JARO
Journal of the Association for Research in Otolaryngology

Auditory Cortical Activity During Cochlear Implant-Mediated Perception of Spoken Language, Melody, and Rhythm

CHARLES J. LIMB,[1,2,3] ANNE T. MOLLOY,[1] PATPONG JIRADEJVONG,[2] AND ALLEN R. BRAUN[1]

[1]NIDCD/NIH, Language Section, 10 Center Drive, Rockville, MD 20892, USA

[2]Department of Otolaryngology-Head and Neck Surgery, Johns Hopkins Hospital, Baltimore, MD 21287, USA

[3]Peabody Institute of Music, Johns Hopkins University, Baltimore, MD 21202, USA

ABSTRACT

Despite the significant advances in language perception for cochlear implant (CI) recipients, music perception continues to be a major challenge for implant-mediated listening. Our understanding of the neural mechanisms that underlie successful implant listening remains limited. To our knowledge, this study represents the first neuroimaging investigation of music perception in CI users, with the hypothesis that CI subjects would demonstrate greater auditory cortical activation than normal hearing controls. $H_2{}^{15}O$ positron emission tomography (PET) was used here to assess auditory cortical activation patterns in ten postlingually deafened CI patients and ten normal hearing control subjects. Subjects were presented with language, melody, and rhythm tasks during scanning. Our results show significant auditory cortical activation in implant subjects in comparison to control subjects for language, melody, and rhythm. The greatest activity in CI users compared to controls was seen for language tasks, which is thought to reflect both implant and neural specializations for language processing. For musical stimuli, PET scanning revealed significantly greater activation during rhythm perception in CI subjects (compared to control subjects), and the least activation during melody perception, which was the most difficult task for CI users. These results may suggest a possible relationship between auditory performance and degree of auditory cortical activation in implant recipients that deserves further study.

Keywords: music, PET, neuroimaging, cochlear implant, melody, rhythm, language

INTRODUCTION

Although cochlear implants (CI) have been remarkably successful in promoting speech-language comprehension to individuals with profound hearing loss (Lalwani et al. 1998), music is significantly more difficult for implant users to perceive than language (Gfeller and Lansing 1991; Gfeller et al. 1997). Indeed, high-level perception of music is rarely attained by CI users (Gfeller et al. 2008), who frequently perceive music as unpleasant (Mirza et al. 2003). While implant processors (which typically extract temporal envelope) can transmit temporal patterns with high fidelity, spectral information such as that needed for pitch is transmitted in degraded fashion. Since pitch relationships between a series of notes (including contour, direction of pitch change, and interval distance) provide the basis for melody perception, melody represents one of the most difficult musical elements for CI users, whereas rhythm is significantly easier (Gfeller and Lansing 1991; Gfeller et al. 2005; Gfeller et al. 2007; Gfeller et al. 1997). Here, we examine the neural substrates of melody and rhythm processing in CI users (and control subjects), and compare cerebral activation

Correspondence to: Charles J. Limb · Peabody Institute of Music · Johns Hopkins University · Baltimore, MD 21202, USA. Telephone: +1-410-5024269; fax: +1-410-9550035; email: climb@jhmi.edu

patterns for these musical stimuli to those elicited by language stimuli.

There have been several positron emission tomography (PET) neuroimaging studies that have examined cortical activation during speech–language perception in implant users. These studies have ranged from measurements of basal metabolism in auditory cortices of deaf individuals using ^{18}F-fluoro-2-deoxy-D-glucose (Ito et al. 1993; Lee et al. 2007) to the study of auditory and non-auditory regional activation in response to speech stimuli using $H_2{}^{15}O$ as a tracer of cerebral blood flow (Fujiki et al. 1999; Giraud et al. 2001a,b; Giraud and Truy 2002; Miyamoto and Wong 2001; Wong et al. 1999; Wong et al. 2002). It has been shown that non-implanted deaf individuals display decreased levels of metabolic activity in auditory cortices in comparison to normal hearing individuals (Green et al. 2005; Ito et al. 1993; Lee et al. 2007). Studies of experienced cochlear implant users suggest that the degree of activation (in terms of both extent and intensity) of auditory cortex in response to speech stimuli corresponds to the degree of success in speech perception (Giraud et al. 2001a,b, 2000; Green et al. 2005; Nishimura et al. 2000). Recently, it was found that good implant users, but not poor users, demonstrated activation of voice selective regions along the superior temporal sulcus in response to spoken language (Coez et al. 2008). This study also identified an overall decrease in functional activation of auditory cortex in individuals with implants compared to controls (Coez et al. 2008). These findings are inconsistent with those of Naito et al. (2000), who reported significant increases in auditory cortical activity in postlingually deafened cochlear implant users for speech (Naito et al. 2000).

Numerous recent studies of normal hearing subjects have been performed over the past decade using a range of functional neuroimaging modalities (including functional MRI, PET, MEG, and EEG). These studies have revealed important differences in hemispheric lateralization for spectral and temporal processing (Zatorre and Belin 2001), activation of Broca's area (inferior frontal gyrus) during detection of musical syntactic violations (Maess et al. 2001), and stimulation of the limbic system during music-induced emotion (Blood et al. 1999), in addition to identifying neuroplastic changes associated with musical training (Gaser and Schlaug 2003). Collectively, these studies and others provide the basis for a model of functional musical processing in the brain in both normal hearing and implant-mediated listening. There have been no studies, however, that have examined perception of music in cochlear implant users using functional neuroimaging. We sought to investigate activation of auditory cortex in CI users and normal hearing subjects for musical stimuli. We used $H_2{}^{15}O$ PET scanning to examine cortical activity in postlingually deafened CI subjects and normal hearing control subjects during perception of melodic, rhythmic, and language stimuli. We hypothesized that implant users, as a function of increased neuronal recruitment during implant-mediated listening, would demonstrate activations of auditory cortex that exceed those of normal hearing control subjects for all categories of stimuli.

METHODS AND MATERIALS

Study subjects

Twenty subjects participated in the study. Ten were normal hearing individuals (seven male, three female; age range 22–56, mean 42.7±13.2 years; abbreviated as Ctrl), and ten were cochlear implant (CI) subjects (six male, four female; age range 29–61, mean 50.2± 10.4 years). All implant subjects were postlingually deafened adults with more than 1 year of implant experience. All subjects were right-handed, English-speakers, and all had right-sided cochlear implants. A range of etiologies was responsible for the hearing loss of the subjects (which varied in duration from 14 to 43 years) and devices from all three implant manufacturers were represented (Table 1, implant subjects). The study was conducted at the National Institutes of Health, Bethesda, MD. The research protocol was approved by the Institutional Review Board of the NIH and informed consent was obtained for all subjects prior to participation in the study. Inclusion criteria for participation included postlingual onset of hearing loss and at least 12 months of implant experience, in addition to right-sided implantation, right-handedness, and English fluency. Control subjects were excluded if they had musical training beyond the amateur level or actively participated in musical activities, due to the potential relationship between musical training and functional cerebral plasticity. No cochlear implant subjects had greater than amateur level musical experience prior to implantation nor did any have significant participation in musical activities following implantation.

