Published in final edited form as:

Hear Res. 2002 April ; 166(1-2): 9–23.

PET imaging of differential cortical activation by monaural speech and nonspeech stimuli

Donald Wong^{a,b,*}, David B. Pisoni^{b,c}, Jennifer Learn^d, Jack T. Gandour^e, Richard T. Miyamoto^b, and Gary D. Hutchins^f

^aDepartment of Anatomy and Cell Biology, Medical Science 5022, Indiana University School of Medicine, 635 Barnhill Drive, Indianapolis, IN 46202, USA

^bDepartment of Otolaryngology-Head Neck Surgery, Indiana University School of Medicine, Indianapolis, IN 46202, USA

^cDepartment of Psychology, Indiana University, Bloomington, IN 47405, USA

^dDepartment of Psychiatry (Medical Neurobiology Program), Indiana University School of Medicine, Indianapolis, IN 46202, USA

^eDepartment of Audiology and Speech Sciences, Purdue University, West Lafayette, IN 47907, USA

^fDepartment of Radiology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

Abstract

Positron emission tomography imaging was used to investigate the brain activation patterns of listeners presented monaurally (right ear) with speech and nonspeech stimuli. The major objectives were to identify regions involved with speech and nonspeech processing, and to develop a stimulus paradigm suitable for studies of cochlear-implant subjects. Scans were acquired under a silent condition and stimulus conditions that required listeners to press a response button to repeated words, sentences, time-reversed (TR) words, or TR sentences. Group-averaged data showed activated foci in the posterior superior temporal gyrus (STG) bilaterally and in or near the anterior insula/frontal operculum across all stimulus conditions compared to silence. The anterior STG was activated bilaterally for speech signals, but only on the right side for TR sentences. Only nonspeech conditions showed frontal-lobe activation in both the *left* inferior frontal gyrus [Brodmann's area (BA) 47] and ventromedial prefrontal areas (BA 10/11). An STG focus near the superior temporal sulcus was observed for sentence compared to word. The present findings show that both speech and nonspeech engaged a distributed network in temporal cortex for early acoustic and prelexical phonological analysis. Yet backward speech, though lacking semantic content, is perceived as speechlike by engaging prefrontal regions implicated in lexico-semantic processing.

Keywords

Speech perception; Positron emission tomography imaging; Reversed speech; Auditory cortical processing; Cochlear implant

1. Introduction

Functional neuroimaging techniques have provided a noninvasive tool for elucidating the neural circuits engaged in speech perception and spoken-language processing (Petersen et al., 1988). These studies of language processing have identified sites in the left inferior frontal cortex and posterior temporal cortex, regions classically implicated as speech/language centers from postmortem studies with aphasic patients (Geschwind, 1979). Furthermore, by comparing brain activation patterns under task conditions requiring different levels of signal processing and analysis, more extensive regions beyond the classical regions have also been identified for speech processing (Petersen and Fiez, 1993; Binder et al., 1997). Based on several recent imaging studies in speech processing, a widely distributed neural network has been hypothesized that links frontal and temporoparietal language regions (Brown and Hagoort, 1999).

The motivation of the present investigation was two-fold. First, we wanted to understand the cortical mechanisms underlying speech perception and spoken language processing. Both speech and nonspeech signals were used in order to identify cortical sites associated with complex-signal processing from sensory to lexicosemantic. Previous imaging studies have used time-reversed (TR) speech as a control condition to compare with speech signals (Howard et al., 1992; Hirano et al., 1997; Binder et al., 2000). TR speech preserves some of the acoustical properties of speech sounds, but these nonspeech signals are devoid of semantic content. Second, we hoped to develop methods with normal-hearing listeners that could be used in future positron emission tomography (PET)-imaging studies of patients with cochlear implants (CIs). Monaural stimulation was employed with these normal-hearing subjects to simulate the clinical conditions in which CI subjects hear on the side fitted with the prosthetic device (Naito et al., 1995, 1997; Wong et al., 1999). Another important consideration was selection of stimuli that have served as standard test batteries for assessing speech perception in hearing-impaired subjects. Lists of words and sentences were presented as test signals under different scanning conditions. Listeners were required to detect a consecutive repetition in a sequence of stimuli by pressing a response button. TR words and TR sentences served as the nonspeech control conditions. Monaural stimulation was used in all acoustic conditions in order to compare the findings with recently emerging imaging data from speech/language studies using binaural presentation.

2. Materials and methods

2.1. Subjects

Five right-handed subjects (three males, two females) with normal-hearing sensitivity (pure-tone air-conduction threshold ≤ 20 dB HL at 0.5, 1.0, 2.0, 4.0 kHz) and with a mean age of 28.3 ± 10.1 years (mean $\pm$ S.D.) participated in this study. Payment was given for their participation. All subjects provided written consent to the experimental protocols for this study, which was approved by The Institutional Review Board of IUPUI and Clarian (Study #9303-29) and was in accordance with the guidelines of the Declaration of Helsinki.

2.2. Stimuli and tasks

An audiotape cassette reproduced the test signals at approximately 75 dB SPL for a total of 3.5 min during scanning under the acoustic conditions. The acoustic signals were originally recorded onto a tape and played back as free-field stimuli delivered by a high-quality, commercial loudspeaker (Acoustic Research) approximately 18 inches from the right ear. The left ear was occluded using an E-A-R foam insert to attenuate sound transmission in this ear by at least 25–30 dB SPL. This monaural stimulation mimics the clinical condition found

with monaural listening by CI subjects hearing with a prosthesis implanted into only one ear. The ambient noise level in the PET room during scanning was ~65 dB SPL.

Prior to the scanning session, subjects were told that they would hear either speech or nonspeech signals in each of the acoustic conditions. However, they were not informed about the specific type of acoustic signal to be presented prior to each scan. Scanning conditions were either active or passive. In the passive, silent baseline condition, the subject was instructed to relax. There was no stimulus presentation or response required during this scanning condition. In the active task conditions, subjects grasped a response device with their right hand, and were instructed to press a button with their right thumb immediately after an acoustic stimulus (e.g., word or sentence) was consecutively repeated. This detection task was designed to direct the subjects' attention to the sound pattern and to monitor the stimulus sequence for repetitions. Thus, the speech and nonspeech tasks were conceptualized as simple auditory discrimination tasks. Each button-press response activated a red light, which the experimenter observed and used to score the number of correct responses during each active task condition. Subjects were debriefed after the imaging session to discuss relative task difficulty and their subjective impressions of the perceived stimuli.

