Zysset, Gabriele Lohmann, and Angela Friederici
2002 Bach speaks: A cortical “Language-network” serves the processing of music. *Neuroimage* 17: 956–966.
- Krishnan, Ananthanarayan, Yisheng Xu, Jackson T. Gandour, and Peter A. Cariani
2004 Human frequency-following response: Representation of pitch contours in Chinese tones. *Hearing Research* 189: 1–12.

- 2005 Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research* 25: 161–168.
- Lehiste, Ilse
1996 Suprasegmental features of speech. In *Principles of experimental phonetics*, N. J. Lass (ed.), 226–244. St. Louis, MO: Mosby.
- Li, Xiaojian, Jackson T. Gandour, Thomas Talavage, Donald Wong, Mario Dzemidzic, Mark Lowe, and Yunxia Tong
2003 Selective attention to lexical tones recruits left dorsal frontoparietal network. *Neuroreport* 14: 2263–2266.
- Li, Xiaojian, D. Wong, Jackson T. Gandour, Mario Dzemidzic, Yunxia Tong, Thomas Talavage, and Mark Lowe
2004 Neural network for encoding immediate memory in phonological processing. *Neuroreport* 15: 2459–2462.
- Meyer, Martin, Kai Alter, Angela Friederici, Gabriele Lohmann, and D. Yves von Cramon
2002 fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping* 17: 73–88.
- Meyer, Martin, Kai Alter, and Angela Friederici
2003 Functional neuroimaging exposes differential brain responses to syntax and prosody during auditory sentence comprehension. *Journal of Neurolinguistics* 16: 277–300.
- Moen, Inger
1993 Functional lateralization of the perception of Norwegian word tones – evidence from a dichotic listening experiment. *Brain and Language* 44: 400–413.
- Moen, Inger, and Kjetil Sundet
1996 Production and perception of word tones (pitch accents) in patients with left and right hemisphere damage. *Brain and Language* 53: 267–281.
- Naeser, Margaret, and Stephen W. C. Chan
1980 Case study of a Chinese aphasic with the Boston Diagnostic Aphasia Exam. *Neuropsychologia* 18: 389–410.
- Paekard, Jerome L.
1993 *A linguistic investigation of aphasic Chinese speech*. Dordrecht, Holland: Kluwer Academic.
- 2006 The manifestation of aphasic syndromes in Chinese. In *Handbook of East Asian psycholinguistics: Chinese*, Vol. 1., Ping Li, Li Hai Tan, Elizabeth Bates, and Ovid J. L. Tzeng (eds.), 330–345. New York: Cambridge University Press.
- Pell, Marc D., and Shari R. Baum
1997 The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults. *Brain and Language* 57: 80–99.
- Pell, Marc D.
1998 Recognition of prosody following unilateral brain lesion: Influence of functional and structural attributes of prosodic contours. *Neuropsychologia* 36: 701–715.
- Perkins, Judy M., Jane A. Baran, and Jackson T. Gandour
1996 Hemispheric specialization in processing intonation contours. *Aphasiology* 10: 343–362.
- Plante, E., Marlina Creusere, and Cynthia Sabin
2002 Dissociating sentential prosody from sentence processing: Activation interacts with task demands. *Neuroimage* 17: 401–410.
- Poeppel, David
2003 The analysis of speech in different temporal integration windows: Cerebral lateralization as ‘asymmetric sampling in time’. *Speech Communication* 41: 245–255.
- Rugg, Michael D.
1999 Functional neuroimaging in cognitive neuroscience. In *The neurocognition of language*, C. M. Brown (ed.), 15–36. Oxford: Oxford University Press.
- Schirmer, Annett, Siu-Lam Tang, Trevor Penney, Thomas Gunter, and Hsuan-Chih Chen
2005 Brain responses to segmentally and tonally induced semantic violations in Cantonese. *Journal of Cognitive Neuroscience* 17: 1–12.
- Shaywitz, Bennett A., Sally E. Shaywitz, Kenneth R. Pugh, Robert K. Fulbright, Pawel Skudlarski, W. Einar Mencil, R. Todd Constable, Karen E. Marchione, Jack M. Fletcher, Rafael Klorman, Cheryl Lacadie, and John C. Gore
2001 The functional neural architecture of components of attention in language-processing tasks. *Neuroimage* 13: 601–612.
- Shen, Xiao-Nan
1990 *The prosody of Mandarin Chinese*. Berkeley, CA: University of California Press.
