

Functional Heterogeneity of Inferior Frontal Gyrus Is Shaped by Linguistic Experience

Li Hsieh,* Jack Gandour,* Donald Wong,† and Gary D. Hutchins‡

**Department of Audiology and Speech Sciences, Purdue University; †Department of Anatomy and Cell Biology, Indiana University School of Medicine; and ‡Department of Radiology, Indiana University School of Medicine*

Published online February 15, 2001

A crosslinguistic, positron emission tomography (PET) study was conducted to determine the influence of linguistic experience on the perception of segmental (consonants and vowels) and suprasegmental (tones) information. Chinese and English subjects (10 per group) were presented binaurally with lists consisting of five Chinese monosyllabic morphemes (speech) or low-pass-filtered versions of the same stimuli (nonspeech). The first and last items were targeted for comparison; the time interval between target tones was filled with irrelevant distractor tones. A speeded-response, selective attention paradigm required subjects to make discrimination judgments of the target items while ignoring intervening distractor tones. PET scans were acquired for five tasks presented twice: one passive listening to pitch (nonspeech) and four active (speech = consonant, vowel, and tone; nonspeech = pitch). Significant regional changes in blood flow were identified from comparisons of group-averaged images of active tasks relative to passive listening. Chinese subjects show increased activity in left premotor cortex, pars opercularis, and pars triangularis across the four tasks. English subjects, on the other hand, show increased activity in left inferior frontal gyrus regions only in the vowel task and in right inferior frontal gyrus regions in the pitch task. Findings suggest that functional circuits engaged in speech perception depend on linguistic experience. All linguistic information signaled by prosodic cues engages left-hemisphere mechanisms. Storage and executive processes of working memory that are implicated in phonological processing are mediated in discrete regions of the left frontal lobe. © 2001 Academic Press

Key Words: PET; pitch; language; speech perception; phonology; prosody; verbal working memory; Chinese; prelexical phonological processing; Broca's area.

A fundamental, yet controversial issue is whether the human brain contains neural circuits that are uniquely engaged in speech perception. While it seems indisputable that language is subserved by left-hemisphere (LH) mechanisms, questions remain about whether the LH is lateralized for speech, language, or something else. Hypothe-

Funding was provided by research grants awarded to J.G. from the James S. McDonnell Foundation, Purdue Research Foundation, Showalter Trust, and Center for Social and Behavioral Sciences of the School of Liberal Arts. We thank Rich Fain and his technical staff of the Physiologic Imaging Center for their invaluable assistance. We also thank Nakarin Sattamnuwong for his programming support. This article was based in part on a doctoral dissertation completed by the first author at Purdue University in August 2000. Portions of this work were presented at the 7th annual meeting of the Cognitive Neuroscience Society, San Francisco, April 2000 and the 6th International Conference on Functional Mapping of the Human Brain, San Antonio, June 2000.

Address correspondence and reprint requests to Jack Gandour, Department of Audiology and Speech Sciences, Heavilon Hall, Purdue University, West Lafayette, IN 47907-1353. Fax: (765) 494-0771. E-mail: gandour@purdue.edu.

ses proposed to account for functional hemispheric asymmetries can generally be classified as either *cue dependent*, i.e., basic neural mechanisms underlie processing of complex auditory stimuli regardless of linguistic relevance (e.g., Ivry & Leiby, 1993), or *task dependent*, i.e., specialized neural mechanisms exist that are activated only by speech (e.g., Imaizumi, Mori, Kiritani, Hosoi, & Tonoike, 1998).

Speech prosody is composed of multiple, phonetic correlates (e.g., pitch, duration, and loudness) that play a critical role in linguistic and paralinguistic functions. One key linguistic function is phonological, whereby variations in pitch patterns minimally signal differences in meaning at the lexical level. Such pitch variations characterize *tone* languages (e.g., Chinese: *ma*^{high level}, ‘‘mother’’; *ma*^{high rising}, ‘‘hemp’’, *ma*^{low falling rising}, ‘‘horse’’; *ma*^{high falling}, ‘‘scold’’). Previous lesion deficit and dichotic listening data (see Gandour, 1998a, 1998b; Baum & Pell, 1999, for reviews), as well as recent positron emission tomography (PET) data (Gandour, Wong, & Hutchins, 1998; Gandour et al., 2000), have consistently supported LH mechanisms specialized for the perception of lexical tones. However, a preponderant role for right-hemisphere (RH) mechanisms has been demonstrated for pitch perception of complex auditory stimuli in nonlinguistic contexts (Johnsrude, Penhume, & Zatorre, 2000; Zatorre, Evans, Meyer, & Gjedde, 1992; Zatorre, Evans, & Meyer, 1994; Zatorre, Meyer, Gjedde, & Evans, 1996, and references therein).

We propose a task-dependent hypothesis in which phonological units of language, whether *suprasegmental* (e.g., tone) or *segmental* (consonants and vowels), are processed in discrete left frontal regions, including the perisylvian region near Broca’s area. LH mechanisms are selectively employed for processing linguistic information irrespective of acoustic cues or subtype of phonological unit, whereas RH mechanisms are employed for prosody-specific cues (Imaizumi et al., 1998). Our PET imaging approach is *crosslinguistic*, which enables us to distinguish brain activation patterns attributable directly to linguistic experience from those reflecting language-independent complex-sound processing. We use *low-pass-filtered speech* as control stimuli, which enables us to distinguish suprasegmental patterns in a language context from those in a nonlanguage context.

As part of our ongoing investigation of speech prosody in tone languages, the objective of this study is to map the functional circuits underlying the processing of segmental and suprasegmental phonological units in Chinese. In an earlier study (Gandour et al., 1998, 2000), Chinese and English listeners did not show the same LH lateralization as Thai listeners when making perceptual judgments of Thai tones. In this study, Chinese and English listeners are asked to make perceptual judgments of Chinese tones, consonants, and vowels. It is hypothesized that Chinese listeners will show LH lateralization for both suprasegmental and segmental phonological units.

METHODS

Subjects

Ten adult, native speakers of Chinese (6 male and 4 female) and 10 adult, native speakers of American English (5 male and 5 female) were closely matched in age [Chinese: $M = 24.9$, $SD = 2.3$; English: $M = 25.6$, $SD = 2.8$; $t(18) = -0.603$, $p < .55$] and years of formal education [Chinese: $M = 18.3$, $SD = 1.6$; English: $M = 18.6$, $SD = 2.8$; $t(18) = -0.290$, $p < .78$]. All Chinese subjects came from the People’s Republic of China and were native speakers of Mandarin Chinese. All subjects were strongly right-handed as measured by the Edinburgh Handedness Inventory (Oldfield, 1971) with no family history of left-handedness (i.e., no history of childhood left-handedness and no left-handed parents, siblings, or children); were musically untrained (i.e., no formal musical training within the past 7 years and 4 years or less total musical training); exhibited normal hearing sensitivity (i.e., pure-tone air conduction thresholds of 20 dB HL or better in both ears) at frequencies of 0.5, 1, 2, and 4 kHz; and exhibited no

significant depressive symptoms as measured by the Beck Depression Inventory (Beck, 1978). None of the subjects had a history of neurological illness or loss of consciousness greater than 30 min. All were free of current use of CNS medications as well as any history of recreational drug use. Subjects gave written informed consent. The study was approved by the Institutional Review Board of Indiana University Purdue University Indianapolis and Clarian.

Stimuli

Two types of stimuli were prepared: (1) (speech condition) natural speech Chinese monosyllabic morphemes and (2) (nonspeech condition) low-pass speech filtered (200 Hz with 50 dB per octave attenuation rate) versions of the same monosyllables. The filtered stimuli were constructed to eliminate higher spectral information while at the same time preserving the acoustic characteristics of the Chinese monosyllables in terms of duration, amplitude, and F_0 contour. As judged by three Chinese listeners, none of the filtered stimuli were recognizable as Chinese words or parts thereof.

Forty-five different lists were used, each consisting of five monosyllabic morphemes. Bisyllabic words predominate in the Chinese lexicon (Chao, 1968). To minimize semantic interference, no two adjacent morphemes in any of the lists matched an existing lexical item. The use of monosyllables also eliminated word-length effects. All monosyllables in the consonant, vowel, and tone conditions were identical, thus eliminating any effects due to differences in frequency of occurrence of consonants, vowels, and tones (Suen, 1986). All monosyllables exhibited a uniform structure, i.e., a single consonant followed by a vowel. Chinese consonants, vowels, and tones occurred with equal frequency in the target syllables in first and last position as well as in the intermediate syllables. In the target syllables, there were 90 occurrences each of consonants, vowels, and tones; in the intermediate syllables, 135 occurrences each. In the syllables in first and last position, 19 of 21 Chinese consonant phonemes were represented, 6 of 6 vowel phonemes, and all 4 tones (Cheng, 1973).

A list length of five items was chosen to investigate short-term retention of consonants, vowels, tones, and pitch patterns under relatively high memory load (cf. Zatorre et al., 1994). Lists were designed so that minimal pairs for Chinese consonants, vowels, or tones occurred in the first and last position. One-third of the lists contained first/last minimal pairs for consonants (e.g., ni^4/li^4); one-third, minimal pairs for vowels (e.g., ni^1/na^1); and one-third, minimal pairs for tone (e.g., na^4/na^2). For consonants, minimal pairs were designed to contrast place of articulation (e.g., gau^1/bau^1), manner of articulation (e.g., lau^3/nau^3), or voicing (e.g., ba^2/pa^2); for vowels, height (e.g., ma^4/mu^4), backness (e.g., bi^3/bu^3), or rounding (e.g., $ji^1/jü^1$); and for tone, height, and direction (e.g., ma^4/ma^3). Within a list, intermediate syllables differed from the first and last in at least two features (consonant, vowel, and tone) to minimize phonological similarity effects (Baddeley, 1966). To eliminate potentially confounding effects of tone sandhi on perceptual judgments (Xu, 1991), Tones 2 and 3, respectively, did not occur on the first and second monosyllables of a list. No list contained any occurrences of two consecutive monosyllables that carried Tone 3.

A randomly ordered list of 285 monosyllabic morphemes was prepared for recording the stimuli. All monosyllables were printed in large Chinese characters and *pinyin* orthography. A 41-year-old male speaker of standard Chinese from the People's Republic of China was instructed to read the words in isolation. He was requested to reread words if they were misread or if they were inconsistent with his normal loudness level or speaking rate. All recording was completed in a soundproof booth in a single 2-h session. His utterances were recorded using an AKG C410 omnidirectional headset microphone and a Sony TCD-D7 DAT cassette recorder at a sampling rate of 48 kHz. The subject was seated and the microphone was maintained at a position 12 cm from the lips.

All of the natural speech and low-pass-filtered speech stimuli were digitally edited to have equal peak energy level in decibels SPL with the remaining data within each of the stimuli scaled accordingly. The sound pressure levels of cognate pairs of natural speech and low-pass-filtered stimuli were then measured at the output of the earphones (E-A-RTONE 3A, 50 Ω) in dBA using a Brüel and Kjaer 2230 sound-level meter. Because of differences in frequency range and bandwidths, the sound pressure level of the natural speech stimuli, on average ($M = 63.4$ dBA, $SD = 1.7$), was 6 dBA higher than that of the low-pass filtered ($M = 57.9$ dBA, $SD = 2.7$). To ensure that all stimuli were presented at about the same loudness level, the sound pressure levels of both natural speech and low-pass-filtered stimuli were digitally attenuated to the minimum dBA level in the low-pass-filtered speech stimuli.