Speech stimuli consisted of lists of isolated English words and meaningful sentences used for speech-intelligibility testing (Egan, 1948). The Word condition used a list of 54 phonetically balanced monosyllables (e.g., 'boost', 'fume', 'each', 'key', 'year', 'rope'). The Sentence condition used a list of 44 Harvard sentences (e.g., 'What joy there is in living.'; 'Those words were cue for the actor to leave.'; 'The wide road shimmered in the hot sun.') selected from the lists developed by Egan (1948). Each sentence comprised six to 10 words. TR versions of the same words and sentences were used for the two nonspeech conditions. The sound spectra for the speech stimuli were similar to their backward nonspeech versions (Fig. 1). These stimuli consisted of the 54 words played backwards in the TR Word condition, and 44 sentences played backwards in the TR Sentence condition. The percentage of consecutively repeated stimuli averaged 21% across all acoustic conditions. TR speech was considered to be devoid of semantic content, and was therefore considered an appropriate nonspeech control. The lists were presented at a rate of one stimulus per 4 s in the Word and TR Word conditions and at one stimulus per 5 s in the Sentence and TR Sentence conditions. The duration of each list was approximately 3.5 min long.

2.3. PET image acquisition and processing

PET scans were obtained using a Siemens 951/31R imaging system, which produced 31 brain image slices at an intrinsic spatial resolution of approximately 6.5 mm full-width-at-half-maximum (FWHM) in plane and 5.5 mm FWHM in the axial direction. During the entire imaging session, the subject lay supine 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 cerebral blood flow (CBF) across task conditions. A peripheral venipuncture and an intravenous infusion line were placed in the subject's left arm. For each condition, about 50 mCi of H₂¹⁵O was injected intravenously as a bolus; upon bolus injection, the scanner was switched on for 5 min to acquire a tomographic image. During the active acoustic conditions, sounds were played over a 3.5 min period followed by 1.5 min of silence. A rapid sequence of scans was performed to enable the selection of a 90-s time window, which begins 35–40 s after the bolus arrived in the brain. The data for image reconstruction are derived from this imaging interval after the 75% level of peak brain activity is determined posthoc. For each condition in the experimental design, instructions were given immediately prior to scanning. Repeated scans were acquired in subjects from the following stimulus conditions: (1) *Silent Baseline*, (2) *Word*, (3) *Sentence*, (4) *TR Word*, (5) *TR Sentence*.

Seven paired-image subtractions were then performed on group-averaged data to reveal statistically significant results in the difference images: (1) Word—Silence, (2) Sentence—Silence, (3) TR Word—Silence, (4) TR Sentence—Silence, (5) Sentence—Word, (6) Sentence—TR Sentence, and (7) Word—TR Word. The Sentence—Word subtraction was designed to dissociate processing of suprasegmental or prosodic cues at the sentence level from those at the level of isolated words. In the Sentence—TR Sentence condition, nonspeech is subtracted from speech. Regions of significant brain activation were identified by performing an analysis process (Michigan software package, Minoshima et al., 1993) that included image registration, global normalization of the image volume data, identification of the intercommissural (anterior commissure—posterior commissure) line on an intrasubject-averaged PET image set for stereotactic transformation and alignment, averaging of subtraction images across subjects, and statistical testing of brain regions demonstrating significant regional CBF changes. Changes in regional CBF were then mapped onto a standardized coordinate system of Talairach and Tournoux (1988). Foci of significant CBF changes were tested by the Hammersmith method (Friston et al., 1990, 1991) and values of $P < 0.05$ (onetailed, corrected) were identified as statistically significant. The statistical map of blood flow changes was then overlaid onto a high-resolution T1-weighted, structural MRI of a single subject solely for display purposes to facilitate identification of activated and deactivated regions with respect to major gyral and sulcal landmarks under each of the subtractions. Each focus of activity was attributed to a specific gyrus and Brodmann's area (BA) on the basis of mapping the peak focus of multi-subject data onto the designated regions of the Talairach atlas.

The significant peak activation localized to the inferior frontal gyrus (IFG), pars orbitalis was based on mapping the focus onto the part of this gyrus designated as BA 47 in the Talairach atlas. The multiple foci of significant peak activation in the superior temporal gyrus were distinguished by arbitrarily grouping these foci into anterior ($y = -5$ mm), middle (y from -5 to -23 mm), and posterior (y from -24 to -35 mm) (Wong et al., 1999). The extent of activation was determined only in the superior temporal gyrus (STG) of each hemisphere by drawing regions of interest (ROIs) around the activation foci at the Hammersmith threshold. A single ROI was drawn on each side to include the extent of activation from all peak foci of STG activation.

3. Results

3.1. Behavioral performance

The subjects scored 100% on the detection tasks for the Word and Sentence conditions. On the nonspeech tasks, a total of one error was scored in the TR Word condition, and two errors in the TR Sentence condition for all subjects.

3.2. Foci of significant blood flow increases

Compared to the silent baseline condition, in both the Sentence and Word conditions, extensive CBF increases were observed bilaterally in the STG (Table 1; Fig. 2, upper two panels). The STG activation pattern was generally more robust and larger on the *left* side for all baseline subtractions in this study using right-ear stimulation (see Table 4). The activated region was elongated in an anterior-to-posterior direction with multiple peak foci distinguishable in the anterior, middle, and posterior parts of the STG (Table 1: foci 4–9, 14–20; Fig. 2). The activations in the posterior half of the STG were often in the superior temporal plane within the Sylvian fissure, presumably encompassing the primary and secondary association auditory cortex (BA 41/42). This activation pattern also extended ventrally onto the lateral (exposed) STG surface as far as the superior temporal sulcus (STS) and middle temporal gyrus (MTG), especially on the left side. This robust activation

presumably included a part of BA 22 on the lateral surface and a part of BA 21 in the banks of the STS or on the MTG. The anterior STG activation was typically observed on the lateral surface, near the STS, and toward BA 38, a region containing the temporal pole. CBF increases were consistently found at the junction between the anterior insula and the frontal operculum (FO) on the left side (Table 1: foci 10–11, 21–22; Fig. 2) (bilateral for Sentence). No CBF increases were found in the IFG of the *left* frontal cortex

