UC Merced
UC Merced Previously Published Works

Title
Cortical Activation Patterns Correlate with Speech Understanding After Cochlear Implantation

Permalink
https://escholarship.org/uc/item/7k67n479

Journal
Ear & Hearing, 37(3)

ISSN
0196-0202

Authors
Olds, Cristen
Pollonini, Luca
Abaya, Homer
et al.

Publication Date
2016-05-01

DOI
10.1097/aud.0000000000000258

Peer reviewed

Cortical Activation Patterns Correlate With Speech Understanding After Cochlear Implantation

Cristen Olds,[1] Luca Pollonini,[2] Homer Abaya,[1] Jannine Larky,[1] Megan Loy,[1] Heather Bortfeld,[3] Michael S. Beauchamp,[4] and John S. Oghalai[1]

Objectives: Cochlear implants are a standard therapy for deafness, yet the ability of implanted patients to understand speech varies widely. To better understand this variability in outcomes, the authors used functional near-infrared spectroscopy to image activity within regions of the auditory cortex and compare the results to behavioral measures of speech perception.

Design: The authors studied 32 deaf adults hearing through cochlear implants and 35 normal-hearing controls. The authors used functional near-infrared spectroscopy to measure responses within the lateral temporal lobe and the superior temporal gyrus to speech stimuli of varying intelligibility. The speech stimuli included normal speech, channelized speech (vocoded into 20 frequency bands), and scrambled speech (the 20 frequency bands were shuffled in random order). The authors also used environmental sounds as a control stimulus. Behavioral measures consisted of the speech reception threshold, consonant-nucleus-consonant words, and AzBio sentence tests measured in quiet.

Results: Both control and implanted participants with good speech perception exhibited greater cortical activations to natural speech than to unintelligible speech. In contrast, implanted participants with poor speech perception had large, indistinguishable cortical activations to all stimuli. The ratio of cortical activation to normal speech to that of scrambled speech directly correlated with the consonant-nucleus-consonant words and AzBio sentences scores. This pattern of cortical activation was not correlated with auditory threshold, age, side of implantation, or time after implantation. Turning off the implant reduced the cortical activations in all implanted participants.

Conclusions: Together, these data indicate that the responses the authors measured within the lateral temporal lobe and the superior temporal gyrus correlate with behavioral measures of speech perception, demonstrating a neural basis for the variability in speech understanding outcomes after cochlear implantation.

Key Words: Cochlea, Hearing loss, Language, Neuroimaging, Speech discrimination.

(Ear & Hearing 2015;XX;00–00)

INTRODUCTION

Cochlear implants (CIs) are commonly used to treat severe-to-profound sensorineural hearing loss when hearing aids are ineffective in helping a patient understand speech. Modern CIs work by splitting the sound frequency spectrum into 12 to 22 discrete channels, each electrically stimulating a different region of the cochlea (Clark 2015). Because auditory neurons are distributed along the tonotopic gradient of the cochlea, sounds of different frequencies thus stimulate different auditory nerve fibers. This is in contrast to normal hearing where approximately 4000 inner hair cells exist to convert sound pressure waves into electrical signals. Nevertheless, most patients demonstrate large improvements in speech understanding after cochlear implantation (Gfeller et al. 2002; Lu et al. 2014). While speech understanding increases with time, it tends to stabilize by 6 to 12 months after implantation (Holden et al. 2013), with the ultimate level of benefit remaining highly variable and difficult to predict (Gates et al. 1995; Perez & Macias 2004). In most cases, the reason any given patient ends up having a poor speech perception outcome after cochlear implantation is unknown. The presumption is that the implant is not able to effectively convey the temporal and frequency characteristics of speech to the auditory nerve.