Total duration of the lists, including an intersyllable interval of 200 ms, ranged from 2.30 to 3.06 s ($M = 2.62$ s; $SD = 0.17$ s). The response interval was 1.5 s, resulting in a delivery rate of about one list every 4 s. Within lists, the syllable presentation rate averaged about one syllable every 0.52 s. Total stimulus duration during the 90-s PET data-acquisition interval was 40 s. These rate and duration characteristics exceeded thresholds for generating a sufficient response magnitude in human auditory cortex (Binder et al., 1994; Price et al., 1992) as well as those employed in a similar study with musical stimuli (Zatorre et al., 1994).

TABLE 1
PET Paradigm

Condition	Stimulus	Response	Example	
Passive Listening	Filtered monosyllables ^a	Alternately	_4 _3 _2 _1 _4	(click L)
		Click button	_3 _4 _3 _2 _3	(click R)
			_2 _4 _4 _3 _2	(click L)
			_1 _4 _1 _3 _4	(click R)
Consonant	Chinese monosyllables ^b	Click button	mu ⁴ ti ³ she ² liu ¹ ma ⁴	(Yes; click L)
			ni ³ ke ⁴ bao ³ fo ² di ³	(No; click R)
			ba ² nao ⁴ cou ⁴ gu ³ pa ²	(No; click R)
			gou ¹ bi ⁴ xiu ¹ shou ³ gou ⁴	(Yes; click L)
Vowel	Chinese monosyllables ^b	Click button	mu ⁴ ti ³ she ² liu ¹ ma ⁴	(No; click R)
			ni ³ ke ⁴ bao ³ fo ² di ³	(Yes; click L)
			ba ² nao ⁴ cou ⁴ gu ³ pa ²	(Yes; click L)
			gou ¹ bi ⁴ xiu ¹ shou ³ gou ⁴	(Yes; click L)
Tone	Chinese monosyllables ^b	Click button	mu ⁴ ti ³ she ² liu ¹ ma ⁴	(Yes; click L)
			ni ³ ke ⁴ bao ³ fo ² di ³	(Yes; click L)
			ba ² nao ⁴ cou ⁴ gu ³ pa ²	(Yes; click L)
			gou ¹ bi ⁴ xiu ¹ shou ³ gou ⁴	(No; click R)
Pitch	Filtered monosyllables ^a	Click button	_4 _3 _2 _1 _4	(Yes; click L)
			_3 _4 _3 _2 _3	(Yes; click L)
			_2 _4 _4 _3 _2	(Yes; click L)
			_1 _4 _1 _3 _4	(No; click R)

Note. Chinese monosyllables are presented in *pinyin* orthography (superscript numbers are defined as follows: ¹ = tone 1, ² = tone 2, ³ = tone 3, ⁴ = tone 4). Correct responses are shown in parentheses (L = left mouse button; R = right mouse button).

^a Nonspeech condition.

^b Speech condition.

Tasks

Three speech conditions were used that involved an active task: *consonant*, *vowel*, and *tone*; one nonspeech condition was used that involved both an active task (*pitch*) and passive listening (Table 1). In the four active conditions, subjects were required to make discrimination judgments of consonants, vowels, tones, and pitch patterns that occur in the *first* and *last* syllables of each list, ignoring the intervening syllables, by clicking a mouse button (left button = same; right button = different). In the passive listening condition, subjects were required to listen to the filtered speech stimuli, alternately clicking the left and right mouse button after each list. Speech stimuli for the consonant, vowel, and tone conditions were identical. Responses to these stimuli are expected to vary according to the instructions given prior to the different conditions. The term *active* means that an overt discrimination judgment is required when listening to these lists of auditory stimuli. The term *passive* means that no overt judgment is required but does not imply that listening to such lists is a passive cognitive process. Subjects were free to make their same–different judgment at any time during the 1.5-s response interval. Tasks were matched for average SPL (75 dBA), trial rate (one trial per 4 s), trial duration (2.62 s), and same/different response rate (two-thirds same, one-third different).

The consonant, vowel, and tone tasks, which require decisions about Chinese phonological units in a nonsyntactic context, were designed to elicit language processing at a phonetic level with the goal of identifying brain areas involved in perceptual processing of segmental (consonants and vowels) and suprasegmental (tones) information. The pitch task, which requires decisions about pitch patterns, was designed to elicit nonlexical processing of pitch information with the goal of identifying brain areas involved in perceptual processing of pitch. Moreover, the pitch task was designed to control for activation of early auditory processing, nonspecific executive functions mediating sustained attention and arousal, working memory, and motor response formation (cf. Binder et al., 1994, 1997). The passive listening condition was designed to address cognitive processes inherent to automatic, nondirected perceptual analysis that might be overlooked if only experimenter-imposed judgments were to be imposed. No first/last minimal pairs were judged more than once in any task, thus eliminating confounding effects of prior exposure.

The five conditions were presented twice (*intrasubject* averaging) during the scanning session. The passive listening condition was presented first and last in scan order for all subjects. Order of presentation of the active conditions was randomized for each subject to eliminate any effect of order of presentation on behavioral or PET data. Within conditions, two different random orders were used for the two scanning periods. In the first scanning period, the first/last minimal pair was presented in the order $a \rightarrow b$, in the second scanning period $b \rightarrow a$. Stimuli were presented in blocked format to ensure homogeneity of cognitive and neuronal activation.

No more than 10 days prior to the scanning session, Chinese subjects were trained to a criterion of 85% correct on a 15-trial practice list for consonant, vowel, tone, and tasks. English subjects were trained to a criterion of 85% correct on the consonant and vowel tasks and 70% on the tone tasks. Both Chinese and English subjects were trained to a criterion of 80% correct on the pitch task.

All stimuli were played out from a PC computer using SuperLab software (Cedrus Corporation) through an NAD 3020A amplifier and presented binaurally through E-A-RTONE 3A foam insert earphones at a comfortable listening level of approximately 75 dBA.

PET/MRI Image Acquisition and Processing

PET scans were obtained using a Siemens 951/31 R, which measures 31 image planes simultaneously over an axial field-of-view of 10.8 cm. The intrinsic image resolution of this system is approximately 6.0 mm FWHM in plane and 5.0 mm FWHM in the axial direction.

During the scanning session, the subject lay supine on the imaging table with his/her eyes blindfolded. Head movement was restricted by placing the subject's head on a custom-fit, firm pillow and by strapping his/her forehead to the imaging table, allowing pixel-by-pixel within-subject comparisons of regional CBF across stimulus conditions. A peripheral venipuncture and an intravenous infusion line were placed in the subject's left arm. The subject's head was positioned in the scanner's gantry to acquire tomographic images beginning inferiorly at 1 cm above and parallel to the canthomeatal line. A transmission scan using three rotating $^{68}\text{Ga}/^{68}\text{Ge}$ rod sources was performed prior to H_2^{15}O intravenous injection to correct the emission scans for photon attenuation due to the skull and cerebrum.

PET scanning was carried out for five separate conditions. Each condition was repeated twice (*intra-subject* averaging) for a total of 10 scanning periods. Fifteen seconds after beginning each task, 40 mCi of H_2^{15}O was injected intravenously as a bolus, and tomographic image acquisition began concurrent with bolus injection and continued for 180 s. A rapid sequence of scans was performed throughout the 180-s scanning period to enable the selection of a 90-s time window beginning 35–40 s after the bolus arrived in the brain. For each condition, instructions were given immediately prior to scanning, and several practice trials were given prior to bolus injection. The passive listening condition was presented in the 1st and 10th scanning periods for all subjects. For the 2nd through 9th scanning periods, the order of the active tasks (consonant, vowel, tone, and pitch) was counterbalanced across subjects. Within subjects for the same active task, two different random orders of trials were used for the two scanning periods. Time between scanning periods was about 12 min to allow for radioactive decay to less than 2% of administered levels and the generation of subsequent doses of H_2^{15}O .

The PET images for each subject were transformed to a stereotaxic coordinate system (Minoshima, Berger, Lee, & Mintun, 1992; Minoshima et al., 1993) and linearly standardized to an atlas brain (Talairach & Tournoux, 1988). After pixel values were normalized for global flow rate differences among scans, the data were averaged across subjects within groups for each condition, giving mean and variance values for each condition. Difference images were then constructed by subtracting an average image for one condition from that for another. The difference image was then analyzed for statistical significance on a pixel-by-pixel basis using the t statistic with a pooled variance estimate (Minoshima et al., 1992, 1993; Minoshima, Koeppe, Frey, & Kuhl, 1994), followed by a multiple comparison adjustment based on the Bonferroni method (Friston et al., 1990; Friston, Frith, Liddle, & Frackowiak, 1991). A one-tailed adjusted value of $p < .05$ was used as a criterion for statistical significance.

Paired-image subtractions were analyzed for significant activation foci: (1) tone – passive listening, (2) pitch – passive listening, (3) consonant – passive listening, (4) vowel – passive listening, (5) tone – pitch, (6) consonant – pitch, and (7) vowel – pitch. The first four comparisons (1–4) employed passive listening to nonspeech as the control task so that any activation seen represents neural responses beyond those already present during passive listening to complex auditory stimuli that are homologous to the speech stimuli in terms of suprasegmental properties. The last three comparisons (5–7) used pitch as the control task so that any activation seen represents neural responses beyond those already present during overt pitch comparisons in nonspeech stimuli that are suprasegmentally homologous to the speech stimuli.

MRI images for anatomical correlation with the functional PET data were acquired using a 1.5T GE Signa Advantage MRI system with echoplanar imaging capabilities. A 3D SPGR pulse sequence was

utilized producing 124 1-mm image planes. The patient lay supine and motionless on the imaging table while the SPGR imaging sequence was performed over a period of approximately 30 min. To spatially align the PET and MRI data, MRI images were segmented to identify gray matter, white matter, and CSF. Pixel intensity values were set in each of the segmented regions to expected PET radionuclide levels for each tissue type and smoothed to the same resolution as the PET data. The image volumes were then registered as rigid bodies (three translation parameters, three rotation parameters, and one scale parameter) using Newton's method based on a least-squares cost function. The registration parameters derived from this algorithm were then applied to the original MRI image data in order to map the PET results onto the anatomical images.

Task Performance

Behavioral measures of task performance (response time and accuracy) were collected online using SuperLab software (Cedrus Corp.) to determine whether subjects were performing the task in the expected manner. Separate three-way (task \times scan \times group) analyses of variance (ANOVA) with task and scan as within-subjects variables and group as the between-subjects variable were performed on arcsin-transformed proportions of correct responses (Winer, Brown, & Michels, 1991) and on response times.

Regions of Interest (ROIs)

ROIs were anatomically defined using sulcal landmarks and complementary anatomic conventions (Rademacher, Galaburda, Kennedy, Filipek, & Caviness, 1992; Ebeling, Steinmetz, Huang, & Kahn, 1989; Ono, Kubik, & Abernathy, 1990). This method does not require complex operations of normalization to a fixed brain atlas and takes into account interindividual anatomical differences as well as hemispheric anatomical asymmetries (Steinmetz & Seitz, 1991). It also makes the ROI boundaries independent of the PET images and provides regional activation values for each individual. Individual ROIs were copied onto coregistered PET images to obtain estimates of normalized cerebral blood flow (Collins, Holmes, Peters, & Evans, 1995). Within each anatomical ROI, normalized cerebral blood flow was estimated as the ratio of the radioactivity concentration in the region to that of the whole brain as measured in the PET images. Five ROIs (pars opercularis, BA 44; pars triangularis, BA 45; supplementary motor cortex, medial BA 6; precentral gyrus, BA 6/4; and lateral cerebellum) were drawn and then applied to the Talairach-transformed, PET/MRI coregistered images to obtain estimates of percentage change in normalized rCBF ($\% \text{Task}_{\text{region } i} = \left| \left\{ \left[\left(\text{Task}_{\text{region } i} / \text{Task}_{\text{global } i} \right) \times 100 \right] / \left[\left(L_{\text{scan } 1} / L_{\text{global } 1} + L_{\text{scan } 2} / L_{\text{global } 2} \right) / 2 \right] \right\} - 100 \right|$, where L = passive listening task and Task = tone, pitch, consonant, and vowel).