Compared to the silent baseline condition, both the TR Sentence and TR Word conditions showed CBF increases in the temporal lobe bilaterally (Table 2; Fig. 2). Compared to baseline, the TR Sentence showed a robust bilateral STG activation pattern (Table 2: foci 13–16) that was similar to that observed for the Sentence minus baseline condition. The strong left posterior STG activation also extended ventrally as far as the STS/MTG (Table 2: focus 13; Fig. 2); this spread of activity was similar to that found for the speech conditions compared to baseline. Noteworthy is the pattern of STG activation on the right side, which contains multiple anterior and posterior foci (Table 2: foci 14–16); the foci extended along the lateral STG surface, but did not spread to the STS. The TR Word condition compared to baseline showed a focal activation confined to the posterior STG activation (mainly left; Table 2: foci 4–6; Fig. 2). The TR Sentence compared to baseline condition showed an elongated swath of STG activation (Fig. 2).

The temporal-lobe activation was mainly on the left side in the posterior STG and MTG. Only a single focus was observed in the posterior STG on the right side. The activated focus observed for TR Sentence minus baseline condition was an elongated swath of activity on the left lateral STG surface along the anterior-to-posterior direction similar to that found in the Sentence minus baseline condition. In contrast, the left STG focus for TR Word minus baseline was more focally confined to the posterior STG, extending ventrally rather than anteriorly (Fig. 2). Examination of the activation patterns of all four baseline comparisons revealed that both the Sentence and TR Sentence conditions evoked larger activations than the Word and TR Word conditions. These larger STG activations occurred extensively along the anterior-to-posterior direction, whereas the smaller activations were confined only to the posterior STG.

Compared to the baseline condition, the two non-speech conditions showed CBF increases in foci of the frontal lobes that were not observed in the speech conditions (Table 2). For example, activation foci were found in the *left* IFG (pars orbitalis, BA 47) (Table 2: foci 1–2, 8–9; Fig. 3), a region often referred to as ventral or inferior prefrontal cortex in imaging studies on language processing (see Fiez, 1997). A second pattern of frontal-lobe activation was also found bilaterally in a part of the frontopolar region; this activation was largely confined to the ventromedial frontal cortex in BA 10/11 (Fig. 3).

CBF increases were also isolated when the speech and nonspeech conditions were compared (Table 3). In the Sentence minus Word condition, a focus of CBF increase was found in the middle part of the left STG (BA 22) near the STS (Table 3: focus 1; Fig. 4). The TR Sentence minus TR Word condition was the only other comparison between two active tasks to show CBF increases in the STG (Table 3: foci 8–10). No CBF increases were found in any speech region of the left frontal lobe for comparisons between speech conditions (Sentence—Word) or between speech and nonspeech conditions (Sentence minus TR Sentence). No significant CBF increase was found for Word—TR Word.

Table 5 summarizes the major similarities and differences in the patterns of CBF increases for speech and nonspeech conditions relative to silent baseline. In brief, activation patterns in posterior STG bilaterally and anterior insula/FO were found consistently across all speech and nonspeech conditions. Bilateral activation in anterior STG was observed for the speech

conditions only. Activation in both the ventromedial prefrontal cortex and left IFG was found only in the nonspeech conditions. When different levels of complex-sound processing were compared, a focus in the left STG/MTG in the STS was observed for the Sentence condition compared to the Word condition.

3.3. Foci of significant blood flow decreases

Compared to the baseline condition, all speech and nonspeech conditions showed CBF decreases typically in the medial parietal and occipital lobe in such regions as the precuneus (BA 7) and posterior cingulate (BA 31, 30, 23) (Table 6: foci 1, 6–8, 10–12). The majority of these foci were found for the TR Sentence minus baseline condition (Fig. 5).

4. Discussion

4.1. Activation in temporal lobe

In neuroimaging studies with normal-hearing listeners, both binaural (e.g., Petersen et al., 1988) and monaural studies (e.g., Lauter et al., 1985; Scheffler et al., 1998) have demonstrated bilateral STG activation. In the monaural studies, a more extensive activation was observed in the hemisphere contralateral to the stimulated ear. The larger temporal-cortex activation contralateral to stimulated ear reflects the known fact that auditory inputs transmitted along the central auditory pathway are sent predominantly to the contralateral auditory cortex (primary and secondary association areas) (e.g., see Jones and Peters, 1985). The present monaural study also demonstrated this general pattern of bilateral STG activation. When speech (words, sentences) and nonspeech (TR words, TR sentences) conditions were compared to the silent baseline condition, the posterior STG was activated bilaterally, with a stronger and more extensive focus in the left hemisphere to right-ear stimulation.

The sound patterns for both the forward and backward sentences evoked the largest overall level of STG activation. The isolated words evoked a smaller level of activation. The TR words evoked the lowest level of activation (Table 4). This differing extent of activation may reflect the greater complexity of the sound patterns (Sentence and TR Sentence) versus the isolated stimuli (Word and TR Word). This interpretation is consistent with binaural studies, in which the activated auditory regions in the temporal lobe became more extensive bilaterally as the task demands increased with complexity of speech stimuli (Just et al., 1996).

The stimulus delivery could also have contributed to the larger overall STG activation for the Sentence (and TR version) than Word (and TR version) condition. For example, presentation rate has been reported to affect the STG activation size (Price et al., 1992; Binder et al., 1994). The rate of word presentation was slower in the Word than Sentence condition (isolated words versus a sequence of words within a sentence during a unit of time). This effect could account for the smaller activation foci for the forward and backward Word conditions. However, both these conditions (compared to baseline) showed a more confined activation in the posterior STG that was as robust on the left as the more extensive STG activation in the forward and backward Sentence conditions. The fact that we find common activation in the posterior STG for both Sentence and Word conditions cannot be explained simply by a presentation-rate effect. Instead, this posterior STG focus appears to reflect processing demands associated with speech stimuli (Price et al., 1992). Second, differences in stimulus duration between conditions could also contribute to differing size of activated foci. The total duration of the stimuli during scanning (i.e., total time that stimuli are heard) was somewhat larger for the forward and backward sentence conditions.