Speech perception occurs within and beyond the auditory cortex, and thus neuroimaging could provide a way to supplement behavioral measures in assessing the ability of the speech information provided by the CI to correctly reach and stimulate the language centers of the brain (Pasley et al. 2012; Steinschneider et al. 2014). In addition, neuroimaging may be able to provide a more immediate measure than behavioral testing is able to provide. The lateral temporal lobe (LTL) and the superior temporal gyrus (STG) are probably the most clinically relevant regions of the cortex to image to study speech perception in individuals hearing through CIs. These core areas process acoustic parameters such as pitch, tone, and spatiotemporal fluctuation (Hall & Plack 2009; Humphries et al. 2010; Lakatos et al. 2013). More recently, these areas have been viewed as a more general purpose acoustic problem solver, with neural tuning to complex sound patterns such as speech (Belin et al. 2002, 2004; Fecteau et al. 2004; Mesgarani et al. 2014). The LTL/STG show selective responses to species-specific vocalizations in humans and other mammals, such as the marmoset (Wang 2000; Belin et al. 2002). Recent studies utilizing functional magnetic resonance imaging (fMRI) and implanted recording electrodes suggest that phonemes, words, and phrases elicit localized responses within the human LTL/STG (DeWitt & Rauschecker 2012; Grodzinsky & Nelken 2014; Mesgarani et al. 2014). Similarly, fMRI has demonstrated rapid neural adaptations in normal-hearing participants exposed to degraded sound similar to what a CI user experiences (Smalt et al. 2013).

However, commonly used neuroimaging techniques do have downsides when used in implanted individuals. While fMRI is the most common technique to measure human brain function, the ferromagnetic components in a CI are not compatible with the strong magnetic fields present in a magnetic resonance imaging (MRI) scanner. Positron emission tomography (PET)

imaging uses radioactive compounds that are not well suited for routine clinical use that might require regular repeated testing as the CI program parameters are adjusted. EEG is a feasible and frequently used technique for measuring cortical activations to speech in CI recipients (Senkowski et al. 2014). Key EEG studies have argued that late auditory evoked potentials provide a useful objective metric of performance in participants hearing through a CI (Firszt et al. 2002; Zhang et al. 2010, 2011). However, EEG does have a disadvantage in that CIs produces electrical noise that can interfere with recordings when long-duration speech stimuli are used although artifact removal techniques can be used to minimize this issue (Viola et al. 2012; Mc Laughlin et al. 2013; Miller & Zhang 2014).

In contrast, functional near-infrared spectroscopy (fNIRS) is amenable to neuroimaging in this patient population because it is CI compatible, safe for repeated testing sessions, and small enough to use in a standard clinic. The technique is based on the differential absorption of near-infrared light by oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) to detect the hemodynamic changes in biological tissues. In the brain, this corresponds with neural activity (Scholkmann et al. 2014). fNIRS has a long history of use in basic science research but is not widely used clinically (Villringer et al. 1993; Zaramella et al. 2001; Bortfeld et al. 2007; Mahmoudzadeh et al. 2013; Lloyd-Fox et al. 2014). Previously, our group demonstrated that fNIRS can detect speech-evoked activation of the LTL/STG, a core region of the auditory cortex in normal-hearing and implanted participants. In addition, we found that adults with normal hearing exhibit greater cortical activation to speech than nonspeech sounds (Sevy et al. 2010; Pollonini et al. 2014). Here, we expand this approach in participants with CIs to explore the use of fNIRS as an objective measure of speech perception.

MATERIALS AND METHODS

Study Design

The study protocol was approved by the Stanford University Institutional Review Board (IRB); all participants signed an informed consent form before participation. Criteria for inclusion and exclusion of participants were established prospectively; criteria for inclusion were age older than 18 years, English fluency, and a functioning CI postactivation. Participants were excluded from the study if they were not fluent in English or had a nonfunctional CI. Normal-hearing controls were healthy adult volunteers who passed a 30-dB hearing level screening test at 500, 1000, and 2000 Hz. In our block design experiment, the participant was exposed to five trials of each sound stimulus. Before beginning the experiment, it was hypothesized that cortical hemodynamic response areas among adults with CIs would correlate with speech perception scores.

Cohorts

The study participants consisted of two cohorts (Table 1). One cohort consisted of 32 postlingually deafened adults with CIs ranging in age from 23 to 86 years, with time from CI activation ranging between 1 day and 12 years. The second cohort consisted of 35 adults with normal hearing, ranging in age from 24 to 65 years. CIs from all three Food and Drug Administration (FDA)-approved brands (Cochlear, Advanced Bionics, and Med-El) were represented in the participants tested. Most adults had only one implant (30 of 32; evenly split between right and

TABLE 1. Demographic characteristics of controls and adults with CIs

	Controls	Adults With CI
	(n = 35)	(n = 32)
Mean age, yr, (SD)	45 (15)	60 (17.5)
Female, n (%)	17 (49)	18 (56)
Right handed, n (%)	33 (94)	29 (91)
Time from implantation, mo (SD)	—	29 (24)
Duration of severe-profound HL, yr (SD)	—	13 (12)
CI side, n (%)		
Right	—	15 (46.8)
Left	—	15 (46.8)
Bilateral	—	2 (6.25)

Demographic variables correspond to mean (SD) or n (%) as indicated.
CI, cochlear implant; HL, hearing loss.

left ears), whereas two adults had bilateral implants. Participants that were unilaterally implanted all had pure-tone averages and speech reception thresholds in the contralateral ear that were >60 dB HL before their implantation surgery. No participant had been implanted for unilateral sensorineural hearing loss.