Auditory stimuli

Three categories of auditory stimuli were presented to each subject. All stimuli were presented at most-comfortable level (MCL) using free-field stimuli through a speaker (M-audio) positioned near the right ear/implant. Most-comfortable levels were established for each subject and condition independently (Rhythm, Melody, and Language) to ensure that differences in timbre did not cause any added discomfort or sensitivity. Peak levels of sound presentation (dB SPL) according to MCL were compared for

Cochlear implant subject characteristics

Subject	Age (years)	Gender	Etiology of deafness	Duration of HL (years)	Duration of implant use (months)	Device type	Side of device
1	40	M	Meningitis/Meniere's disease	26	69	Med-El Combi 40+	Right
2	46	F	Idiopathic progressive SNHL	14	42	Advanced Bionics Clarion CII	Right
3	50	F	Idiopathic progressive SNHL	35	27	Advanced Bionics Clarion CII	Right
4	58	M	Idiopathic progressive SNHL	41	120	Cochlear nucleus 24	Right
5	58	F	Idiopathic progressive SNHL	18	121	Cochlear nucleus 22	Right
6	61	M	Idiopathic progressive SNHL	21	27	Advanced Bionics Clarion CII	Right
7	60	M	Idiopathic progressive SNHL	43	102	Advanced Bionics Clarion CII	Right
8	29	M	Enlarged vestibular aqueduct	25	52	Advanced Bionics Clarion CII	Right
9	44	F	Idiopathic progressive SNHL	28	53	Cochlear Nucleus 24	Right
10	56	M	Idiopathic progressive SNHL	36	12	Advanced Bionics Clarion CII	Right

SNHL sensorineural hearing loss

each stimuli category and group. Levels for control subjects were 81.9±2.6 dB (melody), 80.8±2.5 dB (rhythm), and 83.1±2.8 dB (language). Levels for CI subjects were 81.8±4.4 dB (melody), 81.3±4.7 dB (rhythm), and 83.3±5.3 dB (language); these levels were not statistically different by two-tailed unpaired T test. A foam earplug was used in the left ear for all subjects, such that all subjects were listening either through their healthy right ear or the CI alone, with minimal contributions from the contralateral ear. The foam earplug provided approximately ~30 dB attenuation to the left ear; no masking or further attenuation was provided to the left side in either study group. The Logic Platinum 6 sequencing environment (Apple Inc., Cupertino, CA) was used for all auditory stimuli presentation. In addition to the three conditions, a Rest condition of silence was also included. No other auditory stimuli (e.g. noise) were utilized here.

In the melody (Mel) condition, 18 songs were taken from a source list of popular melodies found to be easily recognized by the general American population (Drayna et al. 2001). These songs represented the possible choices during the melody recognition task (closed-set) and included the following: "Row, Row, Row Your Boat", "Twinkle, Twinkle Little Star", "the Star-Spangled Banner", "Mary Had a Little Lamb", "London Bridges Falling Down", "Beethoven's 9th Symphony (Ode to Joy)", "America, the Beautiful", "My Country 'Tis of Thee", "Jingle Bells", "Three Blind Mice", "Joy to the World", "Yankee Doodle", "Frere Jacques", "Hark, the Herald Angels Sing"; "Frosty the Snowman"; "Auld Lang Syne"; "Old MacDonald Had a Farm"; and "Happy Birthday To You". Of these 18 songs, six were selected for presentation during the PET scanning (one for training and five for acquisition). The songs that were selected ("America the Beautiful", "Yankee Doodle", "Beethoven's 9th Symphony (Ode to Joy)", "Auld Lang Syne", and "Hark the Herald Angels Sing") were chosen based on their identical melody length in bars, similar tempos, use of multiple pitch intervals and ability to be presented as isochronous quarter-notes only at 120 bpm, thereby removing unique rhythmic patterns for each melody. Quarter notes were selected rather than eighth notes so that the stimuli could be better matched acoustically with those presented during the Rhythm condition (below). The songs ranged in fundamental frequency from A1 to E2 (440 Hz to 1,318.5 Hz). A high-quality piano sample (Steinway Grand Piano, Eastwest Pro Samples) was used to present all melodies, which were pre-recorded in musical instrument digital interface (MIDI) format prior to presentation. Each song was exactly 32 bars in length and was repeated twice during each run. Thus, for each melody, the total number of notes heard by the subject remained constant. After listening to each melody, subjects were presented with three song title choices on a computer monitor; the three choices were culled from the pool of 18 songs and presented pseudo-randomly. Patients then verbally selected the one that most closely matched the song. Possible scores ranged from 0% (none correct) to 100% (all correct) in 20% increments. A closed-set task was used because of the high degree of difficulty anticipated for this task in the CI subjects. After being presented with each melody, subjects were also asked to identify if they were unfamiliar with any of the song titles provided as possible choices.

In the rhythm (Rhy) condition, five rhythmic patterns were presented using a high-quality snare drum sample (Mixtended Drums, Wizoo). Each pattern was derived from a basic pattern of four isochronous beats (Fig. 1). The rhythmic patterns were generated by varying the position of one or two notes by one eighth of a measure. This allowed the creation of several unique but simple patterns. The percussive snare sound had a temporal envelope that decayed with a time course similar to piano quarter
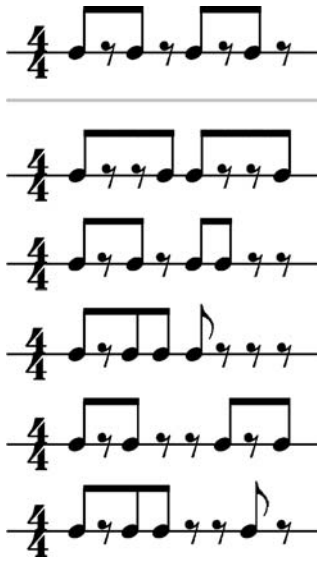
FIG. 1. Rhythmic patterns used during the rhythm condition. Rhythms were generated from an isochronous rhythm pattern, shown on the *top*, by altering the temporal placement of one or two notes by one-eighth of a measure, resulting in five unique rhythmic patterns matched for number of notes per measure.

notes used in the Melody condition. All Rhythm and Melody stimuli were additionally normalized by root mean square energy. For both categories, four notes were presented per measure at 120 bpm. During stimulus presentation, the rhythmic pattern was repeated for a total of 2.5 min (during PET acquisition, described further below). Each run therefore consisted of the same total number of notes and no musical pitch information beyond that present in the pitch envelope of the snare drum sound. After stimulus presentation, scanning was terminated and subjects were asked to reproduce the core rhythmic pattern using a MIDI keyboard (M-Audio Oxygen 8) as a drum trigger. The reproduced patterns played by each subject were recorded as MIDI data. No PET images were acquired during rhythm reproduction, and all subjects were asked to remain perfectly still with eyes closed during stimulus presentation. Following rhythm reproduction, the tester graded the response as correct or incorrect. Criteria for judgment were based on the ability of the subject to properly present the eighth note pattern in terms of temporal sequence, a feature that was easily evaluated by the testers. We did not evaluate tempo or precision of reproduction beyond whether the basic pattern of beats and rests were played back correctly. In addition, subjects reproduced multiple cycles of the rhythm, to ensure consistency in their responses. Possible scores ranged from 0% (none correct) to 100% (all correct) in 20% increments.