In all stimulus conditions compared to silent baseline, the left posterior STG was consistently activated. This focus extended from within the Sylvian fissure to the lateral cerebral surface as far ventrally as the STS and MTG. In a monaural study, Hirano et al. (1997) also observed activated foci in both the STG and the MTG when sentences were compared to backward sentences. In a binaural study, Wise et al. (1991) noted that the bilateral STG activation showed a similar pattern when speech and backward speech were each compared to silent baseline. This pattern of activation suggests that the posterior temporal gyrus of both sides participates in the initial cortical stages of sensory analysis of complex sounds, whether isolated words or sentences perceived as speech or nonspeech. The fact that both speech and nonspeech stimuli similarly activated this region supports the view that listeners can still perceive speech features from nonspeech signals as unintelligible as backward speech (Binder et al., 2000). Furthermore, the left posterior STG, especially near the STS and posterior MTG, has been hypothesized to be involved with prelexical phonological processing (Mazoyer et al., 1993; Price et al., 1999). Our findings suggest that both speech and nonspeech were processed cortically beyond the early sensory level to at least the prelexical phonological stages in left-lateralized speech-specific sites.

The anterior STG (BA 22) was activated bilaterally only when the speech conditions were compared to the silent baseline condition. This pattern of activity was typically located ventrally near the STS/MTG (BA 22/21) and as far anteriorly as BA 38 in a region of the temporal pole (Rademacher et al., 1992). The view that the anterior STG/temporal pole of both sides involves speech-specific processing was previously proposed since bilateral activation of the anterior STG was consistently found when subjects listened to speech sounds (Zatorre et al., 1992; Mazoyer et al., 1993; Binder et al., 1994). In the present study using monaural stimulation, the anterior STG was activated on the right side only when the nonspeech condition (backward sentence) was compared to silence. This finding, in conjunction with the bilateral activation observed in the anterior STG for the speech conditions, suggests that the left anterior STG is engaged in speech-specific processing. The ipsilateral activation of the right anterior STG found in the backward sentence condition strongly suggests that this acoustic stream engaged right-hemispheric mechanisms specialized for the encoding and storage of prosodic and intonation cues (Zatorre et al., 1992; Griffiths et al., 1999). The absence of right-sided activity for the backward word condition compared to baseline is unclear, although this may be due simply to the relatively lower overall level of STG activation observed for this task. The bilateral activation under the speech conditions is also consistent with the interpretation that prosodic processing of the speech stimuli is also lateralized to this right homologous region (Zatorre et al., 1992).

4.2. Frontal-lobe activation

Two separate frontal activations were found almost exclusively in the nonspeech minus silent condition. One activated focus was located in the left IFG, pars orbitalis (BA 47). This focus is over 20 mm rostral to the activation in the frontal operculum/anterior insula, and is at least 18 mm from the second frontal focus located in the ventromedial prefrontal cortex (BA 11/10).

Neuroimaging studies over the last decade have demonstrated that the left inferior prefrontal cortex (BA 44, 45, 46, 47) shows the strongest activation to semantic task demands (for review see Price et al., 1999). The finding that the left IFG was activated only under the *nonspeech* conditions compared to silent baseline was somewhat unexpected. Thus, one would have expected the speech conditions to activate the part of the left IFG that contains the classically defined Broca's area (pars opercularis, pars triangularis: BA 44/45) and is associated mainly with phonological processing. Yet, imaging data obtained across word generation and semantic decision tasks implicate the left IFG, pars orbitalis in semantic processing (Petersen et al., 1988; Binder et al., 1997). This left prefrontal activation has also

been found to extend into the temporal-parietal cortex, from the temporal pole to the angular gyrus (Vanderberghe et al., 1996). Thus, the semantic processing system for spoken language may be mediated by a distributed neural-network linking frontal regions with temporo-parietal language centers on the left hemisphere. The IFG, pars orbitalis, has an executive role for maintaining and controlling the effortful retrieval of semantic information from the posterior temporal areas (Fiez, 1997; Price et al., 1999).

The lack of a left prefrontal activation to speech compared to baseline may be explained by the fact that no explicit semantic association was required for the task in the present study. The subjects treated the task as a simple detection task. Despite the available lexico-semantic information contained in the familiar speech sounds, the speech tasks were apparently performed with minimal effort at the earlier acoustic and prelexical levels of analysis in the brain. This interpretation would account for the observed activation pattern in the superior temporal region that did not extend into the left inferior prefrontal cortex. In fact, similarity of STG activation patterns for forward and backward speech is consistent with the hypothesis that temporal regions are largely used for acoustic-phonetic processing (Binder et al., 2000). The left prefrontal activations observed for the nonspeech signals may be related to greater task demands for backward speech that involved an effortful, although unsuccessful attempt at semantic retrieval and interpretation of these sound patterns. A greater task demand may also be placed on auditory working memory to maintain representations of these nonlinguistic stimuli. BAs 45/47 have been hypothesized as a site for the attentional control of working memory (Smith and Jonides, 1999). The greater effort for the backward speech tasks as subjectively reported during post-scan debriefing may indicate increased attentional or task demands in these auditory discrimination tasks, not otherwise reflected in the scores of speech and nonspeech tasks. Different processes may also be engaged for the nonspeech tasks that resulted in activation of the left prefrontal cortex. The present study cannot dissociate the multiple cortical subsystems, whether task- or speech-specific, that are engaged when listening to backward speech.

The other major prefrontal activation was observed ventromedially. The further recruitment of this prefrontal focus may be related to the generally greater task demands associated with listening to unfamiliar auditory signals. This focus was near frontopolar regions activated by verbal memory tasks involving auditory monitoring (Petrides et al., 1993), or tasks related to retrieval effort and search in memory (Buckner et al., 1996). In the present study, the backward sentence task, presumably the more demanding of the two nonspeech tasks, evoked multiple activated foci extending bilaterally in these areas.