Behavioral Measures of Speech Perception

Each participant hearing through a CI underwent behavioral speech perception testing (Table 2). All behavioral measurements used in this study were obtained on or near the day of fNIRS testing (within 3 months) by a licensed audiologist at the participant's usual clinic appointment. Stimuli were presented in a sound field, and participants with hearing aids in the unimplanted ear removed their hearing aid before the experiment. Participants with CIs were first exposed to speech stimuli with the device on and then the experiment was repeated with the device off (i.e., removed from the side of the head).

The speech recognition threshold (SRT) was measured to assess the hearing level (American Speech-Language-Hearing Association 1988). Postimplantation SRTs ranged from 10 to 35 dB HL, with all but one of the participants being in the range of <30 dB HL, correlating to mild or no hearing loss (American Speech-Language-Hearing Association 1988). Speech perception was measured using monosyllabic consonant-nucleus-consonant (CNC) words scores and AzBio sentence recognition scores (Peterson & Lehiste 1962; Spahr et al. 2014). These stimuli came from audio recordings and were presented in silence at 60 dB SPL. Many participants had both CNC words and AzBio tests performed, whereas some only had one or the other performed.

Acoustic Stimuli

Four types of speech stimuli were presented to all participants. Stimuli were generated in MATLAB (R2013A; The MathWorks), and the volume of each stimulus was set to a comfortable listening level of 60 dB SPL (Fig. 1). *Normal speech* consisted of digital recordings of a male voice narrating an audiobook version of *The War of the Worlds* (by H. G. Wells) in English, which were digitally edited into 20-sec sequential segments (Audio File S1, Supplemental Digital Content 1, http://links.lww.com/EANDH/A243). The second stimulus was *channelized speech*, in which the frequency spectrum was broken into 20 logarithmically spaced channels between 250 and 8000 Hz (the typical range processed by a CI), creating speech that was intelligible

TABLE 2. Audiology test descriptions, grading scales, and examples

Test Name	Description	Grading	Example Words/Phrase
Speech recognition threshold	The quietest level at which a person can understand words. The patient repeats or points to images representing spondees (two-syllable words with equal stress on each syllable).	dB HL	Hotdog, airplane, pinwheel
Consonant-nucleus-consonant words	The patient repeats 50 recorded single-syllable words played at 60 dB SPL; words have the same phonemic distribution as the English language.	0–100%	Jar, tough, patch
AzBio sentence test	The patient repeats sentences spoken by multiple male and female talkers in a conversational, rather than deliberate, speaking style.	0–100%	I could hear another conversation through the cordless phone.

but lacked the fine detail of natural speech (Audio File S2, Supplemental Digital Content 2, http://links.lww.com/EANDH/A244), i.e., vocoded speech (Shannon et al. 1995). The envelope was extracted for each frequency band using the short-time fast Fourier transform "spectrogram" function in MATLAB, with a Hamming window of 31 msec and a 45% overlap between windows. White noise that was band-pass filtered to the frequency range of each channel was then modulated by the envelope signal. The third stimulus was *scrambled speech*, which was the same as channelized speech except that the spectral envelope for each channel was randomly assigned to a different channel. While the frequency range, total energy content, and timing of the speech remained unchanged from channelized speech, the stimulus was not intelligible to normal-hearing participants (Audio File S3, Supplemental Digital Content 3, http://links.lww.com/EANDH/A245). The fourth stimulus was *environmental sounds* and was designed to serve as a nonspeech control. It consisted of 10 two-second clips of animal vocalizations and construction sounds that were spliced together to make a 20-sec stimulus (Audio File S4, Supplemental Digital Content 4, http://links.lww.com/EANDH/A246). Each of the five trials consisted of random permutations of these environmental sounds. The environmental sounds naturally varied in frequency spectrum from speech but were energy matched to the speech stimuli.