In the language (Lang) condition, five series of sentences taken from the Central Institute of the Deaf

(CID) sentence test battery were presented to each subject. There was a 3-s interval between sentences. Twelve sentences were presented per series. A male voice was used to present each sentence, as pre-recorded in the original CID test battery. Following the presentation of sentences, subjects were given a closed-set list of three sentences, from which they were asked to select one sentence that they heard in the preceding series. A closed-set task was chosen here as well because of anticipated difficulties for this task in the CI subjects. Possible scores ranged from 0% (none correct) to 100% (all correct) in 20% increments.

Scanning procedure

PET scans were performed on a GE Advance scanner (GE Medical Systems, Waukesha, WI) operating in 2D-acquisition mode, which acquires 35 slices simultaneously with a spatial resolution of 6.5 mm FWHM in x, y, and z axes. A transmission scan, using a rotating pin source, was performed for attenuation correction. For each scan, 10 mCi of $H_2{}^{15}O$ was injected intravenously through a catheter in the left antecubital vein in a bolus preparation over 10 s. Scans commenced automatically when the count rate in the brain reached a threshold value (approximately 20 s after injection). The scan duration was 1 min with an interscan interval of 5 min. A total of 20 runs (five rest, five melody, five rhythm, and five language) were performed for each subject, with 10 mCi per injection (total 200 mCi per scan session); the stimulus was presented at the time of injection. Stimulus presentation continued for 2 min and 30 s per run. A total of 5 min was allotted per run, for a total of 100 min scanning time.

A baseline scan was acquired 1 min before each injection. Subjects lay supine in the scanner, with access to the MIDI keyboard using their right hand. During the scan procedure, auditory stimuli were presented in pseudo-randomized fashion. Subjects were instructed to keep still, and a thermoplastic mask was used to immobilize the head in the scanner. A video monitor provided task instructions and closed-set task choices for the subject to read. Subjects closed their eyes during rest intervals and during auditory stimuli presentation.

Data analysis

All images were processed using Statistical Parametric Mapping 5 (SPM5; Wellcome Neuroimaging Department, UK). Brain images were spatially realigned, normalized, and smoothed ($15 \times 15 \times 9$ mm kernel in the x, y, and z axes) using pre-processing subroutines within SPM in order to minimize anatomical differ-

ences between subjects. Differences in global activity were controlled using proportional normalization (gray matter average per volume). Following pre-processing, all subject data were entered into a single matrix for two groups and multiple conditions for voxel-by-voxel comparison of activation data. Rest scans were used for baseline data for the appropriate subject group. In order to limit type I error, we report only differences between tasks that were also associated with significant task–rest differences (at these specified thresholds, the conjoint probability of both criteria being reached concurrently by chance is $p<10^{-6}$). Contrast analyses were performed between subject groups and between conditions using the above specifications. Local maxima of activation clusters were identified using the Montreal Neurological Institute (MNI) coordinate system, and then cross-referenced with a standard anatomical brain coordinate atlas (Talairach and Tournoux 1988).

RESULTS

Behavioral parameters and performance scores

Performance results for melody, rhythm, and language tasks were measured in terms of percentage correct. In both groups, subjects scored highest on rhythm tasks (100% in both groups), second highest on language tasks (controls 98.0±6.3%; CI 82.0±31.9%), and poorest on melody (controls 88.0±21.5%; CI 46.0±28.4%). Implant subjects scored significantly poorer on melody than on the other two conditions, just above chance (33.3% for this task). Two-tailed unpaired T tests showed a statistical difference between melody scores for both groups ($p=0.0015$), but no difference for rhythm or language ($p=0.137$) tasks. During the melody condition, both CI subjects and control groups were familiar with 146/150 (97.3%) of the song choices by name. For CI subjects, two subjects were unfamiliar with one of the presented songs by name ("Beethoven's 9th Symphony [Ode to Joy]). For control subjects, one subject was unfamiliar with "Auld Lang Syne", while another subject was unfamiliar with three of the presented songs by name ("Auld Lang Syne", "Hark the Herald Angels Sing", and "Beethoven's 9th Symphony [Ode to Joy])").

PET findings

Contrast analyses were performed within both groups [Condition > Rest], across groups ([Ctrl > CI] and [CI > Ctrl]) for each condition, across condition within groups ([Mel > Rhy], [Mel > Lang], [Rhy > Mel], [Rhy > Lang], [Lang > Mel] and [Lang > Rhy] for both Ctrl and CI groups), and across conditions and

groups (e.g. Ctrl [Mel > Rhy] > CI [Mel > Rhy]). Table 2 shows local maxima of activation clusters within auditory cortices for the [CI > Ctrl] contrast for language, melody and rhythm.

WITHIN GROUP CONTRASTS: [CATEGORY > REST]

Within CI subjects alone, contrasts for each [Category > Rest] were used to measure brain activation patterns in a category specific manner (Fig. 2, left). For these contrasts, the greatest intensity of activation was found during [Lang]. Bilateral extension to the inferior frontal gyrus (IFG) was seen as well. In general, there was left-sided bias in terms of intensity of activation, although the distribution of activity was fairly symmetric. The [Mel > Rest] condition represented the second highest intensity of activation clusters, as well as the greatest spatial extent of activation, with clusters in left superior temporal gyrus (STG), left inferior parietal lobule (IPL), right STG and right middle temporal gyrus

TABLE 2

MNI coordinates of local maxima (bold) and submaxima for the contrasts [CI>control] for language, melody, and rhythm stimuli