4.3. Activation in anterior insula

The anterior insula/FO was activated on the left side for both speech and nonspeech conditions compared to silent baseline. Bilateral activation was found only for the conditions that required listening to stimulus patterns (Sentence or TR Sentence). The auditory task in detecting signal repetition requires maintenance of stimulus patterns in short-term working memory, a function attributed to this region (Paulesu et al., 1993). The possible subvocal rehearsal of meaningful speech signals or pitch patterns in nonspeech may also underlie this activation (Zatorre et al., 1994).

4.4. Activation dissociated from speech and nonspeech comparison

Previous neuroimaging studies have attempted to dissociate sites implicated in prelexical and lexico-semantic stages of cortical processing by directly comparing speech and nonspeech conditions. However, when using backward speech as a nonspeech control for forward speech or pseudowords as a control for real words (Wise et al., 1991; Price et al., 1996; Binder et al., 2000), the brain activation patterns among these subtractions showed

little if any differences, especially in left-hemispheric regions associated with semantic processing (e.g., prefrontal cortex, angular gyrus and ventral temporal lobe). These negative results suggest that these 'speechlike' stimuli, even though they are devoid of semantic content, unavoidably accessed stages of processing up to possibly the lexical level, but produced less activation in this network overall than real words (Norris and Wise, 2000). Consequently, commonly activated foci would be subtracted out in speech versus nonspeech contrasts. The present study also did not isolate auditory/speech centers of significant activation when speech was compared to backward speech (Word minus TR Word; Sentence minus TR Sentence). These findings are consistent with the proposal that backward speech is a complex signal that is heard and processed as speech stimuli (Cutting, 1974), and thus will attempt to engage the distributed network for spoken language as much as possible. Backward speech engaged not only most of the temporal-lobe network that mediates auditory and prelexical phonological stages of analysis of spoken language, but also additional stages of lexico-semantic processing associated with the left frontal lobe.

The TR Sentence compared to TR Word condition revealed activated foci in the left STG, right anterior STG, and basal prefrontal cortex (BA 11). These activated foci may be simply related to the greater complexity and higher presentation rate, and hence greater potency in activation, of a sound pattern associated with a stream (sentences or TR version) than with isolated stimuli (words or TR version). Whereas the activation on the left side may merely reflect a greater activation contralateral to the monaural stimulus, the activation on the right side (anterior STG) probably reflects the relatively greater pitch processing associated with the stimulus stream.

The Sentence compared to Word condition dissociated a focus in the left STS where Mazoyer et al. (1993) implicated in sentence-level processing. In their binaural study, the STG activation became significantly more asymmetric (left-sided) to meaningful stories than to word lists, and the left MTG was activated by stories but not by word lists. Our observations also would support the hypothesis that the cortical stages of processing at the single-word level and higher involve more extensive areas in the temporal lobe outside the classically defined Wernicke's area in the temporo-parietal regions (Binder et al., 1997).

4.5. Deactivation of cortical regions

CBF decreases were commonly found in the medial parietal/occipital lobe for all silent baseline subtractions. Deactivation in these cortical regions had been attributed to suspension of ongoing processes occurring in the uncontrolled, silent condition (Shulman et al., 1997; Binder et al., 1999). The multiple deactivated foci observed in the TR Sentence minus baseline condition may also be related to the greater effort and attentional demands in task performance under this nonspeech condition.

4.6. Implications for neuroimaging of CI patients

Neuroimaging studies of speech and language processing in normal-hearing subjects have recognized that task performance can involve not only the intended auditory processing from early sensory analysis to linguistic processing, but other nonspecific cognitive task-demands that are automatically engaged, such as selective attention and working memory. Yet, no imaging study with CI subjects has considered these more general cognitive demands as they relate to outcomes in speech-perception tasks. Thus, future imaging studies of CI users that attempt to relate their speech-perception levels to the distributed neural network activated in task performance should consider other networks engaged along with those for speech processing. In a recent PET study of a new CI user (Miyamoto et al., 2000), speech stimuli activated prefrontal areas near some of those activated by backward speech in the present study. CI users presumably encounter greater demands on attention and working

memory when listening to speech as compared to normal listeners. Thus, the effortful attempt of CI users to make sense of speech may be modeled in part by observing normal-listeners' efforts to make sense of *backward speech*. These cognitive demands may initially be quite substantial as CI users attempt to recognize degraded signals fed through the device as speech. It remains to be determined whether these frontal circuits will further develop and influence the speech-perception strategies and outcomes of CI users.

Acknowledgments

This study was supported by the Departments of Otolaryngology, Radiology, NIH Grants DC 00064 (R.T.M.) and 000111 (D.B.P.). We thank Rich Fain and PET-facility staff for assistance and radionuclide production, and developers of Michigan PET software for its use.

Abbreviations

BA	Brodmann's area
CBF	cerebral blood flow
CI	cochlear implant
FO	frontal operculum
IFG	inferior frontal gyrus
MTG	middle temporal gyrus
PET	positron emission tomography
ROI	regional of interest
STG	superior temporal gyrus
STS	superior temporal sulcus
TR	time-reversed

References

- Binder J, Rao S, Hammeke T, Frost J, Bandettini P, Hyde J. Effects of stimulus rate on signal response during functional magnetic resonance imaging of auditory cortex. *Cogn. Brain Res.* 1994; 2:31–38.
- Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T. Human brain language areas identified by functional magnetic resonance imaging. *J. Neurosci.* 1997; 17:353–362. [PubMed: 8987760]
- Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Rao SM, Cox RW. Conceptual processing during the conscious rest state: A functional MRI study. *J. Cogn. Neurosci.* 1999; 11:80–93. [PubMed: 9950716]
- Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Springer JA, Kaufman JN, Possing T. Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex.* 2000; 10:512–528. [PubMed: 10847601]
- Brown, CM.; Hagoort, P., editors. *The Neurocognition of Language*. Oxford University Press; New York: 1999.
- Buckner RL, Raichle ME, Miezin FM, Petersen SE. Functional anatomic studies of memory retrieval for auditory words and visual pictures. *J. Neurosci.* 1996; 16:6219–6235. [PubMed: 8815903]
- Cutting JE. Two left-hemisphere mechanisms in speech perception. *Percept. Psychophys.* 1974; 16:601–612.
- Egan JP. Articulation testing methods. *Laryngoscope.* 1948; 58:955–991. [PubMed: 18887435]
- Fiez JA. Phonology, semantics, and the role of the left inferior prefrontal cortex. *Hum. Brain Mapping.* 1997; 5:79–83.