A four-way (task \times hemisphere \times ROI \times group) ANOVA with hemisphere, ROI, and task as within-subjects variables and group as the between-subjects variable was performed on percentage changes in normalized blood flow. Such a multifactorial design permits us to make direct comparisons of brain activity in selected ROIs *between* groups.

RESULTS

Task Performance

Results of the ANOVA performed on proportions of correct responses revealed a two-way interaction between task and group [$F(3, 54) = 5.61, p < .002$]. Post hoc Student–Newman–Keuls (SNK) multiple comparisons ($\alpha = .01$) indicated no significant differences in response accuracy for Chinese subjects across tasks, whereas English subjects made more errors on the tone and pitch tasks than on the consonant and vowel tasks (Table 2). English subjects also exhibited a significantly lower level of proficiency than Chinese subjects on the tone and pitch tasks.

The ANOVA performed on response times similarly showed a two-way interaction between task and group [$F(3, 54) = 11.69, p < .0001$]. SNK comparisons indicated no significant differences in response times for Chinese subjects, whereas English subjects were slower on the tone task than on the others (Table 2).

TABLE 2
Behavioral Performance

	Response accuracy ^a				Response time ^b			
	C	V	T	P	C	V	T	P
Chinese	94.8 (5.2)	96.0 (3.9)	95.1 (4.5)	88.6 (8.0)	418 (89)	453 (109)	458 (102)	377 (100)
English	92.1 (4.1)	94.1 (4.3)	66.5 (10.3)	76.2 (7.5)	345 (97)	366 (88)	526 (149)	404 (140)

Note. C = consonant; V = vowel; T = tone; and P = pitch.

^a Mean and standard deviation values (in parentheses) are expressed in percentages.

^b Mean and standard deviation values (in parentheses) are expressed in milliseconds.

PET Scanning: CBF Increases

Comparison of the consonant, vowel, tone, and pitch tasks to passive listening for the Chinese group (Fig. 1) yielded significant CBF increases in the premotor cortex across all tasks (Table 3, focus 1; Table 4, focus 1; Table 5, focus 5; Table 6, focus 1) and in the left inferior frontal gyrus (IFG) (pars opercularis) across the consonant, vowel, and tone tasks only (Table 3, focus 3; Table 4, foci 2–3; Table 5, focus 3). Significant CBF increases were also observed in the left IFG (pars triangularis) on the consonant and vowel tasks (Table 4, focus 4; Table 5, focus 4) and in the left frontal operculum and anterior insula on the vowel task (Table 5, focus 8). It is noteworthy that CBF increases were observed for the Chinese group in the left premotor cortex on all four active tasks; in the pars opercularis on the consonant, vowel, and tone tasks; and in the pars opercularis *and* pars triangularis on the consonant and vowel tasks only. However, there was not a complete absence of activity in the pars triangularis on the tone ($-37, 32, 9; t = 4.2$) and pitch ($-42, 39, 4; t = 4.0$) tasks. Although t values are nonsignificant, the stereotaxic coordinates are quite close to those found in the consonant and vowel comparisons to passive listening (Table 4, focus 4; Table 5, focus 4; Fig. 1, cf. Figs. 1A–1B to Figs. 1C–1D), thus making them unlikely to be spurious observations.

Of these three left frontal regions (premotor cortex, pars opercularis, and pars triangularis), the English group demonstrated significant activation singularly in the pars opercularis on the vowel task only (Table 5, focus 10; Fig. 1D). No significant activity was detected in the anterior pars triangularis, even using the lower t threshold values.

Comparison of the pitch task to passive listening for the English group yielded significant CBF increases predominantly in the RH. These regions included the right premotor cortex, IFG, frontal operculum/anterior insula, and frontal pole (Table 6, foci 7, 9, and 13–14; Fig. 1B). Somewhat unexpectedly, no significant CBF increases in right frontal regions were observed for the Chinese group on the pitch task. Indeed, a CBF increase was observed in the left premotor cortex (Table 6, focus 1). Although t values are nonsignificant, activity was observed in the left pars opercularis ($-48, 14, 27; t = 4.2$; cf. Table 3, focus 3; Fig. 1, cf. Fig. 1B to Fig. 1A), and in the left pars triangularis ($-42, 39, 4; t = 4.0$; cf. Table 4, focus 4; Table 5, focus 4; Fig. 1, cf. Fig. 1B to Figs. 1C–1D). It is noted that the stereotaxic coordinates of these two IFG foci are in close proximity to those that exceeded the t threshold set a priori.

A significant CBF increase in the anterior cingulate gyrus and supplementary motor area was also noted for the Chinese group when comparing the tone task to passive listening (Table 3, focus 5). And finally, both groups demonstrated activation in the left fusiform gyrus, the English group in the consonant task (Table 4, focus 6), and the Chinese group in the vowel task (Table 5, focus 1).

TABLE 3
Tone minus Passive Listening

Region	Brodmann's area	Coordinates (mm)			<i>t</i> value
		<i>x</i>	<i>y</i>	<i>z</i>	
Blood flow increases					
Chinese					
1. L premotor cortex	6	-30	-4	45	6.1
2. M vermal cerebellum	—	1	-64	-25	5.3
3. L inferior frontal gyrus (pars opercularis)	44	-48	10	25	5.0
4. M vermal cerebellum	—	6	-49	-14	5.0
5. M anterior cingulate gyrus/SMA	32/6	-6	8	45	4.4
English					
6. R lateral cerebellum	—	26	-62	-36	5.7
7. M vermal cerebellum	—	6	-49	-25	5.0
8. M vermal cerebellum	—	3	-71	-27	4.9
Blood flow decreases					
Chinese					
9. M precuneus	7	1	-60	32	-5.0
10. M medial frontal gyrus	11/10	1	44	-11	-4.6
11. M anterior cingulate gyrus	24	-3	32	-4	-4.5
12. R superior frontal gyrus	8	21	35	43	-4.5
13. R inferior parietal lobule	39	42	-64	27	-4.4
14. R posterior cingulate gyrus	30	19	-51	14	-4.4
15. L superior frontal gyrus	8	-21	30	45	-4.3
16. M medial frontal gyrus/anterior cingulate gyrus	10/32	-6	50	9	-4.3
English					
17. M posterior cingulate gyrus	23	-1	-53	20	-6.0
18. M posterior cingulate gyrus	31	3	-46	38	-4.9
19. L inferior parietal lobule	39/19	-44	-64	22	-4.6
20. L uncus	28	-21	3	-27	-4.5
21. R temporal pole	38	35	12	-27	-4.4

Note. Significant activation foci (blood flow increases and decreases in order of decreasing significance) that exceed the Hammersmith statistical criterion of significance (adjusted *p* threshold = .05) in normalized CBF for subtraction of tone condition minus passive listening condition. Stereotaxic coordinates, in millimeters, are derived from the human brain atlas of Talairach and Tournoux (1988). The *x* coordinate refers to medial-lateral position relative to midline (negative = left); the *y* coordinate refers to anterior-posterior position relative to the anterior commissure (positive = anterior); and the *z* coordinate refers to superior-inferior position relative to the CA-CP (anterior commissure-posterior commissure) line (positive = superior). Designation of Brodmann's areas are referenced to this atlas and are approximate only. L = left; R = right; M = at or near midline; SMA = supplementary motor area.

Across all four active tasks relative to passive listening, significant CBF increases were identified in vermal, paravermal, and right lateral cerebellar regions for Chinese and English groups alike (Table 3, foci 2, 4, 6, and 7–8; Table 4, foci 5 and 7; Table 5, foci 2, 7, 9, and 13; Table 6, foci 2–3, 5–6, 8, and 12).

Comparison of the tone, consonant, and vowel tasks to pitch revealed significant CBF increases consistently in the medial frontal gyrus for the English group (Table 7, foci 1, 3, and 6). A CBF increase was observed for the Chinese group in the left inferior temporal gyrus when the pitch discrimination task was subtracted from the vowel task (Table 7, focus 5). It is noteworthy that the Chinese group failed to demonstrate any significant CBF increases in the left inferior frontal regions when comparing the tone, consonant, and vowel tasks to pitch.

TABLE 4
Consonant minus Passive Listening

Region	Brodmann's area	Coordinates (mm)			t value
		x	y	z	
Blood flow increases					
Chinese					
1. L premotor cortex	6/4	-39	-4	43	5.5
2. L inferior frontal gyrus (pars opercularis)	44/6	-53	14	18	5.1
3. L inferior frontal gyrus (pars opercularis)	44/6	-51	10	25	4.9
4. L inferior frontal gyrus (pars triangularis)	45/46	-42	37	4	4.4
5. R lateral cerebellum	—	33	-64	-25	4.4
English					
6. L fusiform gyrus	37	-48	-51	-18	5.0
7. M vermal cerebellum	—	-1	-71	-29	4.4
8. Pons	—	-1	-35	-22	4.4
Blood Flow Decreases					
Chinese					
9. M posterior cingulate gyrus	31	-1	-55	34	-5.2
10. R posterior cingulate gyrus	31/23	8	-55	22	-4.9
English					
11. M posterior cingulate gyrus	31/23	-1	-51	25	-4.8
12. R posterior cingulate gyrus	31	12	-44	36	-4.4

Note. Significant activation foci (blood flow increases and decreases) for subtraction of consonant condition minus passive listening condition. See also note under Table 3 for further details.

PET Scanning: CBF Decreases

The posterior cingulate and/or precuneus showed significant CBF decreases for either the Chinese or English group across all active tasks compared to passive listening (Table 3, foci 9, 14, and 17–18; Table 4, foci 9–12; Table 5, focus 14; and Table 6, focus 15, 17, and 23–24). When comparing the tone or pitch tasks to passive listening, CBF decreases were observed in the medial frontal gyrus and/or anterior cingulate (Table 3, foci 10–11 and 16; Table 6, foci 16, 18–19, and 21–22).

Other foci of CBF decreases were task-specific and varied between groups: inferior parietal areas (Table 3, foci 13 and 19; Table 5, focus 15), bilateral superior frontal gyrus (Table 3, foci 12 and 15), left inferior temporal gyrus (Table 6, focus 20), right middle frontal gyrus (Table 5, foci 16–17), left uncus (Table 3, focus 20), right temporal pole (Table 3, focus 21), right premotor cortex (Table 7, focus 7), left posterior insula (Table 7, focus 2), and right supramarginal gyrus (Table 7, focus 8). For example, in the comparison of the tone task to passive listening, the Chinese group demonstrated CBF decreases in the right inferior parietal area and bilateral superior frontal gyrus, whereas the English group showed decreased activations in the left inferior parietal area, left uncus, and right temporal pole. The English group also exhibited a decreased activation in the left posterior insula when comparing the tone to the pitch task. With regard to the vowel task, the Chinese group exhibited CBF decreases in the right inferior parietal area and right middle frontal gyrus in the comparison of the vowel task to passive listening, whereas the English group showed decreased activations in the right premotor cortex and right supramarginal gyrus in the comparison of the vowel task to the pitch task. In the comparison of the pitch task to passive listening, one CBF decrease was observed in the left inferior temporal gyrus for the Chinese group only.