fNIRS Testing Procedure

fNIRS testing was performed in a quiet, darkened room equipped with a computer possessing all necessary data-processing software. For the duration of the experiment, participants were seated in a chair in front of the computer screen displaying a silent visual stimulus consisting of moving geometric shapes intended to maintain their attention and minimize head movement. Participants were exposed to speech stimuli from speakers placed directly in front of the listener. Five trials of each sound stimulus were presented in a pseudorandom order, alternating with 20-sec blocks of silence. Each session took 20 to 30 min to complete. Just like with the behavioral testing procedures, participants with hearing aids in the unimplanted ear removed their hearing aid before the experiment. Participants with CIs were first exposed to speech stimuli with the device on and then the experiment was repeated with the device off (removed from the side of the head). During data collection,
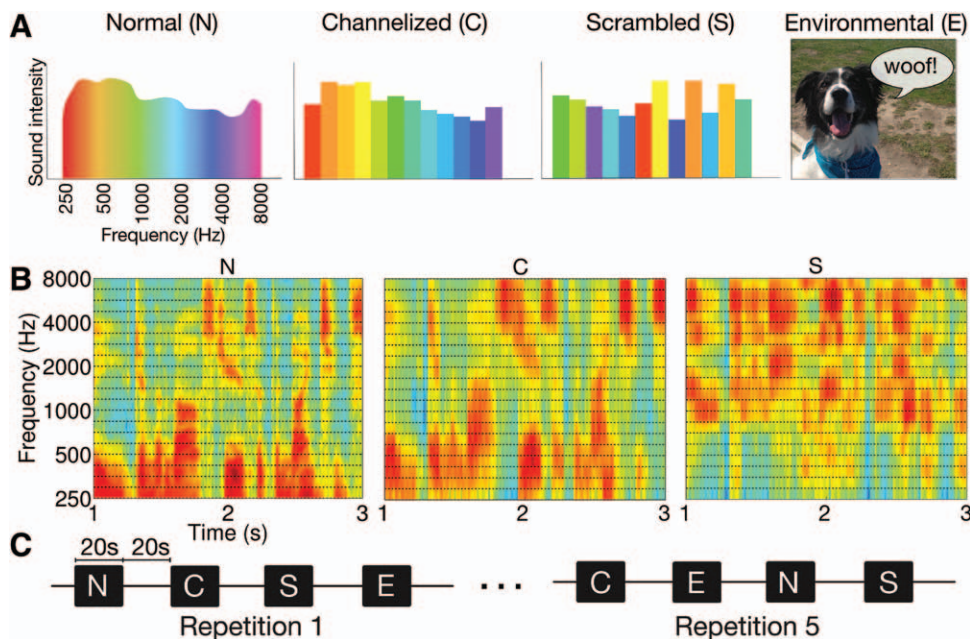

Fig. 1. Sound stimuli and experimental paradigm. A, Schematized illustration of the Fourier transform at one-time segment for normal speech, channelized speech, and scrambled speech. Environmental sounds were shuffled 2-sec clips of animal vocalizations and construction sounds. B, Spectrograms of segments of speech stimuli created from the same audio file for normal speech (N, left), channelized speech (C, center), and scrambled speech (S, right). C, Auditory stimulus exposure protocol, illustrating five repetitions of each stimulus type in pseudorandom order.

the start and end of auditory stimuli were synchronized with the incoming fNIRS data and recorded in an event file, which recorded the timing of the beginning and end of each stimulus.

fNIRS Hardware

We used an NIRScout 1624 (NIRx Medical Technologies, LLC, Glen Head, NY) instrument containing 16 dual-wavelength infrared light sources and 24 detectors. Each illumination optode consisted of two light-emitting diodes (LEDs) emitting at 760 and 850 nm near-infrared radiation. Backscattered light returning to the scalp's surface was collected and carried back to the instrument by fiber-optic cables and transduced into electrical signals by the photo-detectors. Signals at the two wavelengths were separated by sequentially activating the sources. The software that came with the device was used to collect the data from all channels at a rate of 6.25 Hz.

A custom headpiece was created to hold the light sources and detectors (optodes) in place against the participant's scalp. Sources and detectors were alternated to create a checkerboard pattern with 15-mm center-to-center distance between adjacent optodes (Fig. 2). The optodes for each hemisphere were secured in a scaffold of flexible black polypropylene film (0.30-in thick; part #1451T21 [3-288-1143-20T21]; McMaster-Carr, Santa Fe Springs, CA) by rubber O-rings supplied by the instrument's manufacturer. This arrangement yielded two symmetrical optode holders, one for each hemisphere. The optode holders were connected to one another with Velcro straps, making the headpiece adjustable to the participant's head size and shape.