CI>control

x	y	z	T	Voxels	Region
Language					
−66	**−36**	**6**	**7.33**	487	**L STG/STS**
−60	−12	−6	5.16		L MTG
69	**−27**	**6**	**6.65**	674	**R STG/STS**
60	−9	−9	6.27		R MTG
57	**33**	**18**	**3.96**	24	**R IFG**
Melody					
−66	**−30**	**6**	**4.83**	69	**L STG/STS**
63	**−3**	**−9**	**4.53**	57	**R MTG**
39	**60**	**15**	**4.15**	34	**R MFG**
63	**−45**	**0**	**3.83**	56	**R MTG**
69	−36	6	3.61		R STG/STS
−33	**−54**	**42**	**3.66**	36	**L IPL**
−39	**48**	**−3**	**3.56**	21	**L MFG**
Rhythm					
60	**−3**	**−9**	**5.06**	138	**R MTG**
−63	**−30**	**6**	**5**	207	**L STG/STS**
−63	**−3**	**−6**	**4.54**		**L MTG**
69	**−27**	**9**	**4.84**	327	**R STG**
57	0	24	3.68		R PoCG
51	**−39**	**60**	**4.7**	97	**R IPL**
51	−12	54	3.37		R PoCG, PrCG
−54	**−54**	**3**	**4.4**	106	**L MTG**
−51	**9**	**−18**	**3.93**	35	**L MTG**
−39	**−9**	**57**	**3.41**	20	**L PrCG**

STG superior temporal gyrus, *STS* superior temporal sulcus, *MTG* middle temporal gyrus, *IFG* inferior frontal gyrus, *MFG* middle frontal gyrus, *IPL* inferior parietal lobule, *PoCG* postcentral gyrus, *PrCG* precentral gyrus, *R* right, *L* left
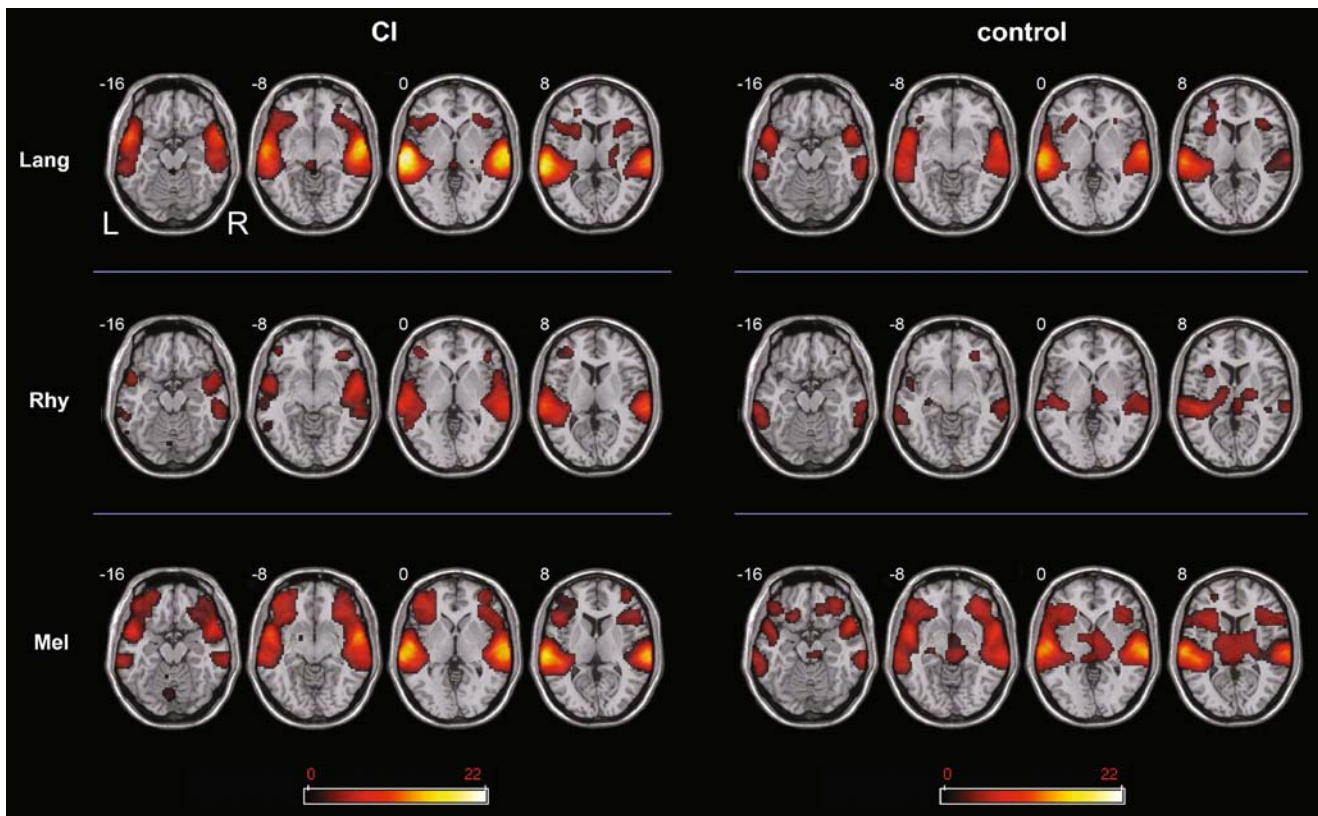
FIG. 2. Activation maps for [Lang–Rest], [Melody–Rest], and [Rhythm–Rest] for the cochlear implant patient group (CI) and controls. All 1activations were obtained using a threshold of *p*<0.001 for significance. The *scale bar* shows *t* score intensity of activation (range 0 to 22). *Abbreviations L* left; *R* right (applies to all slices); *Lang* language, *Mel* melody, *Rhy* rhythm.

(MTG). Bilateral extension to the IFG was also seen during [Mel] condition, with a leftward predominance for temporal cortical intensity of activation. The least intense activation was observed during the [Rhy > Rest] condition. In addition, this condition was associated with the least extent of activation. Activity was essentially bilaterally distributed, with left MTG, STG, inferior temporal gyrus (ITG), IPL, supramarginal gyrus (SMG), and right STG, ITG, and IPL clusters. Bilateral temporal poles (right > left) were active, in addition to small foci of activity in left IFG and right IFG.

For control subjects alone, [Category > Rest] contrasts were also used to identify category-specific activation patterns (Fig. 2, right). For these contrasts, control subjects demonstrated similar intensity maxima for both [Lang] and [Mel] conditions, although melody was associated with the greatest extent of spatial distribution for all conditions. [Rhy > Rest] was associated with the least intense and smallest spatial network of activation. During [Lang > Rest], auditory activations were seen bilaterally in STG, MTG, and temporal polar regions, with extension to the IFG, as well as left ITG. There was an overall leftward bias in both intensity and extent of distribution. The [Mel > Rest] condition showed bilateral STG and temporal polar activity as

well, in addition to bilateral extension into the IFG. These activations were overall symmetric in extent, with a slight leftward predominance in intensity for the temporal cortices. For control subjects during [Rhy > Rest], unlike in CI listeners, there was leftward dominance of activity in MTG. This condition showed both the least intense and smallest extent of activation.