- Shipley-Brown, Frances, William O. Dingwall, Charles I. Berlin, Grace Yeni-Komshian, and Sandra Gordon-Salant
1988 Hemispheric processing of affective and linguistic intonation contours in normal subjects. *Brain and Language* 33: 16–26.
- Tallal, Paula, Steve Miller, and R. Holly Fitch
1993 Neurobiology of speech perception. *Annals of New York Academy of Science* 682: 27–47.
- Tingsabhadh, M. R. Kalaya, and Arthur S. Abramson
1993 Thai. *Journal of the International Phonetic Association* 23: 25–28.
- Tong, Yunxia, Jackson T. Gandour, Thomas Talavage, Donald Wong, Mario Dzemidzic, Yisheng Xu, Xiaojian Li, and Mark Lowe
2005 Neural circuitry underlying sentence-level linguistic prosody. *Neuroimage* 28: 417–428.
- Van Lancker, Diana, and Victoria Fromkin
1973 Hemispheric specialization for pitch and tone: Evidence from Thai. *Journal of Phonetics* 1: 101–109.

- 1978 Cerebral dominance for pitch contrasts in tone language speakers and in musically untrained and trained English speakers. *Journal of Phonetics* 6: 19–23.
- Walker, Judy P., Kirk Fongemie, and Tracy Daigle
2001 Prosodic facilitation in the resolution of syntactic ambiguities in subjects with left and right hemisphere damage. *Brain and Language* 78: 169–196.
- Walker, Judy P., Tracy Daigle, and Melissa Buzzard
2002 Hemispheric specialisation in processing prosodic structures: Revisited. *Aphasiology* 16: 1155–1172.
- Wang, Yue, Allard Jongman, and Joan A. Sereno
2001 Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language* 78: 332–348.
- Wang, Yue, Joan A. Sereno, Allard Jongman, and Joy Hirsch
2003 fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience* 15: 1019–1027.
- Wang, Yue, Dawn Behne, Allard Jongman, and Joan A. Sereno
2004 The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics* 25: 449–466.
- Weintraub, Sharon, M. Marsel Mesulam, and L. Kramer
1981 Disturbances in prosody. A right-hemisphere contribution to language. *Archives of Neurology* 38: 742–744.
- Whalen, Douglas H., and Yi Xu
1992 Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica* 49: 25–47.
- Wildgruber, Dirk, Ingo Hertrich, Axel Riecker, Michael Erb, S. Anders, Wolfgang Grodd, and Hermann Ackermann
2004 Distinct frontal regions subserved evaluation of linguistic and emotional aspects of speech intonation. *Cerebral Cortex* 14: 1384–1389.
- Wong, Patrick C.
2002 Hemispheric specialization of linguistic pitch patterns. *Brain Research Bulletin* 59: 83–95.
- Wong, Patrick C., Lawrence M. Parsons, Michael Martinez, and Randy L. Diehl
2004 The role of the insular cortex in pitch pattern perception: The effect of linguistic contexts. *Journal of Neuroscience* 24: 9153–9160.
- Xu, Yi
1997 Contextual tonal variations in Mandarin. *Journal of Phonetics* 25: 61–83.
1999 Effects of tone and focus on the formation and alignment of f_0 contours. *Journal of Phonetics* 27: 55–105.
2006 Tone in connected discourse. In *Encyclopedia of language and linguistics*, 2nd ed., Vol. 12., K. Brown (ed.). Oxford: Elsevier.
- Xu, Yisheng, Jackson T. Gandour, Thomas Talavage, Donald Wong, Mario Dzemidzic, Yunxia Tong, Xiaojian Li, and Mark Lowe
2006 Activation of the left planum temporale in pitch processing is shaped by language experience. *Human Brain Mapping*: 27 (2): 173–183.
- Yiu, Edward, and Angela Fok
1995 Lexical tone disruption in Cantonese aphasic speakers. *Clinical Linguistics and Phonetics* 9: 79–92.
- Yip, Moira
2003 *Tone*. New York: Cambridge University Press.
- Zatorre, Robert J., Alan C. Evans, and Ernst Meyer
1994 Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience* 14: 1908–1919.
- Zatorre, Robert J., Todd A. Mondor, and Alan C. Evans
1999 Auditory attention to space and frequency activates similar cerebral systems. *Neuroimage* 10: 544–554.