TABLE 5
Vowel minus Passive Listening

Region	Brodmann's area	Coordinates (mm)			<i>t</i> value
		<i>x</i>	<i>y</i>	<i>z</i>	
Blood flow increases					
Chinese					
1. L fusiform gyrus	37	-51	-49	-14	5.3
2. M vermal cerebellum	—	6	-71	-25	5.2
3. L inferior frontal gyrus (pars opercularis)	44/6	-44	5	27	5.1
4. L inferior frontal gyrus (pars triangularis)	45/46	-39	37	2	4.8
5. L premotor cortex	6	-33	-4	45	4.8
6. L inferior colliculus	—	-12	-33	-9	4.5
7. R paravermal cerebellum	—	8	-55	-16	4.4
8. L frontal operculum/anterior insula	45/—	-35	30	14	4.3
English					
9. M vermal cerebellum	—	1	-71	-29	4.8
10. L inferior frontal gyrus (pars opercularis)	44/6	-55	12	25	4.6
11. L superior temporal gyrus	22/42	-64	-24	7	4.5
12. L thalamus	—	-15	-10	16	4.4
13. R lateral cerebellum	—	26	-60	-34	4.4
Blood flow decreases					
Chinese					
14. M posterior cingulate gyrus	30/23	1	-51	16	-5.0
15. R angular gyrus	39	46	-55	27	-4.8
16. R middle frontal gyrus	8	24	21	45	-4.3
17. R middle frontal gyrus	9	21	32	34	-4.3

Note. Significant activation foci (blood flow increases and decreases) for subtraction of vowel condition minus passive listening condition. See also note under Table 3 for further details.

Of particular interest is the significant CBF *decrease* in the right premotor cortex for the English group when making a direct comparison of the vowel and the pitch tasks (Table 7; focus 7 = 39, 5, 27). These stereotaxic coordinates are virtually identical to those associated with a CBF *increase* in the same region of the right hemisphere when comparing the pitch task to passive listening (Table 6; focus 13 = 39, 3, 25). It is noteworthy that a CBF decrease was not observed in the right premotor cortex for the tone or consonant minus pitch subtraction.

Regions of Interest (ROIs)

Results of a four-way (hemisphere \times region \times task \times group) repeated-measures ANOVA performed on percentage changes in rCBF showed a three-way interaction among task, group, and ROI [$F(12, 204) = 2.80, p < .0015$]. A subsequent series of one-way ANOVAs was performed at each task for the five ROIs for the purpose of evaluating differences *between* groups. In the left pars opercularis, results indicated that the Chinese group showed significantly greater rCBF increases than the English group on the tone task [$F(1, 36) = 4.17, p < .0485$] as well as a strong tendency toward significance on the pitch [$F(1, 36) = 4.01, p < .0528$] and consonant tasks [$F(1, 36) = 4.11, p < .0502$]. In the right pars opercularis, the Chinese group showed significantly greater activation on the consonant task [$F(1, 36) = 5.81, p < .0212$] and pitch tasks [$F(1, 36) = 4.97, p < .0321$]. The Chinese group also showed significantly greater rCBF increases than the English group in the left pars triangularis on the pitch task [$F(1, 36) = 4.53, p < .0402$] and a tendency toward significance on

TABLE 6
Pitch minus Passive Listening

Region	Brodmann's area	Coordinates (mm)			t value
		x	y	z	
Blood flow increases					
Chinese					
1. L premotor cortex	6	-33	-6	45	4.8
2. M vermal cerebellum	—	-8	-46	-20	4.6
3. R lateral cerebellum	—	28	-64	-25	4.4
4. R thalamus	—	3	-19	-2	4.3
English					
5. R lateral cerebellum	—	24	-60	-36	5.1
6. R lateral cerebellum	—	26	-53	-25	5.0
7. R inferior frontal gyrus	44/45/9	37	26	20	4.7
8. M vermal cerebellum	—	-6	-73	-29	4.7
9. R frontal pole	10	39	53	-9	4.6
10. L basal ganglia	—	-17	3	16	4.4
11. L lingual gyrus	18	-1	-89	-7	4.3
12. M vermal cerebellum	—	3	-51	-22	4.3
13. R premotor cortex	6/44	39	3	25	4.3
14. R frontal operculum/anterior insula	47/—	35	23	0	4.3
Blood flow decreases					
Chinese					
15. M precuneus	31/7	-1	-53	36	-5.2
16. M medial frontal gyrus	11/10	3	46	-11	-4.8
17. M precuneus	31	-3	-62	25	-4.7
18. M medial frontal gyrus	10	-1	59	-4	-4.5
19. M medial frontal gyrus/anterior cingulate	10/32	-6	46	11	-4.5
20. L inferior temporal gyrus	20	-55	-13	-20	-4.5
English					
21. M medial frontal gyrus	10	1	48	2	-5.6
22. M medial frontal gyrus	10/9	-3	57	16	-4.9
23. M posterior cingulate gyrus	31/23	-1	-51	25	-4.4
24. M posterior cingulate gyrus	31/23	6	-42	32	-4.4

Note. Significant activation foci (blood flow increases and decreases) for subtraction of pitch condition minus passive listening condition. See also note under Table 3 for further details.

the consonant [$F(1, 36) = 3.77, p < .0602$], vowel [$F(1, 36) = 3.67, p < .0633$], and tone tasks [$F(1, 36) = 3.22, p < .0810$]. No other differences in rCBF changes between the Chinese and English groups reached the significance threshold.

DISCUSSION

Task-Dependent Laterality during Language Processing

Our findings support a task-dependent hypothesis of cerebral lateralization whereby LH mechanisms selectively mediate processing of linguistic information irrespective of acoustic cues or type of phonological unit (i.e., segmental or suprasegmental) and RH mechanisms are engaged for prosody-specific cues (Imaizumi et al., 1998). Left IFG regions are activated for Chinese listeners across the consonant, vowel, tone, and pitch tasks; right IFG areas are activated for English listeners on the pitch task. These findings are in accord with earlier lesion deficit (Gandour & Dardarananda, 1983; Hughes, Chan, & Ming, 1983; Yiu & Fok, 1995; Eng, Obler,

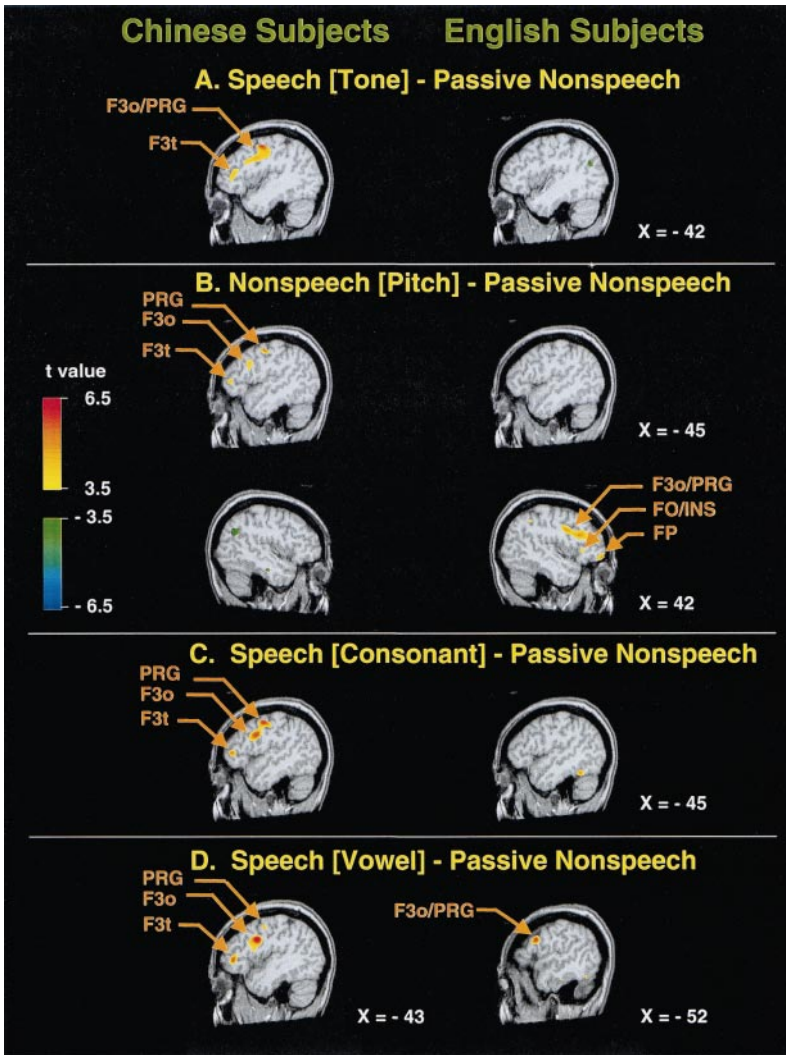


FIG. 1. Averaged PET subtraction images of Chinese (left) and English (right) subjects for the most pertinent foci are shown superimposed on parasagittal Talairach-transformed T1-weighted high-resolution anatomic images. Focal CBF changes are shown as a t statistic image; t values are coded by a color scale ranging from blue for the lowest values to red for the highest values (see Tables 3–7 and Results for precise t values of each focus). Stereotaxic coordinates, in millimeters, are derived from the human brain atlas of Talairach and Tournoux (1988) and refer to the medial–lateral position (x) relative to midline (negative = left; positive = right). Abbreviations: F3o = inferior frontal gyrus/pars opercularis; PRG = precentral gyrus/premotor cortex; F3t = inferior frontal gyrus/pars triangularis; FO = frontal operculum; INS = insula; FP = frontal pole. (A–D) Activity in the left inferior frontal gyrus (premotor cortex, pars opercularis, and pars triangularis) of Chinese subjects across all four active tasks relative to passive listening to pitch. With the exception of the vowel minus passive listening subtraction (D, right), there is an absence of activity for English subjects in left inferior frontal regions comparable to those shown for Chinese subjects. In the pitch minus passive listening subtraction, activation patterns in the left and right frontal lobes are complementary between language groups. Activation is visible in the right frontal lobe (premotor cortex, inferior frontal gyrus/pars opercularis, frontal operculum/anterior insula, and frontal pole) unilaterally for English subjects (cf. B, bottom right to top right), whereas activation is visible in the left frontal lobe unilaterally for Chinese subjects (cf. B, top left to bottom left). In the consonant and vowel minus passive listening subtractions, also visible is activation in the left fusiform gyrus for English subjects (C and D, right).