BETWEEN GROUP CONTRASTS: [CI > CONTROL] AND [CONTROL > CI]

For all conditions, the [CI > Ctrl] contrast showed activity in the temporal cortices, consistent with heightened functional activation in these regions for implant subjects in comparison to normal hearing subjects (Fig. 3, left). Of the three conditions, the non-musical condition [Lang] demonstrated the greatest difference between groups, with the highest *t* score of any cluster for the [CI > Ctrl] contrasts. In particular, CI subjects demonstrated greater activity during [Lang] in bilateral temporal poles, left MTG and right STG than control subjects. In addition, a small cluster of activation in right IFG (p. triangularis) was observed. For the [Rhy] condition, CI subjects demonstrated greater activity than controls in right STG, as well as L MTG and STG.
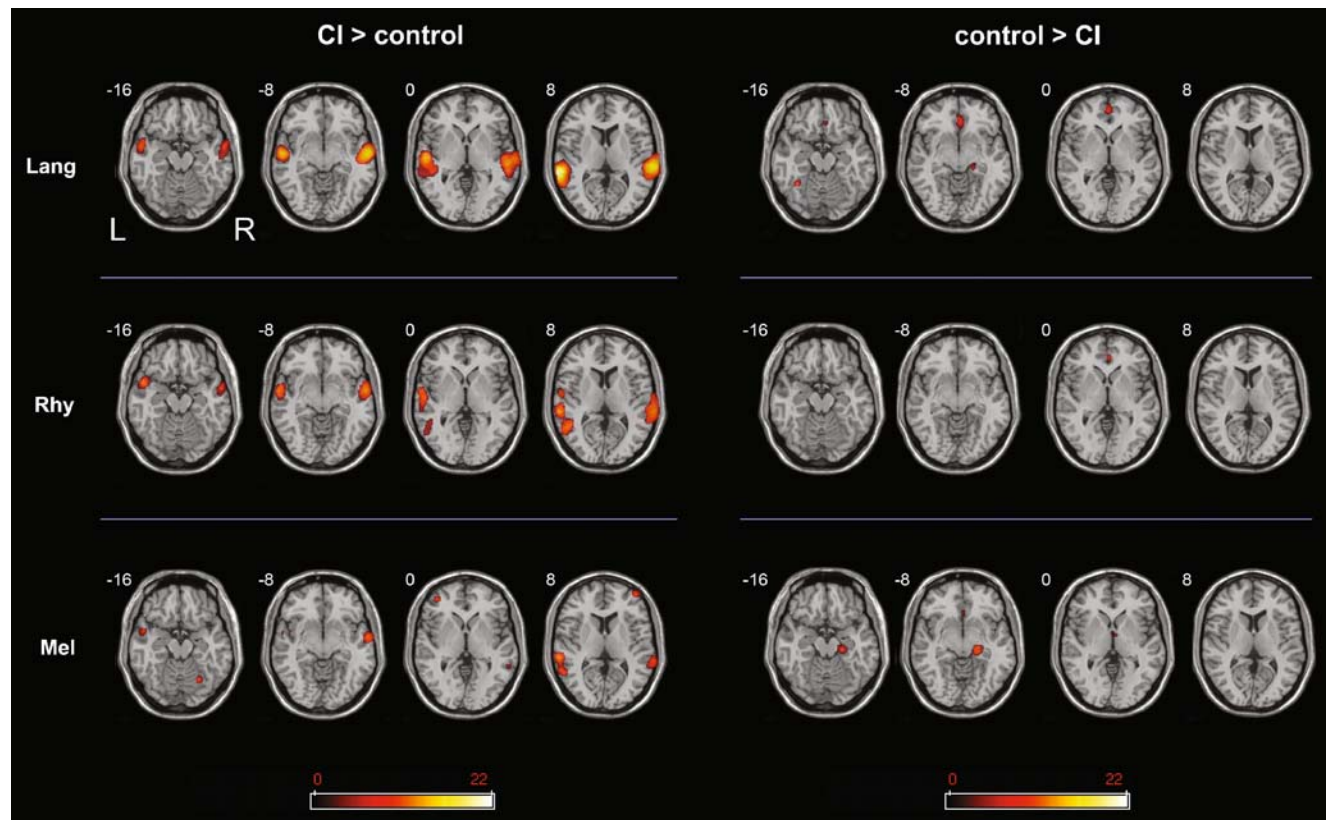
FIG. 3. Activation maps for [Lang–Rest], [Melody–Rest], and [Rhythm–Rest] for the contrasts between [CI>control] (*left*) and [control>CI] (*right*). All activations were obtained using a threshold of *p*<0.001 for significance. The *scale bar* shows *t* score intensity of activation (range 0 to 22). *Abbreviations L* left, *R* right (applies to all slices), *Lang* language, *Mel* melody, *Rhy* rhythm.

Bilateral temporal poles were active as well, while extratemporal activations were seen in the postcentral gyrus bilaterally (right > left) and right MFG. The weakest between-group differences ([CI > Ctrl] contrast) were observed during [Melody]; activations were seen in left MTG, STG, and right MTG, with additional foci of activation within both prefrontal cortices (MFG).

In comparison to the [CI > Ctrl] contrasts, in which temporal cortical activation was found for all conditions, the [Ctrl > CI] contrasts revealed no activity within the temporal cortices for any condition (Fig. 3, right). Hence, CI subjects showed greater temporal cortex activation for all conditions compared to control subjects, and there were no foci of temporal cortex that showed greater activity in control subjects than CI subjects under any condition. Instead, the most significant extemporal activity was seen in cingulate cortex, precuneus, parahippocampus, and hippocampus.

WITHIN GROUP/ACROSS CATEGORY CONTRASTS

Within CI subjects, [Lang > Mel] and [Lang > Rhy] contrasts both revealed significant activation of bilateral

temporal poles, with a left > right dominance in terms of intensity and extent for both contrasts (Fig. 4, top). For [Lang > Mel], bilateral MTG and left ITG activation were seen, while for [Lang > Rhy] left MTG and right
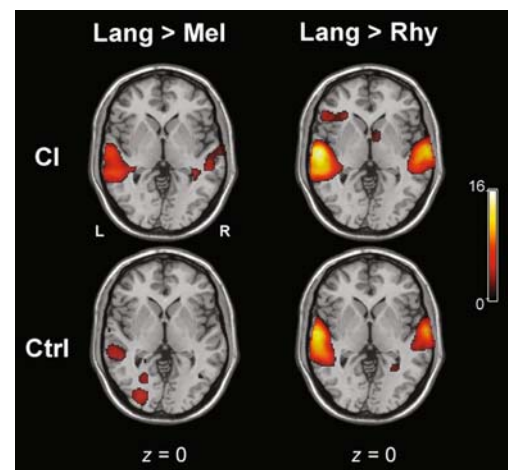

FIG. 4. Representative axial slices for the contrasts of [Language > Melody] and [Language > Rhythm] for both CI and control groups. The *scale bar* shows *t* score intensity of activation (range 0 to 16). *Abbreviations L* left, *R* right (applies to all slices), *Lang* language, *Mel* melody, *Rhy* rhythm.