- Friston K, Frith C, Liddle P, Dolan R, Lamerstma A, Frackowiak R. The relationship between global and local changes in PET scans. *J. Cereb. Blood Flow Metab.* 1990; 10:458–466. [PubMed: 2347879]
- Friston K, Frith C, Liddle P, Frackowiak R. Comparing functional PET images: The assessment of significant changes. *J. Cereb. Blood Flow Metab.* 1991; 11:81–95.
- Geschwind N. Specializations of the human brain. *Sci. Am.* 1979; 241:158–168.
- Griffiths TD, Johnsrude I, Dean JL, Green GGR. A common neural substrate for the analysis of pitch and duration patterns in segmented sounds. *NeuroReport.* 1999; 10:3825–3830. [PubMed: 10716217]
- Hirano S, Naito Y, Okazawa H, Kojima H, Honjo I, Ishizu K, Yenokura Y, Nagahama Y, Fukuyama H, Konishi J. Cortical activation by monaural speech sound stimulation demonstrated by positron emission tomography. *Exp. Brain Res.* 1997; 113:75–80. [PubMed: 9028776]
- Howard D, Patterson K, Wise R, Brown WD, Friston K, Weiller C, Frackowiak R. The cortical localization of the lexicon. *Brain.* 1992; 115:1769–1782. [PubMed: 1486460]
- Jones, EG.; Peters, A. Association and Auditory Cortices. Vol. 4. Plenum Press; New York: 1985. Cerebral Cortex.
- Just MA, Carpenter PA, Keller TA, Eddy WF, Thulborn KR. Brain activation modulated by sentence comprehension. *Science.* 1996; 274:114–116. [PubMed: 8810246]
- Lauter JL, Herscovitch P, Formby C, Raichle ME. Tonotopic organization in human auditory cortex revealed by positron emission tomography. *Hear. Res.* 1985; 20:199–205. [PubMed: 3878839]
- Mazoyer B, Dehaene S, Tzourio N, Frak V, Cohen L, Murayama N, Levrier O, Salamon G, Mehler L. *J. Cogn. Neurosci.* 1993; 5:467–479.
- Minoshima S, Koeppe R, Mintum M, Berger KL, Taylor SF, Frey KA, Kuhl DE. Automated detection of the inter-commissural line for stereotaxic localization of functional brain imaging. *J. Nucl. Med.* 1993; 34:322–329. [PubMed: 8429356]
- Miyamoto, RT.; Wong, D.; Pisoni, DB.; Hutchins, GD. Changes induced in brain activation in a prelingually-deaf, adult cochlear implant user: A PET study. Presented at Annual Convention of ASHA; Nov 16–19. 2000
- Naito Y, Okazawa H, Honjo I, Hirano S, Takahashi H, Shiomi Y, Hoji W, Kawano M, Ishizu K, Yonekura Y. Cortical activation with sound stimulation in cochlear implant users demonstrated by positron emission tomography. *Cogn. Brain Res.* 1995; 2:207–214.
- Naito Y, Okazawa H, Hirano S, Takahashi H, Kawano M, Ishizu K, Yonekura Y, Konishi J, Honjo I. Sound induced activation of auditory cortices in cochlear implant users with post- and prelingual deafness demonstrated by positron emission tomography. *Acta Oto-Laryngol.* 1997; 117:490–496.
- Norris, D.; Wise, R. The study of prelexical and lexical processes in comprehension: Psycholinguistics and functional neuro-imaging. In: Gazzaniga, M., editor. *The New Cognitive Neurosciences*. 2nd edn. MIT Press; Cambridge, MA: 2000. p. 867–880.
- Paulesu E, Frith C, Frackowiak R. The neural correlates of the verbal component of working memory. *Nature.* 1993; 362:342–345. [PubMed: 8455719]
- Petersen SE, Fiez JA. The processing of single words studied with positron emission tomography. *Annu. Rev. Neurosci.* 1993; 13:25–42. [PubMed: 2183676]
- Petersen SE, Fox PT, Posner MI, Mintum M, Raichle ME. Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature.* 1988; 331:585–589. [PubMed: 3277066]
- Petrides M, Alivasatos B, Meyer E, Evans AC. Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proc. Natl. Acad. Sci. USA.* 1993; 90:878–882. [PubMed: 8430101]
- Price C, Wise R, Ramsay S, Friston K, Howard D, Patterson K, Frackowiak R. Regional response differences within the human auditory cortex when listening to words. *Neurosci. Lett.* 1992; 146:179–182. [PubMed: 1491785]
- Price CJ, Wise RJS, Warburton EA, Moore CJ, Howard D, Patterson K, Frackowiak RSJ, Friston KJ. Hearing and saying. The functional neuroanatomy of auditory word processing. *Brain.* 1996; 119:919–931. [PubMed: 8673502]