TABLE 7
Tone, Consonant, and Vowel minus Pitch

Region	Brodmann's area	Coordinates (mm)			<i>t</i> value
		<i>x</i>	<i>y</i>	<i>z</i>	
Tone minus pitch					
Blood flow increases					
English					
1. M medial frontal gyrus	11/32	6	30	-16	5.0
Blood flow decreases					
English					
2. L posterior insula	—	-39	-17	-4	-4.5
Consonant minus pitch					
Blood flow increases					
English					
3. M medial frontal gyrus	11/32	1	26	-16	6.1
4. L fusiform gyrus	37	-48	-49	-18	4.6
Vowel minus pitch					
Blood flow increases					
Chinese					
5. L inferior temporal gyrus	20	-35	-10	-32	5.2
English					
6. M medial frontal gyrus	11/32	1	30	-16	6.5
Blood flow decreases					
English					
7. R premotor cortex	6/44	39	5	27	-4.6
8. R supramarginal gyrus	40	48	-44	34	-4.3

Note. Significant activation foci (blood flow increases and decreases) for subtraction of tone, consonant, and vowel conditions minus pitch condition. See also note under Table 3 for further details.

Harris, & Abramson, 1996; Moen & Sundet, 1996), dichotic listening (Van Lancker & Fromkin, 1973; Ip & Hoosain, 1993; Moen, 1993), and functional neuroimaging (Gandour et al., 1998, 2000) data on tone perception. They do not support the cue-dependent hypothesis that left inferior frontal regions are specialized for rapid temporal processing of complex acoustic stimuli (cf. Schwartz & Tallal, 1980; Tallal, Miller, & Fitch, 1993). Chinese tones are of relatively long durations (250–350 ms) yet lateralized to the LH instead of the RH. Another cue-dependent hypothesis claims the RH mediates acoustic information contained in relatively low frequencies of the speech spectrum (Ivry & Leiby, 1993; Ivry & Robertson, 1998). Even with low-pass-filtered speech below 200 Hz, tone perception by Chinese listeners remains lateralized to the LH. Since these filtered stimuli have no linguistic significance for English listeners, their RH activation patterns argue for mediation of pitch processing per se in the RH, as predicted from previous lesion deficit (Johnsrude et al., 2000; Sidtis & Volpe, 1988; Robin, Tranel, & Damasio, 1990; Zatorre & Samson, 1991), dichotic listening (Sidtis, 1980), and functional neuroimaging (Démonet, Price, Wise, & Frackowiak, 1994; Zatorre et al., 1994; Binder et al., 1997) data. Yet another cue-dependent hypothesis argues that the LH mediates the temporal cues, while the RH mediates spectral cues (Robin et al., 1990; Van Lancker & Sidtis, 1992). To the contrary, our findings indicate that Chinese listeners process the predominant spectral cues associated with lexical tones (i.e., pitch patterns) in the LH instead of in the RH.

When judging pitch patterns in low-pass-filtered speech, Chinese listeners are apparently able to extract phonological information associated with the Chinese tones via recruitment of the left IFG regions. Presumably the absence of a natural speech context accounts for slightly diminished activation of the pars opercularis ($-48\ 14\ 27$, $t = 4.2$, n.s.; cf. Table 3, focus 3). In the consonant, vowel, and tone tasks, both segmental and suprasegmental information is accessible via natural speech. In the pitch task, only suprasegmental information is accessible via low-pass-filtered speech. Nevertheless, we still observe activation of left inferior frontal regions in Chinese listeners, as confirmed by the absence of CBF increases when subtracting the pitch task from the consonant, vowel, or tone task. Conversely, English listeners do not associate any pitch patterns in low-pass-filtered speech with Chinese tones and exhibit activation in right IFG areas. Such contrasting brain activation patterns between the two language groups undoubtedly reflect the powerful effect of a listener's language experience on brain organization for language. This functional lateralization effect is not predicted by acoustic cues but instead by the specific task depending on the listener's language experience. Pitch processing is lateralized to the LH only when the pitch patterns are phonologically significant to the listener; otherwise to the RH when the pitch patterns are linguistically irrelevant.

These findings strengthen the view that encoding of auditory characteristics of the speech signal is not only based on complex acoustic properties, but also takes into account their functional significance in a particular language. The LH is activated during tasks that require subjects to extract linguistic information based on their language experience. Acoustic properties of Chinese tones notwithstanding, LH lateralization is strongly shaped by a subject's *analytic* listening strategy (Mazziotta, Phelps, Carson, & Kuhl, 1982). Chinese listeners appear to use a phonologically based, analytic strategy in making perceptual judgments, resulting in left IFG activations. English listeners, on the other hand, employ a nonlinguistic strategy in making perceptual judgments of pitch patterns, resulting in right IFG activations.

It is worth noting that we also observe activation of the pars opercularis/premotor cortex (BA 44/6) *in the RH* for Chinese listeners at levels a little below the significance threshold on both the tone (35, 8, 25; $t = 3.6$) and pitch (37, 10, 27; $t = 3.2$) tasks. This finding is consistent with the hypothesis that pitch processing, in and of itself, not only engages RH mechanisms, but also that it may be "attracted" to the LH when it is used to signal linguistic prosody (Shipley-Brown, Dingwall, Berlin, Yeni-Komshian, & Gordon-Salant, 1988). In our study, pitch processing is still implemented in the RH but to a lesser degree than in the LH because of its linguistic function. This hypothesis accounts for the collaboration between the two hemispheres and assumes that differential lateralization occurs as a result of an interaction between the auditory property and its function. Interhemispheric cooperation in the perception of linguistic prosody may be accomplished by connections via the corpus callosum (Klouda, Robin, Graff-Radford, & Cooper, 1988).

Convergence with Previous Neuroimaging and Behavioral Studies

Chinese listeners show increased activation of the premotor cortex and pars opercularis across all four active tasks. This swath of activity, extending from the premotor area to the posterior Broca's area, possibly reflects subvocal rehearsal of phonological information (Paulesu, Frith, & Frackowiak, 1993). This phonological rehearsal circuit is consistently activated regardless of whether subjects are focusing on segmental or suprasegmental characteristics of the syllable. These foci in the left posterior frontal lobe (BA 44/6) are consistent with previous neuroimaging studies of consonant perception in English and French (Demonet et al., 1992, 1994; Zatorre et al., 1992,

1996), tone perception in Thai (Gandour et al., 1998, 2000), and phonemic verbal fluency in English (Paulesu et al., 1997). Thus, when either segmental or suprasegmental extraction and identification of phonetic units is required, articulatory recoding (Liberman & Mattingly, 1985, 1989) is implemented in the posterior portion of Broca's area.

Chinese listeners extract and process phonological information about tones from both natural speech and low-pass-filtered speech. Extraction of phonological information is confirmed by the absence of CBF increases in the left IFG regions when subtracting the pitch task from the consonant, vowel, or tone task. In the pitch task, Chinese listeners are required to extract tonal contours in the absence of lexical context. This phenomenon of prelexical phonological processing of prosody is consistent with a behavioral study of lexical stress perception in English (Mattys, 2000), showing fine stress distinctions are possible without lexical information and with minimal speech normalization. This finding differs from our PET study of Thai tone perception (Gandour et al., 2000), in which activation of the left frontal operculum was observed for Thai listeners when comparing the tone task to the pitch task. Methodological differences between these two studies, however, are not inconsequential, especially in terms of task, memory demands, and stimuli. It is therefore difficult to say to what extent this discrepancy between Thai listeners' processing of Thai tones and Chinese listeners' processing of Chinese tones reflects language-specific effects on tonal processing.

These findings of lateralization to the LH for lexical tone perception are consistent with previous lesion deficit (Gandour & Dardarananda, 1983; Hughes et al., 1983; Yiu & Fok, 1995; Eng et al., 1996; Moen & Sundet, 1996), dichotic listening (Van Lancker & Fromkin, 1973; Ip & Hoosain, 1993; Moen, 1993), and PET (Gandour et al., 1998, 2000) data.

As predicted, no activation in left frontal regions occurs for English listeners on either the tone or pitch task because they were judging auditory features of Chinese tones that are not phonologically significant in the English language. Similarly, when judging Thai tones, Chinese as well as English listeners fail to show activation of the left IFG (Gandour et al., 2000). English is not a tone language. The lack of activation in this region for English subjects reflects their unfamiliarity with linguistically significant pitch patterns at the syllable level. Although Chinese is a tone language, Chinese listeners are unfamiliar with the specific pitch patterns associated with Thai tones and are not expected to process Thai tones the same way as Chinese tones. The lack of neural activation in the left IFG for Chinese listeners unequivocally points to language-specific phonological information as the trigger for lateralization to left frontal regions.

English listeners show activation in right frontal regions on the pitch task, including premotor cortex, IFG, frontal operculum/anterior insular, and frontal pole. For English subjects, pitch processing is lateralized to the RH since the pitch patterns do not carry any linguistic meaning. RH lateralization for pitch perception is consistent with earlier findings from the lesion deficit (Sidtis & Volpe, 1988; Zatorre & Samson, 1991), dichotic listening (Sidtis, 1980; Wioland, Rudolf, Metz-Lutz, Mutschler, & Marescaux, 1999), and functional neuroimaging (Jones, Longe, & Pato, 1998; Zatorre et al., 1992, 1994) literature. English listeners do not show similar activation in the right IFG area when comparing Chinese tones with passive listening to filtered speech. Because of their unfamiliarity with Chinese tones, English listeners are likely to be distracted by segmental information (i.e., consonants and vowels) in their attempt to tease apart the tonal patterns. Indeed, in a PET study of Thai tone perception (Gandour et al., 2000), English listeners show activation in the left IFG even when comparing Thai tones to resting baseline. Thus, the degree of interference between

segmental and suprasegmental processing is likely to vary depending on an individual's language experience.

Functional Heterogeneity in the Left Frontal Lobe for Working Memory

Our findings reveal a functional separation between anterior and posterior areas of the left frontal lobe corresponding to executive and storage processes, respectively, of verbal working memory. Chinese listeners show increased activation of the premotor cortex and pars opercularis across all four active tasks. These clusters of activation, running from the premotor area to Broca's area, arguably reflect subvocal rehearsal processes that refresh the contents of a phonological buffer for temporary storage of phonological information (Baddeley, 1992; Smith & Jonides, 1999). This rehearsal circuit is consistently activated regardless of whether subjects are focusing on segmental or suprasegmental characteristics of the syllable. These foci in the left posterior frontal lobe (BA 44/6) are consistent with previous PET studies of consonant perception in English (Demonet et al., 1992, 1994; Zatorre et al., 1992, 1996) and tone perception in Thai (Gandour et al., 1998, 2000) and with a theory of speech perception which involves articulatory recoding when segmental or suprasegmental extraction and identification of phonetic units is required (Liberman & Mattingly, 1985, 1989).

Interestingly, Chinese listeners show activation in the anterior pars triangularis of the left frontal lobe, extending beyond the rostral tip of the anterior horizontal ramus (BA 45/46). This anteroventral focus is clearly distinct from the activations that cluster posterodorsally in the frontal lobe and may reflect one or more executive processes implemented by regions in prefrontal cortex (PFC) (Smith & Jonides, 1999). Two putative, executive processes—attention/inhibition and active comparison—are likely to have controlled Chinese listeners' on-line encoding and retrieval of long-term phonological representations in guiding their performance on tasks that require the maintenance of information across a distracted delay. To perform the tasks successfully, subjects are required to focus attention selectively on segmental information and inhibit tonal information, or vice versa, in addition to inhibiting the intervening distractor syllables during the delay period. Subjects are also required to make active comparisons of selected components of Chinese syllables. These executive processes are to be distinguished from others (e.g., manipulation and monitoring). Different executive processes may be subserved by functionally distinct subdivisions of lateral PFC in both spatial and nonspatial memory tasks (Petrides, 1994). Active comparisons of information held in working memory may be localized to ventral regions of PFC (D'Esposito et al., 1998). Indeed, the coordinates on this verbal memory task are in close proximity to those reported on a spatial working memory task (Owen, Evans, & Petrides, 1996, p. 35: -46, 41, 6; cf. Tables 3–6). In both tasks, subjects selectively attend to stimuli in predictable locations, inhibit distracting elements, and make an active comparison. Our findings lend support to a dorsal/ventral subdivision of the PFC that depends on the type of processing performed on information held in working memory.