STG activations were seen, with extratemporal activations in left IFG (p. orbitalis and p. triangularis) and left insula. For [Mel > Lang] in CI subjects alone, temporal cortical activation was minimal (Fig. 5), with small foci of activations in left STG, right STG and right temporal pole. In comparison to [Mel > Lang], the contrast for [Mel > Rhy] for CI subjects revealed significantly greater activity within the temporal cortices, in right STG, temporal pole, and left STG. For both [Rhy > Lang] and [Rhy > Mel] contrasts within CI subjects, the temporal cortical activation was located more posteriorly and inferiorly than the other conditions (Fig. 6, top). For [Rhy > Lang], activation of left MTG and right ITG was seen, whereas [Rhy > Mel] revealed bilateral ITG and MTG activation.

For control subjects, the [Lang > Mel] contrast revealed a high degree of activity in the temporal cortices, including bilateral temporal poles, ITG, left MTG, and right STG (Fig. 4). For [Lang > Rhy], bilateral temporal pole activation was observed, in addition to left MTG, ITG, and right STG. Co-activation of the IFG was observed, as well as extratemporal activations of left hippocampus, PHC, lingual gyrus, and fusiform gyrus. The [Mel > Lang] contrast revealed activity within bilateral STG, as well as right IFG (Fig. 5). The temporal cortical activation for the [Mel > Rhy] contrast was located superiorly within bilateral STG with bilateral extension (left > right) to the IFG, temporal poles, IFG, and left SMG (Fig. 5). For [Rhy > Lang], there was little temporal cortical activity within controls, similar to CI subjects (Fig. 6). [Rhy > Mel] also showed little overall temporal cortical activity (Fig. 6), with bilateral (but left dominant) MTG activation was seen.
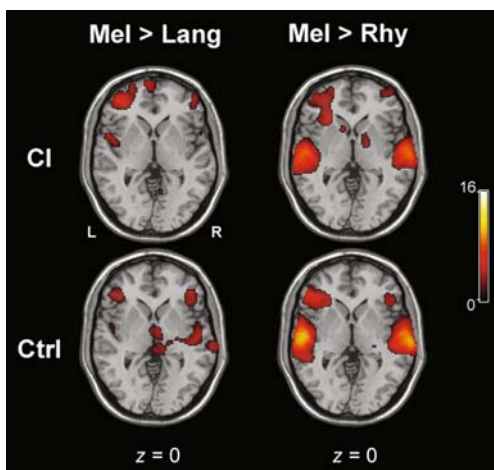

FIG. 6. Representative axial slices for the contrasts of [Rhythm > Language] and [Rhythm > Melody] for both CI and control groups. The *scale bar* shows *t* score intensity of activation (range 0 to 16). *Abbreviations L* left, *R* right (applies to all slices), *Lang* language, *Mel* melody, *Rhy* rhythm.
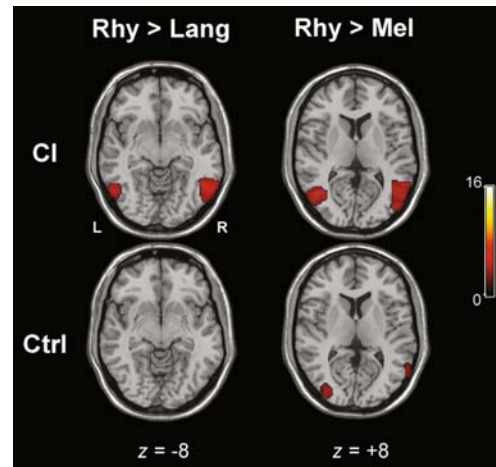
BETWEEN GROUP/ACROSS CATEGORY CONTRASTS

Contrast analyses across both groups and categories (e.g. CI [Lang > Mel] > Ctrl [Lang > Mel] and so forth) were performed. These analyses reveal that for the [CI > Ctrl] group comparison, it is the specific contrast of CI [Lang > Mel] > Ctrl [Lang > Mel] (which is mathematically equivalent to Ctrl [Mel > Lang] > CI [Mel > Lang]), that reveals the strongest auditory cortical activity, with maxima in right STG (maxima at $x=74$, $y=-26$, $z=2$; $t=4.8$; 781 voxel cluster size) and left MTG (maxima at $x=-68$, $y=-14$, $z=-10$; $t=4.65$; 338 voxel cluster size).

DISCUSSION

The relationship between functional brain activity within auditory cortex and perception of auditory stimuli is highly germane to our understanding of how individuals with CIs perceive musical stimuli. This study represents the first neuroimaging study of music perception in cochlear implant users, a category of stimulus that implant users typically find significantly more challenging to perceive than language. Due to the presence of indwelling ferromagnetic implants, functional MRI is not permissible for CI subjects. In addition, PET scanning is a relatively quiet imaging modality, which is useful for both CI subjects and for the study of music. Hence, PET scanning, despite its limitations in spatial and temporal resolution (particularly relevant for music perception), remains the best method of functional brain imaging for CI users.

FIG. 5. Representative axial slices for the contrasts of [Melody > Language] and [Melody > Rhythm] for both CI and control groups. The *scale bar* shows *t* score intensity of activation (range 0 to 16). *Abbreviations L* left, *R* right (applies to all slices), *Lang* language, *Mel* melody, *Rhy* rhythm.

In addition to difficulties in pitch perception, which remain the fundamental obstacle for melody perception in CI users, there are several additional reasons why music may be more difficult to perceive than language through an implant (Limb 2006b). First, music is an inherently abstract stimulus without clear semantic implications, unlike language, which is explicitly designed to convey specific semantic ideas. Second, music is characterized by a level of spectral, dynamic, and temporal heterogeneity that is rarely found in language. Third, music often contains multiple streams of information with widely varying spectral and temporal characteristics that must be integrated in parallel, unlike conversation, which is typically processed serially. It should be pointed out that while the individual elements that comprise music can be studied in isolation (as was done here), most forms of music contain elements of pitch/ melody and rhythm blended together and presented synchronously, together with other musical elements such as harmony and timbre. Therefore, while the component of rhythm may be easily perceived in isolation, it must also be integrated together with other components that are poorly perceived for an accurate perception of a musical excerpt or piece. How well this process of integration takes place in CI users is an area for future study.