- Price, C.; Indefrey, P.; Turrenoul, M. The neural architecture underlying the processing of written and spoken word forms. In: Brown, CM.; Hagoort, P., editors. *The Neurocognition of Language*. Oxford University Press; New York: 1999. p. 211-240.
- Rademacher J, Galaburda AM, Kennedy DN, Filipek PA, Caviness VS. Human cerebral cortex: localization, parcellation, and morphometry with magnetic resonance imaging. *J. Cogn. Neurosci.* 1992; 4:352–374.
- Scheffler K, Bilecen D, Schmid N, Tschopp K, Seelig J. Auditory cortical responses in hearing subjects and unilateral deaf patients as detected by functional magnetic resonance imaging. *Cereb. Cortex.* 1998; 8:156–163. [PubMed: 9542894]
- Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Petersen SE. Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *J. Cogn. Neurosci.* 1997; 9:648–663.
- Smith E, Jonides J. Storage and executive processes in the frontal lobes. *Science.* 1999; 283:1657–1661. [PubMed: 10073923]
- Talairach, J.; Tournoux, P. 3-Dimensional Proportional System: An Approach to Cerebral Imaging. Thieme Medical; New York: 1988. Co-planar Stereotaxic Atlas of the Human Brain.
- Vanderberghe R, Price CJ, Wise R, Josephs O, Frackowiak RSJ. Functional anatomy of a common semantic system for words and pictures. *Nature.* 1996; 383:254–256. [PubMed: 8805700]
- Wise R, Chollet F, Hadar U, Friston KJ, Hoffner E, Frackowiak RSJ. Distribution of cortical networks involved in word comprehension and word retrieval. *Brain.* 1991; 114:1803–1817. [PubMed: 1884179]
- Wong D, Miyamoto RT, Pisoni DB, Sehgal M, Hutchins GD. PET imaging of cochlear-implant and normal-hearing subjects listening to speech and nonspeech. *Hear. Res.* 1999; 132:34–42. [PubMed: 10392545]
- Zatorre RJ, Evan AC, Meyer E, Gjedde A. Lateralization of phonetic and pitch discrimination in speech processing. *Science.* 1992; 256:846–849. [PubMed: 1589767]
- Zatorre RJ, Evans AC, Meyer E. Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 1994; 14:1908–1919. [PubMed: 8158246]

Fig. 1. Spectra of a sample word (A), sentence (C), and their backward versions (B, D). The acoustic nature of the speech stimulus was similar to its backward versions.

Fig. 2. Group-averaged PET images showing activated foci for the speech and nonspeech conditions minus silent baseline. In all subtractions activation is observed in the STG bilaterally with a more extensive activation on the left side. The STG activation is more extensive for the Sentence and TR Sentence conditions than the Word and TR Word conditions. Activation is also found at the junction of the anterior insula (INS) and FO. The TR Sentence minus Baseline condition shows an activated focus in the left IFG (bottom row; see Fig. 3). Each row of images contains an axial section flanked on each side by a sagittal section. The axial and left sagittal images are cut through the focus of peak activation in the left STG. The right sagittal image is cut through the focus of peak activation in the right STG. The left (L, x = negative) and right (R, x = positive) sides of the brain are designated in all images. PET images are overlaid onto a high-resolution, T1-weighted anatomic reference image in all figures. Monaural stimuli were presented into the *right* ear.

Fig. 3. Group-averaged PET images showing prefrontal activation for the nonspeech conditions minus silent baseline. Activation of the IFG, pars orbitalis, is observed only on the *left* side. Other prefrontal activations include the orbital gyrus on both sides. Each subtraction shows two coronal images (y-coordinate with larger value is more anterior).

Fig. 4. Group-averaged PET images showing activation for the Sentence minus Word condition. Activation of the temporal lobe is observed only on the left side in the STS at the junction between the midportion of the STG and MTG. The sagittal (top left), axial (top right), coronal (bottom right) images are cut through the focus of peak activation.

Fig. 5.

Group-averaged PET images showing deactivated foci for the TR Sentence condition minus silent baseline. Deactivated foci are observed mainly in the precuneus and posterior cingulate of the medial parietal lobe. This subtraction shows the most extensive pattern of deactivation. The sagittal (top left), axial (top right), coronal (bottom right) images are cut through a point near several deactivated foci.

Table 1

Speech task compared to silent baseline: regions of significant blood flow increases^a

Region	Brodmann's area	Coordinates (mm)			Z score
		x	y	z	
Word—baseline					
Frontal lobe					
1. R inferior frontal gyrus	47	48	30	-7	4.8
2. L orbital gyrus/gyrus rectus	11	-10	12	-18	4.9
3. L middle frontal gyrus	11	-19	19	-16	4.8
Temporal lobe					
4. L anterior superior temporal gyrus/STS	22/21	-46	-4	-7	4.8
5. L posterior superior temporal gyrus/STS	22	-55	-28	4	6.6
6. R anterior superior/middle temporal gyrus/STS	22/21/38	48	5	-11	4.3
7. R anterior superior temporal gyrus/STS	22/21	57	-4	-2	4.9
8. R mid superior temporal gyrus	22	53	-15	2	4.5
9. R post superior temporal gyrus/STS	22	57	-28	4	5.9
Other					
10. L anterior insula/frontal operculum	–	-35	19	0	5.2
11. L anterior insula	–	-33	5	-11	4.5
12. L posterior insula	–	-33	-31	9	5.4
13. R thalamus (dorsomedial nucleus)	–	1	-17	2	4.6
Sentence—baseline					
Temporal lobe					
14. L anterior superior temporal gyrus/STS	22/21	-48	3	-9	6
15. L transverse gyrus of Heschl	42/41	-33	-31	14	5.5
16. L posterior superior temporal gyrus/STS	22	-53	-28	4	8.4
17. R anterior superior temporal gyrus	38	46	10	-9	5
18. R mid superior temporal gyrus	22/42	53	-15	2	6.1
19. R mid superior temporal gyrus	22	55	-19	4	6.1
20. R posterior superior temporal gyrus/STS	22/21	55	-28	2	6.1
Other					

Region	Brodmann's area	Coordinates (mm)			Z score
		x	y	z	
21. L anterior insula/frontal operculum	–	–35	21	2	5.4
22. R anterior insula/frontal operculum	–	39	14	0	5.1

^aSignificant activation foci that exceed the Hammersmith statistical criterion of significance (adjusted P threshold = 0.05) in normalized CBF for all subtractions. 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); y -coordinate refers to anterior–posterior position relative to the anterior commissure (positive = anterior); z -coordinate refers to superior–inferior position relative to the CA–CP (anterior commissure–posterior commissure) line (positive = superior). Regions and Brodmann's areas listed are based on mapping the peak foci onto these designated areas in the atlas. L = left; R = right.