It is well known that verbal working memory is sensitive to different aspects of phonology (Baddeley, 1992), including aspects of tonal phonology (e.g., Chinese: Xu, 1991; Cantonese: Cutler & Chen, 1997). Our findings suggest that these executive processes may be influenced by the structure of phonological representations. The prevailing view is that phonological information is specified on autonomous, interconnected *tiers* (Goldsmith, 1990). In Chinese, consonant and vowel information is specified on the segmental tier, tones on the tonal tier. The slightly different levels

of activation on the consonant and vowel tasks as compared to the tone and pitch tasks may reflect sensitivity of executive processes to tier representations.

The English group shows no significant activity in the pars triangularis on any of the tasks. Chinese is an unfamiliar, foreign language for English listeners. Without knowledge of Chinese phonology, they would not be expected to exploit executive processes sensitive to segmental and suprasegmental components of Chinese words, such as encoding and retrieval of Chinese phonological representations from long-term memory.

Prelexical Phonological Processing of Tones

An autonomous, prosodic tier has been well established in phonological theory (Goldsmith, 1990). In Chinese, tones are assigned to the prosodic or suprasegmental tier, consonants and vowels to the segmental tier. Prosody has also been encapsulated in the phonological component of psycholinguistic processing models of both speech perception (Cutler & Clifton, 1999) and production (Levelt, 1999). Prosodic information is important in *prelexical* processes of phonological encoding and has been shown to constrain word activation in tonal (Cutler & Chen, 1997), pitch accent (Cutler & Otake, 1999), and lexical stress languages (Mattys, 2000). The difficulty in identifying candidate areas for a prelexical phonological processing stage with functional neuroimaging is that speech input also triggers lexicosemantic processing. “To avoid this confound, the candidate areas should be demonstrated to be insensitive to the lexical status of the stimuli, yet respond to phonemes, illegal non-words, pseudo-words, or words” (Price, Indefrey, & van Turenhout, 1999, p. 221). Our findings suggest that the LH pars opercularis and pars triangularis may be involved in prelexical phonological processing of Chinese tones. These candidate areas appear to be insensitive to the lexical status of the stimuli. The tone task required judgments of Chinese pseudowords; the pitch task required judgments of nonwords. Yet Chinese listeners apparently respond to tonal phonemes at the prelexical stage of processing in both tasks, as evidenced by the absence of any significant increases in rCBF on the tone task relative to pitch.

A comment is in order about the absence of bilateral activation in the STG for the Chinese group in any of paired-image subtractions. In this study, due to task requirements, in which passive listening served as the baseline condition, activation in the STG would be subtracted out in any of the comparisons between active tasks and the passive listening control. The absence of bilateral activation may arguably reflect that the STG is involved in early perceptual analysis of complex incoming auditory signals, as suggested in Zatorre et al. (1996). However, limitations of PET experimental design do not permit us to tease apart different stages of perceptual processing of auditory information.

Revising the Classic Model of Language Localization

Due to limitations of the classic model of language localization, which fail to account for the functional overlap between Broca’s and Wernicke’s areas in producing and comprehending spoken language as well as for the involvement of other regions, both cortical and subcortical, a refined and extended model of language specialization is being developed to modify the classic Wernicke–Geschwind model (see Dronkers, Pinker, & Damasio, 2000, for review). Instead of a unidirectional pathway from posterior (Wernicke’s) to anterior (Broca’s) regions, the arcuate fasciculus provides for a bidirectional pathway between frontal and temporoparietal regions. Neuroimaging data show that left frontal areas participate in language comprehension (Binder, Frost,

Hammeke, Rao, & Cox, 1996; Binder et al., 1997) and speech perception (Zatorre et al., 1992, 1996; Gandour et al., 2000). Similarly, our findings support the notion that left frontal regions mediate *perception* as well as production of both segmental and suprasegmental aspects of the speech signals. Involvement of left frontal regions in speech perception is compatible with the notion that articulatory recoding of speech signals occurs during speech perception (Liberman & Mattingly, 1985, 1989). Executive functions (selective attention, inhibition, and active comparison) that mediate phonological processing appear to be a function of the rostral portion of the left pars triangularis (BA 45). Conversely, the left planum temporale and left posterior perisylvian cortex, i.e., secondary auditory cortex, have been implicated in a speech production task that masked out auditory input (Paus, Perry, Zatorre, Worsley, & Evans, 1996). These audio–vocal interactions in the auditory system are seen to be analogous to visuo-oculomotor interactions in the visual system. Converging evidence from lesion studies show that Broca’s aphasics have comprehension problems in processing grammatical aspects of language (see Berndt, Mitchum, & Wulfeck, 1996; Caplan, Hildebrandt, & Makris, 1996, for review) and that Wernicke’s and conduction aphasics exhibit ‘subtle phonetic deficits’ in speech production (see Vijayan & Gandour, 1995, for review).

According to this revised model, other cortical regions in the LH may be critically involved in language processing, including ‘prefrontal and cingulate areas that implement executive control and mediation of necessary memory and attentional processes’ (Dronkers et al., 2000, p. 1175). Our findings suggest that the rostral portion of the left pars triangularis may play an executive role in mediating between long-term phonological representations and a short-term phonological buffer. Consistent with previous studies on the attentional functions of anterior cingulate cortex and medial frontal gyrus (Cabeza & Nyburg, 1997; Carter et al., 1998), the anterior cingulate gyrus (BA 32/6) is activated for the Chinese group in tone minus passive listening; the medial frontal gyrus (BA 11/32) for the English group in consonant, vowel, or tone minus pitch.

The role of the insula in language remains somewhat controversial (see Ardila, Benson, & Flynn, 1997; Ardila, 1999; Flynn, Benson, & Ardila, 1999, for review). Anatomically, insular cortex is connected to pre- and postcentral opercular cortex; inferior frontal cortex; adjacent auditory cortex (BA 22 and BA 41/42); several limbic structures, including anterior cingulate gyrus; as well as several thalamic nuclei (Mesulam & Mufson, 1985). Thus, it is well positioned anatomically to play a crucial role in neural networks involved in verbal communication. Damage to the insula has been proposed as one of the possible neural substrates of conduction aphasia (Damasio & Damasio, 1980), apraxia of speech (Dronkers, 1996), auditory agnosia (Habib et al., 1995), dyslexia (Paulesu et al., 1996), and verbal memory (Manes, Springer, Jorge, & Robinson, 1999). Previous neuroimaging studies have implicated insular cortex in tasks involving articulation and speech coordination (Wise, Greene, Buchel, & Scott, 1999), short-term verbal memory (Paulesu et al., 1993), and music (Zatorre et al., 1994). Its precise function has been variously interpreted to be associated with coordination of speech articulation (Dronkers, 1996; Wise et al., 1999), subvocal rehearsal in verbal short-term memory (Paulesu et al., 1993), subvocal rehearsal in vocal pitch (Zatorre et al., 1994), and conversion of auditory-visual signals to phonological codes (Paulesu et al., 1996).

Our findings may be brought to bear on the issue of which functions are governed by the left anterior insula vis-à-vis the posterior portion of the IFG and premotor cortex. Left anterior insular activations for Chinese listeners exceed statistical significance on the vowel task and are just below the threshold of significance on the consonant task ($-42, 19, -4; t = 3.8$). The absence of activation on the tone and

pitch tasks is likely due to Chinese listeners' being able to extract pitch patterns of Chinese tones in the nonspeech condition. Assuming that the left anterior insula is necessary for speech articulation, articulatory recoding of tones in these two perception tasks is apparently subtracted out by comparison to a passive listening condition in which homologous pitch patterns were presented in low-pass-filtered speech. Thus, insular activation for consonants and vowels arguably reflects Chinese listeners' articulatory recoding of spectral information above 200 Hz as compared to low-pass-filtered speech. Regardless of whether speech is overt or covert, articulatory rehearsal processes assist in the maintenance of information in the phonological buffer. Assuming that covert rehearsal processes engage Broca's area (BA 44/6), activation of this region for Chinese listeners is fully predictable when they are asked to perform a task that requires extraction of Chinese segmental (consonants and vowels) and suprasegmental (tones) information in natural speech and suprasegmental (tones) information in low-pass-filtered speech.

These left anterior insular activations also support the notion of possibly two separate pathways projecting from the temporo-parietal to the frontal language area, the first via the arcuate fasciculus, the second via the insula (Lichtheim, 1885; McCarthy & Warrington, 1984). All the Chinese stimuli (i.e., meaningless sequences of Chinese monosyllables) carried a minimum of lexical information. Left anterior insular activation may have resulted from increased demand placed on speech articulation within a context devoid of semantic content. Under normal circumstances, both processing routes are assumed to be available to listeners. Access to one route versus another will vary depending upon task demands. In this study, the more accessible route for Chinese listeners would likely be the nonsemantic route. The insula appears to be a crucial element in this nonsemantic processing route (cf. Gandour et al., 2000). It is hypothesized to be a link between an auditory input signal and a parsed string of phonological units, i.e., consonants and vowels, and by extension, tones (Paulesu et al., 1996).

Right anterior insular activation for English listeners on the pitch task relative to passive listening supports the idea that structures within the RH are specialized for at least some aspects of pitch processing (Imaizumi et al., 1998; Robin et al., 1990; Zatorre et al., 1992, 1994). English listeners presumably do not associate these pitch patterns with distinctive features of Chinese tones and consequently we see no attraction to the LH. Instead, complex pitch perception for English listeners elicited activation predominantly in the right anterior insula as well as other IFG regions in the RH. Perhaps this region of the insula is critical for production, covert as well as overt, of all complex volitional vocalizations, including vocal fundamental frequency and articulation (Perry et al., 1999). This is just the opposite of what we observe for Chinese listeners. These opposite hemispheric lateralization effects at the level of the anterior insula are likely due to the linguistic significance of pitch patterns. These findings on auditory perception are also consistent with a recent fMRI study of overt speaking and singing (Riecker, Ackermann, Wildgruber, Dogil, & Grodd, 2000). Reproduction of a nonlyrical tune elicited activation predominantly in the right motor cortex, the *right anterior insula*, and the left cerebellum, whereas the opposite response pattern emerged during the speech task.

The cerebellum may play a role in attention coordination, imagery, covert action, and working memory (see Ivry & Fiez, 2000, for review). Medial and right lateral cerebellum are typically activated during verbal working memory tasks. One component of Baddeley's (1992) model of working memory involves maintenance of linguistic information in a short-term buffer via covert articulatory rehearsal. Our data are consistent with the notion that the cerebellum, in addition to Broca's area and the supplementary motor cortex, may comprise the neural basis of an articulatory

rehearsal system (Fiez et al., 1996). Cerebellar activations are found across all active tasks compared to passive listening for both Chinese and English groups. Even though overt motor demands (i.e., clicking a mouse button) are equated in the active and passive tasks, two sets of foci are distinguishable from each other. One set of foci is located laterally in the right cerebellum ($x \geq +24$ mm); the other set is located medially in the vermal cerebellum ($x \leq \pm 8$ mm). In terms of covert articulatory rehearsal, the cerebellum “may contribute to motoric aspects of internal speech representation in much the same way that it contributes to overt speech production” (Ivry & Fiez, 2000, p. 1008). Cerebellar activation also correlates with task difficulty (Ivry, 1997) and coordination of attention (Courchesne & Allen, 1997). Increased cerebellar activation in the active tasks compared to the passive baseline might reflect the concurrent preparation and planning of candidate responses (i.e., same or different response) as well as ensuring that the selected information (i.e., the first and last targeted syllables in each list) is processed with maximal efficiency to satisfy current task demands in a limited time (i.e., response time of 1.5 s).