The data we present here support the idea that heightened auditory cortical responses to sound are observed in postlingually deafened implant users, particularly during speech comprehension (Naito et al. 2000). It should be noted that brain maps derived from functional neuroimaging paradigms do not prove a causal relationship between areas that are active and a given task, a factor that must be considered during the interpretation of findings such as those described here. As shown in Table 2, the clusters of activity in the auditory region were bilaterally distributed, with the strongest activity in superior temporal gyrus/sulcus and middle temporal gyrus. These regions comprise auditory association areas along the presumptive belt and parabelt regions, which are thought to receive input directly from the primary auditory cortex in normal hearing primates and humans (Hackett et al. 2001). In our study, we never observed greater activity in auditory cortex in control subjects than in implant users, regardless of the nature of the stimulus. This observation may reflect either successful recruitment or more intense activation of neuronal substrates to assist in processing sensory information received through an implant, and such recruitment may be a prerequisite for high-level CI use. These findings may also reflect an expansion of voice-selective regions along the superior temporal sulcus (STS; Belin et al., 2000) in implant-mediated listening as an example of neural

plasticity in response to the degraded quality of incoming stimuli. It is plausible that heightened prefrontal cortical responses reflect an increase in effort required to perceive stimuli that are perceived as difficult, such as those transmitted through an implant. We observed pronounced co-activation of temporal cortex and the inferior frontal gyrus on both sides in implant users during melody, and this extension may reflect effortful processing as well as the recruitment of areas that subserve working memory (Muller et al. 2002). These data support the hypothesis that activity in the temporal cortices is greater in CI subjects in comparison to normal hearing subjects regardless of the stimulus category (language, melody, or rhythm here).

The distinct categories of stimuli that we used allowed us to evaluate differences between normal hearing and implant-mediated hearing during perception of complex linguistic and non-linguistic auditory stimuli. During melody perception, CI users demonstrated the least significant auditory cortical differences in comparison to normal listeners (Fig. 3). This finding may provide a possible explanation for the poor performance of implant users during the melody perception task. As described above, we used purely instrumental melodies here without any lyrics, which likely made the stimuli more difficult (but allowed us to distinguish pitch-based musical identification from linguistic-based identification). However, the melodies that we utilized are commonly associated with lyrics (e.g. "Happy Birthday To You"), the processing of which has shown to be related to left hemispheric top-down mechanisms for highly familiar songs (Yasui et al. 2009). It is plausible that the extratemporal activations observed during melody perception in the inferior frontal gyrus were due to lexicosemantic access in an effort to match lyrics to the presented melody (Platel et al. 1997). The rhythm task, in contrast, was found to be significantly easier than the melody task for all subjects and was accompanied in CI users by greater activations in auditory cortex than was observed during melody perception (in comparison to control subjects).

Although no other functional neuroimaging data on music perception in CI users are available for comparison, results in normal hearing listeners can provide a useful basis for comparison. While it appears that both right and left hemispheres are important for perception of pitch and melody, a PET study by Zatorre and Belin (2001) showed that spectrally varying stimuli (thought to be important for melody recognition) showed a bias towards the anterior region of right primary auditory cortex while temporally varying stimuli (thought to be important for rhythm perception) showed a leftward bias, an intriguing finding in the context of well-described

left-dominant language processing in right-handed individuals within the perisylvian language cortices (superior temporal gyrus/sulcus, inferior frontal gyrus, and inferior parietal lobule). These findings were further refined by a functional imaging study by Sakai et al. of metric and non-metrical rhythm perception, in which it was found that perception of non-metrical rhythm patterns led to increased right hemispheric activation in prefrontal, premotor, and parietal cortex in comparison to metrical rhythm patterns, which were left hemisphere dominant in premotor, prefrontal, and cerebellum (Sakai et al. 1999). A study by Platel et al. utilized PET imaging to study perception of several basic musical elements in six normal hearing volunteers without music training. This study primarily examined differences across conditions (rather than for each condition vs. rest, thereby eliminating common regions of auditory cortex activation), and showed extratemporal activations in left inferior Broca's area and adjacent insular cortex for rhythm, and left cuneus/precuneus for pitch, interpreted by the authors as a visual mental imagery strategy employed by individuals during the pitch task (Platel et al. 1997).

It should be pointed out here that melody, rhythm, and language are intrinsically different stimuli that cannot be naturally matched to one another temporally and spectrally without rendering the stimuli unmusical. We chose therefore to minimize potential confounds within individual categories by eliminating rhythmic aspects from melody identification and eliminating pitch processing features from rhythm reproduction. Since control subjects performed most poorly on melody recognition, it is plausible that the rhythm task used here was relatively easy and that our melody task was relatively difficult, thereby reducing the validity of a direct comparison between performance scores for different categories. For this reason, we focused most of our data analysis on the differences between groups for identical stimuli or category contrasts, rather than on the differences in stimulation between categories within each group alone. Overall, implant subjects demonstrated the greatest difference in activity in comparison to control subjects during language perception. This finding is particularly relevant in light of the fact that cochlear implants are essentially designed for speech processing, rather than music processing. We found that language stimuli evoked the greatest activity in auditory cortices in implant users, a mechanism that may underlie the typically high level of speech performance of most postlingually deafened CI users. Interestingly, we observed left-hemispheric lateralization in CI users when comparing cortical responses for language to responses for

melody or rhythm, which likely reflects the fact that implant devices are optimized for speech perception, and that even in the CI population, cortical responses to language rely upon a left-lateralized network of auditory cortex that is specialized for language. The exact relationship between degree of auditory cortical activation and auditory performance (for example, during perception of melodies, which was both difficult for CI subjects and associated with the least amount of cortical activation in comparison to controls) is an important subject that requires further investigation.

In conclusion, numerous differences exist between cortical mechanisms of sound processing in individuals with CIs and individuals with normal hearing. These differences are attributable to the multifactorial effects of auditory deprivation on neural structure and function, diminished auditory experience during conditions of deafness, and imperfect transmission of sound by the implant. It should be emphasized that our subjects were all postlingual implant recipients. Further studies remain to be performed on prelingual implant users or extremely poor implant users, for whom functional neuroimaging studies with $H_2^{15}O$ PET may provide a clinically useful tool. Open-set language acquisition, once considered the benchmark for successful implant performance during the early history of cochlear implantation, is now a reasonable expectation for most postlingually deafened implant users (Lalwani et al. 1998), whereas perception of nonlinguistic stimuli such as music has been relatively neglected in clinical settings (Limb 2006a). Through further studies, high-level music perception may eventually become the new benchmark for implant success. Ultimately, technological advances and the development of processing strategies geared specifically for music should lead to improvements in the ability of CI users to hear the full range of sounds that comprise the auditory world.