Table 2

Nonspeech task compared to silent baseline: regions of significant blood flow *increases*^a

Region	Brodmann's area	Coordinates (mm)			Z score
		X	y	z	
Time-reversed Word—baseline					
Frontal lobe					
1. L. Inferior frontal gyrus, pars orbitalis	47	-35	28	-2	4.4
2. L. Inferior frontal gyrus, pars orbitalis	47	-35	39	-7	4.5
3. R. orbital gyrus	11	10	46	-18	4.4
Temporal lobe					
4. L. posterior superior temporal gyrus	22/42	-51	-31	7	6.2
5. L. posterior middle temporal gyrus	21/20	-51	-40	-9	4.2
6. R. posterior superior temporal gyrus	42	57	-26	7	4.2
Other					
7. L. insula	-	-371	1	-165	5
Time-reversed Sentence—baseline					
Frontal lobe					
8. L. inferior frontal gyrus, pars orbitalis	47	-30	41	-9	5.2
9. L. inferior frontal gyrus (pars orbitalis)/middle frontal gyrus	47/11/10	-39	41	2	4.8
10. L. orbital gyrus	11	-12	39	-18	5.1
11. R. middle frontal gyrus	11	26	48	-11	4.5
12. R. middle frontal gyrus	11/10	37	48	-4	4.5
Temporal lobe					
13. L. posterior superior temporal gyrus	22/42	-53	-22	2	10
14. R. anterior superior temporal gyrus	38	48	8	-9	5.3
15. R. anterior superior temporal gyrus	22	57	-6	-2	6.8
16. R. posterior superior temporal gyrus	22/42	57	-26	4	7.1
Other					
17. L. anterior insula/frontal operculum	-	-35	23	0	5.1
18. R. anterior insula/frontal operculum	-	39	19	2	5.7
19. R. inferior parietal lobule	40	44	-49	45	4.3

^aSignificant activation foci that exceed the Hammersmith statistical criterion of significance (adjusted P -threshold = 0.05) in normalized CBF for all subtractions. 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); y -coordinate refers to anterior–posterior position relative to the anterior commissure (positive = anterior); z -coordinate refers to superior–inferior position relative to the CA–CP (anterior commissure–posterior commissure) line (positive = superior). Regions and Brodmann's areas listed are based on mapping the peak foci onto these designated areas in the atlas. L = left; R = right.

Table 3

Comparison of speech and nonspeech tasks: regions of significant blood flow *increases*^a

Region	Brodmann's area	Coordinates (mm)			Z score
		X	y	z	
Sentence—Word					
Temporal lobe					
1. L mid superior temporal gyrus (in STS)	22	-53	-13	-2	4.3
2. L post thalamus (pulvinar)	-	-12	-35	7	4.5
Other					
3. R frontal operculum/anterior insula	-	35	-1	16	4.2
Sentence—TR Sentence					
Parietalooccipital lobe					
4. L post cingulate	23	-3	-55	18	4.4
5. L post cingulate	23/31	-3	-46	27	4.4
6. L precuneus	31	-6	-62	20	4.3
TR Sentence—TR Word					
Frontal lobe					
7. L orbital gyrus	11	-15	41	-18	4.1
Temporal lobe					
8. L mid superior temporal gyrus	22/42	-55	-17	2	6.5
9. L transverse gyrus of Heschl	41	-37	-31	7	4.3
10. R anterior superior temporal gyrus	22/38	53	-1	-7	4.1

^aSignificant activation foci that exceed the Hammersmith statistical criterion of significance (adjusted P threshold = 0.05) in normalized CBF for all subtractions. 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); y-coordinate refers to anterior-posterior position relative to the anterior commissure (positive = anterior); z-coordinate refers to superior-inferior position relative to the CA-CP (anterior commissure-posterior commissure) line (positive = superior). Regions and Brodmann's areas listed are based on mapping the peak foci onto these designated areas in the atlas. L = left; R = right.

Table 4

ROI analysis of STG

Subtraction	Volume (ml)		L/R ratio
	Left	Right	
Word-baseline	10.82	4.88	2.20
Sentence—baseline	18.81	8.04	2.30
Time-reversed Word-baseline	3.90	0.02	195.00
Time-reversed Sentence3baseline	24.66	9.58	2.60

Table 5

Summary of key regions engaged in speech and nonspeech tasks compared to silence

Region	Brodmann's area	Speech		Nonspeech	
		Sentence	Word	TR Sentence	TR Word
Anterior superior temporal gyrus	22/21/38	L/R	L/R	R	-
Posterior superior temporal gyrus	22	L/R	L/R	L/R	L/R
Anterior insula/frontal operculum	-	L/R	L	L/R	L
Inferior frontal gyrus (pars orbitalis)	47	-	R	L	L
Frontopolar region	10/11	-	-	L/R	L

Table 6

Regions of significant blood flow decreases^a

Region	Brodmann's area	Coordinates (mm)			Z score
		x	y	z	
Word—baseline					
1. R precuneus	31/30	8	-64	11	-5.4
2. R superior parietal lobule	7	21	-49	54	-4.4
3. R superior parietal lobule	7	15	-60	47	-4.8
4. R inferior temporal gyrus/fusiform gyrus	19	37	-67	-2	-4.7
5. R medial occipital gyrus	19	28	-76	14	-4.7
Sentence—baseline					
6. R precuneus	7	3	-67	9	-4.4
TR Word—baseline					
7. R precuneus	7	8	-55	45	-4.2
TR Sentence—baseline					
8. L precuneus	7	-6	-69	20	-5.3
9. L superior parietal lobule	7	-17	-49	52	-4.5
10. R post cingulate	23	1	-46	25	-5.6
11. R post cingulate	23/30/31	6	-55	14	-6.1
12. R precuneus	7	1	-64	40	-5.1
13. R superior parietal lobule	7	17	-49	54	-5.6
14. R cuneus	17/18	3	-69	9	-5.7
Sentence—Word					
15. R anterior cingulate	32	6	39	14	-4.5
Sentence—TR Sentence					
16. L transverse gyrus of Heschl	42/41	-57	-177	7	-4.1
TR Sentence—TR Word					
17. R fusiform gyrus	18	30	-85	-20	-4.2
Word—TR Word					
18. L fusiform gyrus	18/19	-24	-76	-16	-4.6

^aSignificant deactivation foci that exceed the Hammersmith statistical criterion of significance (adjusted $P_{\text{threshold}} = 0.05$) in normalized CBF for all subtractions. See Table 2 footnote regarding Talairach coordinates and Brodmann's areas.