Blood Flow Decreases

Active minus passive decreases in rCBF may reflect decreased activity related to active task processes or suspension of activity related to ongoing processes in the passive condition (Shulman et al., 1997). The interpretation of blood flow decreases during active tasks depends on their consistency across tasks. “If each active task inhibits different areas . . . consistent decreases should not be found. If each active task inhibits the same area . . . similar decreases should be found across tasks. If decreases represent ongoing processes in the passive state, decreases should also generalize active tasks, since the passive condition is similar across tasks” (p. 649).

In this study, the posterior cingulate gyrus and/or precuneus (BA 31/7) showed consistent decreases across all active tasks for both Chinese and English listeners as compared to a common, passive listening baseline. Ongoing processes in the passive state that may have been suspended during active tasks include unconstrained verbal thought processes, monitoring of the external environment, or monitoring of the emotional state (Shulman et al., 1997, pp. 660–662).

Decreases were observed in the medial frontal gyrus (BA 10) and anterior cingulate (BA 32/24) on the tone and pitch tasks, but not on the consonant and vowel tasks, for both English and Chinese listeners. These inconsistent decreases across tasks may reflect varying attentional demands of suprasegmental vis-à-vis segmental processing.

Other foci of CBF decreases that appear inconsistently across tasks or groups might reflect processes related to active tasks and/or linguistic experience. For example, when comparing pitch to passive listening, rCBF deactivation in the left inferior temporal gyrus was found for the Chinese group only. This region has been identified as playing an intermediary or mediational role in lexical retrieval (Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996). Stimuli for the pitch task were nonwords, i.e., not recognizable as speech; for the consonant, vowel, and tone tasks, stimuli were pseudowords and recognizable as speech. Chinese listeners might be expected to inhibit lexical retrieval processes during the pitch task.

Other interpretations of blood flow decreases may require comparisons of active tasks to one another as well as active tasks to a passive condition. For the English group, decreased CBF in the premotor cortex of the RH was found when comparing two active tasks to one another, namely vowel minus pitch. This deactivation makes sense only when you consider that there was an increase in the same area of the RH

during the pitch task relative to passive listening as well as an increase in the homologous region of the LH during the vowel task relative to passive listening. This blood flow decrease in the RH presumably reflects English listeners' LH bias when attempting to parse vowels in Chinese speech.

Effects of Task Performance on Brain Activation Patterns

Task performance differences between Chinese and English groups notwithstanding, it is unlikely that such task-related differences can account for their differential patterns of brain activation. Both Chinese and English subjects perform at a high level of proficiency on both consonant and vowel tasks. Yet differences in brain activation patterns are found between the two language groups. Chinese listeners show increased activation in the left premotor cortex, pars opercularis, and pars triangularis on both consonant and vowel tasks, whereas English listeners show similar activation in the left premotor cortex and pars opercularis on the vowel task only. English listeners perform at a significantly lower level of accuracy on the tone and pitch tasks than on the consonant and vowel tasks and take longer to respond on the tone task than on the others. Yet significant activation in right IFG regions is found on the pitch task only.

When judging Chinese vowels, but not Chinese consonants, English listeners are evidently able to extract phonetic information via recruitment of left IFG regions. Task difficulty is an unlikely explanation because of the absence of significant differences in response accuracy or response time between the consonant and vowel tasks. This dissociation between vowels and consonants for English listeners may be attributed to several factors. First, it could be due to the different size of Chinese consonant and vowel inventories. Mandarin Chinese has 21 consonants and six vowels (Cheng, 1973). In this study, 19 of 21 Chinese consonant phonemes were represented and six of six vowel phonemes. The fewer number of Chinese vowels than consonants may have made it easier for English listeners to make discrimination judgments of vowels. Second, Chinese vowels (/i y u ə o a/) are phonetically more similar to English vowels than Chinese consonants are to English consonants. Of Chinese vowels presented in this study, only occurrences of /y/ (3.5%), a high front rounded vowel, are likely to be totally unfamiliar to English listeners. This phonetic similarity may have enabled English listeners to extract vowel information from the Chinese stimuli. Of Chinese consonants (insert 21 consonant symbols here) presented in this study, 35% are likely to sound very foreign to English listeners: voiceless retroflex (/ʂ/), alveolopalatal (/ç/), and velar fricatives (/x/); and voiceless unaspirated and aspirated alveolar (/ts, ts^h/), retroflex (/tʂ/), and alveolopalatal (/tç, tç^h/) affricates. Third, durational differences between Chinese consonants and vowels may be relevant. Acoustic information for Chinese vowels is transmitted over a longer temporal domain than for Chinese initial consonants. The longer duration of Chinese vowels might provide a more salient perceptual cue to naïve English listeners and contribute to a more stable memory trace when making discrimination judgments.

The group difference (Chinese vs. English) in insular activation cannot be accounted for by the degree to which the task is learned or automatic (cf. Raichle et al., 1994). Both groups received the same instructions and same amount of exposure to the auditory stimuli prior to scanning. Therefore, it is unlikely that differences between Chinese and English listeners are due to a practice effect. It appears that the degree to which the insula is engaged in these auditory discrimination tasks reflects different cognitive strategies that listeners employ depending on their language experience.

CONCLUSIONS

Our findings support a task-dependent hypothesis of functional lateralization as well as illustrate how memory and language may interact in phonological processing. In agreement with previous imaging studies of working memory (Smith & Jonides, 1999), our data point to the localization of the phonological loop and central executive processes to posterodorsal and anteroventral regions, respectively, within the left IFG. Also borne out is the promise of crosslinguistic approaches to functional imaging (cf. Gandour et al., 1998, 2000). Between-group differences in brain activation support a *language-specific* fractionation of executive and storage processes that would otherwise be obscured in a single-language study. Consistent with other functional imaging studies of auditory processing (e.g., Binder et al., 1996, 1997; Zatorre et al., 1992, 1996), our crosslinguistic findings on tonal perception support a revision of the classic model of language localization. Left frontal regions appear to mediate *perception* as well as production of both segmental and suprasegmental aspects of the speech signal.

REFERENCES

- Ardila, A., Benson, F., & Flynn, F. (1997). Participation of the insula in language. *Aphasiology*, **11**, 1159–1169.
- Ardila, A. (1999). The role of the insula in language: An unsettled question. *Aphasiology*, **13**, 79–87.
- Baddeley, A. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, **18**, 362–365.
- Baddeley, A. (1992). Working memory. *Science*, **255**, 556–559.
- Baum, S., & Pell, M. (1999). The neural bases of speech prosody: Insights from lesion studies and neuroimaging. *Aphasiology*, **13**, 581–608.
- Beck, A. (1978). *Beck Depression Inventory*. New York: The Psychological Corporation/Harcourt Brace Jovanovich.
- Berndt, R., Mitchum, C., & Wulfeck, B. (1996). Comprehension of reversible sentences in “agrammatism”: A meta-analysis. *Cognition*, **58**, 289–308.
- Binder, J., Rao, S., Hammeke, T., Frost, J., Bandettini, P., & Hyde, J. (1994). Effects of stimulus rate on signal response during functional magnetic resonance imaging of auditory cortex. *Cognitive Brain Research*, **2**, 31–38.
- Binder, J., Frost, J., Hammeke, T., Rao, S., & Cox, R. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain*, **119**, 1239–1247.
- Binder, J., Frost, J., Hammeke, T., Cox, R., Rao, S., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience*, **17**, 353–362.
- Cabeza, R., & Nyberg, L. (1997). Imaging cognition: An empirical review of PET studies with normal subjects. *Journal of Cognitive Neuroscience*, **9**, 1–26.
- Caplan, D., Hildebrandt, N., & Makris, N. (1996). Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain*, **119**, 933–949.
- Carter, C., Braver, T., Barch, D., Botvinick, M., Noll, D., & Cohen, J. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, **280**, 747–749.
- Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley, CA: Univ. of California Press.
- Cheng, C. (1973). *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton.
- Collins, L., Holmes, C., Peters, T., & Evans, A. (1995). Automatic 3-D model-based neuroanatomical segmentation. *Human Brain Mapping*, **3**, 190–208.
- Courchesne, E., & Allen, G. (1997). Prediction and preparation, fundamental functions of the cerebellum. *Learning and Memory*, **4**, 1–35.
- Cutler, A., & Chen, H.-C. (1997). Lexical tone in Cantonese spoken-word processing. *Perception and Psychophysics*, **59**, 165–179.
- Cutler, A., & Clifton, C. (1999). Comprehending spoken language: A blueprint of the listener. In C. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 123–166). New York: Oxford Univ. Press.