ACKNOWLEDGMENTS

REFERENCES

Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. Nature. 403:309–312, 2000.

Blood AJ, Zatorre RJ, Bermudez P, Evans AC. Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. Nat Neurosci. 2:382–387, 1999.

Coez A, Zilbovicius M, Ferrary E, Bouccara D, Mosnier I, Ambert-Dahan E, et al. Cochlear implant benefits in deafness rehabilitation: PET study of temporal voice activations. J Nucl Med. 49:60–67, 2008.

DRAYNA D, MANICHAIKUL A, DE LANGE M, SNIEDER H, SPECTOR T. Genetic correlates of musical pitch recognition in humans. Science. 291:1969–1972, 2001.

FUJIKI N, NAITO Y, HIRANO S, KOJIMA H, SHIOMI Y, NISHIZAWA S, ET AL. Correlation between rCBF and speech perception in cochlear implant users. Auris Nasus Larynx. 26:229–236, 1999.

GASER C, SCHLAUG G. Brain structures differ between musicians and non-musicians. J Neurosci. 23:9240–9245, 2003.

GFELLER K, LANSING CR. Melodic, rhythmic, and timbral perception of adult cochlear implant users. J Speech Hear Res. 34:916–920, 1991.

GFELLER K, WOODWORTH G, ROBIN DA, WITT S, KNUTSON JF. Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users. Ear Hear. 18:252–260, 1997.

GFELLER K, OLSZEWSKI C, RYCHENER M, SENA K, KNUTSON JF, WITT S, ET AL. Recognition of "real-world" musical excerpts by cochlear implant recipients and normal-hearing adults. Ear Hear. 26:237–250, 2005.

GFELLER K, TURNER C, OLESON J, ZHANG X, GANTZ B, FROMAN R, ET AL. Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. Ear Hear. 28:412–423, 2007.

GFELLER K, OLESON J, KNUTSON JF, BREHENY P, DRISCOLL V, OLSZEWSKI C. Multivariate predictors of music perception and appraisal by adult cochlear implant users. J Am Acad Audiol. 19:120–134, 2008.

GIRAUD AL, TRUY E, FRACKOWIAK RS, GREGOIRE MC, PUJOL JF, COLLET L. Differential recruitment of the speech processing system in healthy subjects and rehabilitated cochlear implant patients. Brain. 123(Pt 7):1391–1402, 2000.

GIRAUD AL, PRICE CJ, GRAHAM JM, FRACKOWIAK RS. Functional plasticity of language-related brain areas after cochlear implantation. Brain. 124:1307–1316, 2001a.

GIRAUD AL, PRICE CJ, GRAHAM JM, TRUY E, FRACKOWIAK RS. Cross-modal plasticity underpins language recovery after cochlear implantation. Neuron. 30:657–663, 2001b.

GIRAUD AL, TRUY E. The contribution of visual areas to speech comprehension: a PET study in cochlear implants patients and normal-hearing subjects. Neuropsychologia. 40:1562–1569, 2002.

GREEN KM, JULYAN PJ, HASTINGS DL, RAMSDEN RT. Auditory cortical activation and speech perception in cochlear implant users: effects of implant experience and duration of deafness. Hear Res. 205:184–192, 2005.

HACKETT TA, PREUSS TM, KAAS JH. Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. J Comp Neurol. 441:197–222, 2001.

ITO J, SAKAKIBARA J, IWASAKI Y, YONEKURA Y. Positron emission tomography of auditory sensation in deaf patients and patients with cochlear implants. Ann Otol Rhinol Laryngol. 102:797–801, 1993.

LALWANI AK, LARKY JB, WAREING MJ, KWAST K, SCHINDLER RA. The Clarion Multi-Strategy Cochlear Implant–surgical technique, complications, and results: a single institutional experience. Am J Otol. 19:66–70, 1998.

LEE HJ, GIRAUD AL, KANG E, OH SH, KANG H, KIM CS, ET AL. Cortical activity at rest predicts cochlear implantation outcome. Cereb Cortex. 17:909–917, 2007.

LIMB CJ. Cochlear implant-mediated perception of music. Curr Opin Otolaryngol Head Neck Surg. 14:337–340, 2006a.

LIMB CJ. Structural and functional neural correlates of music perception. Anat Rec A Discov Mol Cell Evol Biol. 288:435–446, 2006b.

MAESS B, KOELSCH S, GUNTER TC, FRIEDERICI AD. Musical syntax is processed in Broca's area: an MEG study. Nat Neurosci. 4:540–545, 2001.

MIRZA S, DOUGLAS S, LINDSEY P, HILDRETH T, HAWTHORNE M. Appreciation of music in adult patients with cochlear implants: a patient questionnaire. Cochlear Implants Int. 4:85–95, 2003.

MIYAMOTO RT, WONG D. Positron emission tomography in cochlear implant and auditory brainstem implant recipients. J Commun Disord. 34:473–478, 2001.

MULLER NG, MACHADO L, KNIGHT RT. Contributions of subregions of the prefrontal cortex to working memory: evidence from brain lesions in humans. J Cogn Neurosci. 14:673–686, 2002.

NAITO Y, TATEYA I, FUJIKI N, HIRANO S, ISHIZU K, NAGAHAMA Y, ET AL. Increased cortical activation during hearing of speech in cochlear implant users. Hear Res. 143:139–146, 2000.

NISHIMURA H, DOI K, IWAKI T, HASHIKAWA K, OKU N, TERATANI T, ET AL. Neural plasticity detected in short- and long-term cochlear implant users using PET. Neuroreport. 11:811–815, 2000.

PLATEL H, PRICE C, BARON JC, WISE R, LAMBERT J, FRACKOWIAK RS, ET AL. The structural components of music perception. A functional anatomical study. Brain. 120(Pt 2):229–243, 1997.

SAKAI K, HIKOSAKA O, MIYAUCHI S, TAKINO R, TAMADA T, IWATA NK, ET AL. Neural representation of a rhythm depends on its interval ratio. J Neurosci. 19:10074–10081, 1999.

TALAIRACH J, TOURNOUX P. Co-planar Stereotaxic Atlas of the Human Brain. New York, Thieme Medical Publishers, 1988.

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. 132:34–42, 1999.

WONG D, PISONI DB, LEARN J, GANDOUR JT, MIYAMOTO RT, HUTCHINS GD. PET imaging of differential cortical activation by monaural speech and nonspeech stimuli. Hear Res. 166:9–23, 2002.

YASUI T, KAGA K, SAKAI KL. Language and music: differential hemispheric dominance in detecting unexpected errors in the lyrics and melody of memorized songs. Hum Brain Mapp. 30:588–601, 2009.

ZATORRE RJ, BELIN P. Spectral and temporal processing in human auditory cortex. Cereb Cortex. 11:946–953, 2001.