- Cutler, A., & Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese. *Journal of the Acoustical Society of America*, **105**, 1877–1888.
- Damasio, H., & Damasio, A. (1980). The anatomical basis of conduction aphasia. *Brain*, **103**, 337–350.
- Damasio, H., Grabowski, T., Tranel, D., Hichwa, R., & Damasio, A. (1996). A neural basis for lexical retrieval. *Nature*, **380**, 499–505.
- Demonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.-L., Wise, R., Rascol, A., & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, **115**, 1753–1768.
- Demonet, J., Price, C., Wise, R., & Frackowiak, R. (1994). A PET study of cognitive strategies in normal subjects during language tasks: Influence of phonetic ambiguity and sequence processing on phoneme monitoring. *Brain*, **117**, 671–682.
- D'Esposito, M., Aguirre, G., Zarahn, D., Ballard, D., Shin, R., & Lease, J. (1998). Functional MRI studies of spatial and nonspatial working memory. *Cognitive Brain Research*, **7**, 1–13.
- Dronkers, N. (1996). A new brain region for coordinating speech articulation. *Nature*, **384**, 159–161.
- Dronkers, N., Pinker, S., & Damasio, A. (2000). Language and the aphasias. In E. Kandel, J. Schwartz, & T. Jessell (Eds.), *Principles of neural science* (4th ed., pp. 1169–1187). New York: McGraw-Hill.
- Ebeling, U., Steinmetz, H., Huang, Y., & Kahn, T. (1989). Topography and identification of the inferior precentral sulcus in MR imaging. *American Journal of Neuroradiology*, **10**, 937–942.
- Eng, N., Obler, L., Harris, K., & Abramson, A. (1996). Tone perception deficits in Chinese-speaking Broca's aphasics. *Aphasiology*, **10**, 649–656.
- Fiez, J., Raife, D., Balota, J., Schwarz, J., Raichle, M., & Petersen, S. (1996). A positron emission tomography study of the short-term maintenance of verbal information. *Journal of Neuroscience*, **16**, 808–822.
- Flynn, F., Benson, F., & Ardila, A. (1999). Anatomy of the insula—Functional and clinical correlates. *Aphasiology*, **13**, 55–78.
- Friston, K., Frith, C., Liddle, P., Dolan, R., Lamertsma, A., & Frackowiak, R. (1990). The relationship between global and local changes in PET scans. *Journal of Cerebral Blood Flow Metabolism*, **10**, 458–466.
- Friston, K., Frith, C., Liddle, P., & Frackowiak, R. (1991). Comparing functional PET images: The assessment of significant change. *Journal of Cerebral Blood Flow Metabolism*, **11**, 81–95.
- Gandour, J. (1998a). Aphasia in tone languages. In P. Coppens, A. Basso, & Y. Lebrun (Eds.), *Aphasia in atypical populations* (pp. 117–141). Hillsdale, NJ: Erlbaum.
- Gandour, J. (1998b). Phonetics and phonology. In B. Stemmer & H. Whitaker (Eds.), *Handbook of neurolinguistics* (pp. 207–219). New York: Academic Press.
- Gandour, J., & Dardarananda, R. (1983). Identification of tonal contrasts in Thai aphasic patients. *Brain and Language*, **18**, 98–114.
- Gandour, J., Wong, D., & Hutchins, G. D. (1998). Pitch processing in the human brain is influenced by language experience. *NeuroReport*, **9**, 2115–2119.
- Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., & Hutchins, G. D. (2000). A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience*, **12**, 207–222.
- Goldsmith, J. (1990). *Autosegmental & metrical phonology*. Cambridge, MA: Basil Blackwell.
- Habib, M., Daquin, G., Milandre, L., Royere, M., Rey, M., Lanteri, A., Salamon, G., & Khalil, R. (1995). Mutism and auditory agnosia due to bilateral insular damage: Role of the insula in human communication. *Neuropsychologia*, **33**, 327–339.
- Hughes, C., Chan, J., & Ming, S. (1983). Aprosodia in Chinese patients with right cerebral hemisphere lesion. *Archives of Neurology*, **40**, 732–736.
- Imaizumi, S., Mori, K., Kiritani, S., Hosoi, H., & Tonoike, M. (1998). Task-dependent laterality for cue decoding during spoken language processing. *NeuroReport*, **9**, 899–903.
- Ip, K., & Hoosain, R. (1993). Dichotic listening of Chinese and English words. *Psychologia*, **36**(3), 140–143.
- Ivry, R. (1997). Cerebellar timing systems. *International Review of Neurobiology*, **41**, 555–573.
- Ivry, R., & Fiez, J. (2000). Cerebellar contributions to cognition and imagery. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 999–1011). Cambridge, MA: MIT Press.
- Ivry, R., & Leiby, P. (1993). Hemispheric differences in auditory perception are similar to those found in visual perception. *Psychological Science*, **4**, 41–45.
- Ivry, R., & Robertson, L. (1998). *The two sides of perception*. Cambridge, MA: MIT Press.

- Johnsrude, I., Penhume, V., & Zatorre, R. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, **123**, 155–163.
- Jones, S., Longe, O., & Pato, M. (1998). Auditory evoked potentials to abrupt pitch and timbre change of complex tones: Electrophysiological evidence of 'streaming'? *Electroencephalography and Clinical Neurophysiology*, **108**, 131–142.
- Klouda, G., Robin, D., Graff-Radford, N., & Cooper, W. (1988). The role of callosal connections in speech prosody. *Brain and Language*, **35**, 154–171.
- Lichtheim, L. (1885). On aphasia. *Brain*, **7**, 433–484.
- Levelt, W. (1999). Producing spoken language: A blueprint of the listener. In C. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83–122). New York: Oxford Univ. Press.
- Liberman, A., & Mattingly, I. (1985). The motor theory of speech perception revised. *Cognition*, **21**, 1–36.
- Liberman, A., & Mattingly, I. (1989). A specialization for speech perception. *Science*, **243**, 489–494.
- Manes, F., Springer, J., Jorge, R., & Robinson, R. (1999). Verbal memory impairment after left insular cortex infarction. *Journal of Neurology Neurosurgery Psychiatry*, **67**, 532–534.
- Mattys, S. (2000). The perception of primary and secondary stress in English. *Perception & Psychophysics*, **62**, 253–265.
- Mazziotta, J., Phelps, M., Carson, R., & Kuhl, D. (1982). Tomographic mapping of human cerebral metabolism: Auditory stimulation. *Neurology*, **32**, 921–937.
- McCarthy, R., & Warrington, E. (1984). A two-route model of speech production: Evidence from aphasia. *Brain*, **107**, 463–485.
- Mesulam, M., & Mufson, E. (1985). The insula of Reil in man and monkey: Architectonics, connectivity and function. In A. Peters & E. Jones (Eds.), *Cerebral cortex*, (pp. 179–226). New York: Plenum.
- Minoshima, S., Berger, K., Lee, K., & Mintun, M. (1992). An automated method for rotational correction and centering of three-dimensional functional brain images. *Journal of Nuclear Medicine*, **33**, 1579–1585.
- Minoshima, S., Koeppe, R., Mintun, M., Berger, K., Taylor, S., Frey, K., & Kuhl, D. (1993). Automated detection of the intercommissural line for stereotactic localization of functional brain images. *Journal of Nuclear Medicine*, **34**, 322–329.
- Minoshima, S., Koeppe, R., Frey, K., & Kuhl, D. (1994). Anatomic standardization: Linear scaling and nonlinear warping of functional brain images. *Journal of Nuclear Medicine*, **35**, 1528–1537.
- Moen, I. (1993). Functional lateralization of the perception of Norwegian word tones—Evidence from a dichotic listening experiment. *Brain and Language*, **44**, 400–413.
- Moen, I., & Sundet, K. (1996). Production and perception of word tones (pitch accents) in patients with left and right hemisphere damage. *Brain and Language*, **53**, 267–281.
- Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, **9**, 97–113.
- Ono, M., Kubik, S., & Abernathy, C. (1990). *Atlas of the cerebral sulci*. New York: Thieme Medical.
- Owen, A., Evans, A., & Petrides, M. (1996). Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: A positron emission tomography study. *Cerebral Cortex*, **6**, 31–38.
- Paus, T., Perry, D., Zatorre, R., Worsley, K., & Evans, A. (1996). Modulation of cerebral blood flow in the human auditory cortex during speech: Role of motor-to-sensory discharges. *European Journal of Neuroscience*, **8**, 2236–2246.
- Paulesu, E., Frith, C., & Frackowiak, R. (1993). The neural correlates of the verbal component of working memory. *Nature*, **362**, 342–345.
- Paulesu, E., Frith, U., Snowling, M., Gallagher, A., Morton, J., Frackowiak, R., & Frith, C. (1996). Is developmental dyslexia a disconnection syndrome? *Brain*, **119**, 143–157.
- Paulesu, E., Goldacre, B., Scifo, P., Cappa, S., Gilardi, M., Castiglioni, I., Perani, D., & Fazio, F. (1997). Functional heterogeneity of left inferior frontal cortex as revealed by fMRI. *NeuroReport*, **8**, 2011–2016.
- Perry, D., Zatorre, R., Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. (1999). Localization of cerebral activity during simple singing. *NeuroReport*, **10**, 3979–3984.
- Petrides, M. (1994). Frontal lobes and working memory: Evidence from investigations of the effects of cortical excisions in nonhuman primates. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 9, pp. 59–82). Amsterdam: Elsevier.
- Price, C., Wise, R., Ramsay, S., Friston, K., Howard, D., Patterson, K., & Frackowiak, R. (1992). Re-

- gional response differences within the human auditory cortex when listening to words. *Neuroscience Letters*, **146**, 179–182.
- Price, C., Indefrey, P., & van Turennout, M. (1999). The neural architecture underlying the processing of written and spoken word forms. In C. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 211–240). New York: Oxford Univ. Press.
- Rademacher, J., Galaburda, A., Kennedy, D., Filipek, P., & Caviness, V., Jr. (1992). Human cerebral cortex: Localization, parcellation, and morphometry with magnetic resonance imaging. *Journal of Cognitive Neuroscience*, **4**, 352–374.
- Raichle, M., Fiez, J., Videen, T., MacLeod, A., Pardo, J., Fox, P., & Petersen, S. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, **4**, 8–26.
- Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., & Grodd, W. (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula, and cerebellum. *NeuroReport*, **11**, 1–4.
- Robin, D., Tranel, D., & Damasio, H. (1990). Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain and Language*, **39**, 539–555.
- Shipley-Brown, F., Dingwall, W., Berlin, C., Yeni-Komshian, G., & Gordon-Salant, S. (1988). Hemispheric processing of affective and linguistic intonation contours in normal subjects. *Brain and Language*, **33**, 16–26.
- Schwartz, M., & Tallal, P. (1980) Rate of acoustic change may underlie hemispheric specialization for speech perception. *Science*, **207**, 1380–1381.
- Shulman, G., Fiez, J., Corbetta, M., Buckner, R., Miezin, F., Raichle, M., & Petersen, S. (1997). Common blood flow changes across visual tasks: II. decreases in cerebral cortex. *Journal of Cognitive Neuroscience*, **9**, 648–663.
- Sidtis, J. (1980). On the nature of the cortical function underlying right hemisphere auditory perception. *Neuropsychologia*, **18**, 321–330.
- Sidtis, J., & Volpe, B. (1988). Selective loss of complex-pitch or speech discrimination after unilateral lesion. *Brain and Language*, **34**, 235–245.
- Smith, E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, **283**, 1657–1661.
- Steinmetz, H., & Seitz, R. (1991). Functional anatomy of language processing: Neuroimaging and the problem of individual variability. *Neuropsychologia*, **12**, 1149–1161.
- Suen, C. (1986). *Computational studies of the most frequent Chinese words and sounds*. Philadelphia: World Scientific.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain, 3-D proportional system: An approach to cerebral imaging*. New York: Springer-Verlag.
- Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiology of speech perception. *Annals of New York Academy of Science*, **682**, 27–47.
- Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Papers in Linguistics*, **13**(2), 201–277.
- Van Lancker, D., & Fromkin, V. (1973). Hemispheric specialization for pitch and tone: Evidence from Thai. *Journal of Phonetics*, **1**, 101–109.
- Van Lancker, D., & Sidtis, J. (1992). The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: All errors are not created equal. *Journal of Speech and Hearing Research*, **35**, 963–970.
- Vijayan, A., & Gandour, J. (1995). On the notion of a “subtle phonetic deficit” in posterior aphasia. *Brain and Language*, **48**, 106–119.
- Winer, B., Brown, D., & Michels, K. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.
- Wioland, N., Rudolf, G., Metz-Lutz, M., Mutschler, V., & Marescaux, C. (1999). Cerebral correlates of hemispheric lateralization during a pitch discrimination task: An ERP study in dichotic situation. *Clinical Neurophysiology*, **110**, 516–523.
- Wise, R., Greene, J., Buchel, C., & Scott, S. (1999). Brain regions involved in articulation. *Lancet*, **353**, 1057–1061.
- Xu, Y. (1991). Depth of phonological recoding in short-term memory. *Memory & Cognition*, **19**, 263–273.
- Yiu, E., & Fok, A. (1995). Lexical tone disruption in Cantonese aphasic speakers. *Clinical Linguistics & Phonetics*, **9**, 79–92.

- Zatorre, R., & Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain*, **114**, 2403–2417.
- Zatorre, R., Evans, A., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch processing in speech perception. *Science*, **256**, 846–849.
- Zatorre, R., Evans, A., & Meyer, E. (1994). Cortical mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, **14**, 1908–1919.
- Zatorre, R., Meyer, E., Gjedde, A., & Evans, A. (1996). PET studies of phonetic processing of speech: Review, replication, and re-analysis. *Cerebral Cortex*, **6**, 21–30.