The optode holders were placed against the scalp centered at the T7/T8 position (based on the modified international 10 to 20 system suggested by the American Electroencephalographic Society), with the second column of optodes being situated directly superior to the tragus and the bottom of the headset sitting in the sulcus between the pinna and the skull (Klem et al. 1999). The light from each source was received by neighboring detectors in the array, giving a potential of 96 source–detector pairings (or channels) on each side of the head. To optimally probe the cortex, we only analyzed the 31 channels in which the source–detector distance was 3.0 to 3.3 cm (Sevy et al. 2010; Beauchamp et al. 2011; Pollonini et al. 2014).
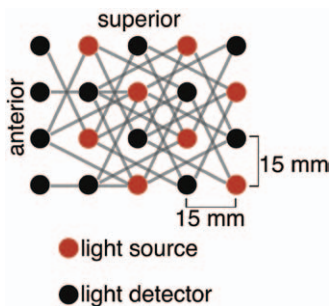

Fig. 2. Channel configuration, left hemisphere. Channels (source–detector pairs; gray lines) used for our data analysis were those with a source–detector distance of 3.0 to 3.3 cm. Channels with shorter distances were not analyzed as they predominantly probe the superficial tissues of the scalp. Channels with larger distances were not analyzed because the signal-to-noise ratio was too low to give meaningful data.

The Localization of Cortical Responses Measured by fNIRS

With this type of headset localization, our fNIRS setup probes regions within the auditory cortex, specifically the LTL/STG. Previously, we used a mathematical model (Dehghani et al. 2003; Huppert et al. 2006) to estimate the coverage area for the simple optode probe set configuration we were using at that time (Sevy et al. 2010). This supported our prediction that the headset provided excellent coverage of the LTL that extended into to the STG. This prediction was confirmed after fNIRS and fMRI demonstrated similar response patterns within the LTL/STG to auditory stimuli. We then validated the expanded fNIRS headset we use here by comparing measurements made using fNIRS and fMRI and finding a similar pattern of responses within the LTL/STG to the various stimuli (Pollonini et al. 2014).

Data Analysis

The analysis of the fNIRS data was performed in a similar manner to our previous publication (Pollonini et al. 2014). In essence, this involved removing channels from the analysis that did make good contact with the scalp, removing motion artifacts, band-pass filtering to the time course of the stimuli, and the calculation of changes in oxyhemoglobin and deoxyhemoglobin using the modified Beer–Lambert Law. We then performed statistical analysis to compare the change in hemoglobin over time to the predicted response. Statistically significant responses were then interpolated over a two-dimensional (2D) grid to quantify the area of activation to each stimuli. These steps are detailed below.

Preprocessing • All fNIRS data were processed identically using custom software created in MATLAB (R2013A; The MathWorks). First, optodes with satisfactory scalp contact were identified, so channels with favorable signal-to-noise ratios could be analyzed further, excluding those channels resulting from optodes with poor scalp contact (Fig. 3A). Adequate scalp contact is denoted by a synchronous cardiac pulse signal detected by both wavelengths of light directed from a single probe. Raw data were filtered between 0.5 and 2.5 Hz to isolate the portion of the signal attributable to the cardiac pulse (for heart rates between 30 and 150 beats/min) and normalized to account for differences in amplitude between the signals from the two wavelengths. The two wavelengths' signals were then used to generate a correlation coefficient between 0 and 1.0, the scalp contact index (SCI). Channels with an SCI below 0.70 were discarded.

Next, signal artifacts due to motion were removed (Fig. 3B). Each channel was band-pass filtered between 0.1 and 3.0 Hz to remove slow signal drift, and the channel's voltage was normalized to the intensity of the highest peak. Peaks in the signal exceeding 20% of the maximum peak intensity were identified as motion artifact. We went back to the raw data tracing and removed these artifacts by linearly interpolating between the start and stop time points. To prevent flattening a clean signal (without motion artifact), only those channels where <2% of the data samples were occupied by peaks were processed in this manner.

To remove cardiac and respiratory contributions to the hemodynamic signal, the raw data from each channel were band-pass filtered between 0.016 and 0.25 Hz (Fig. 3C), and the modified
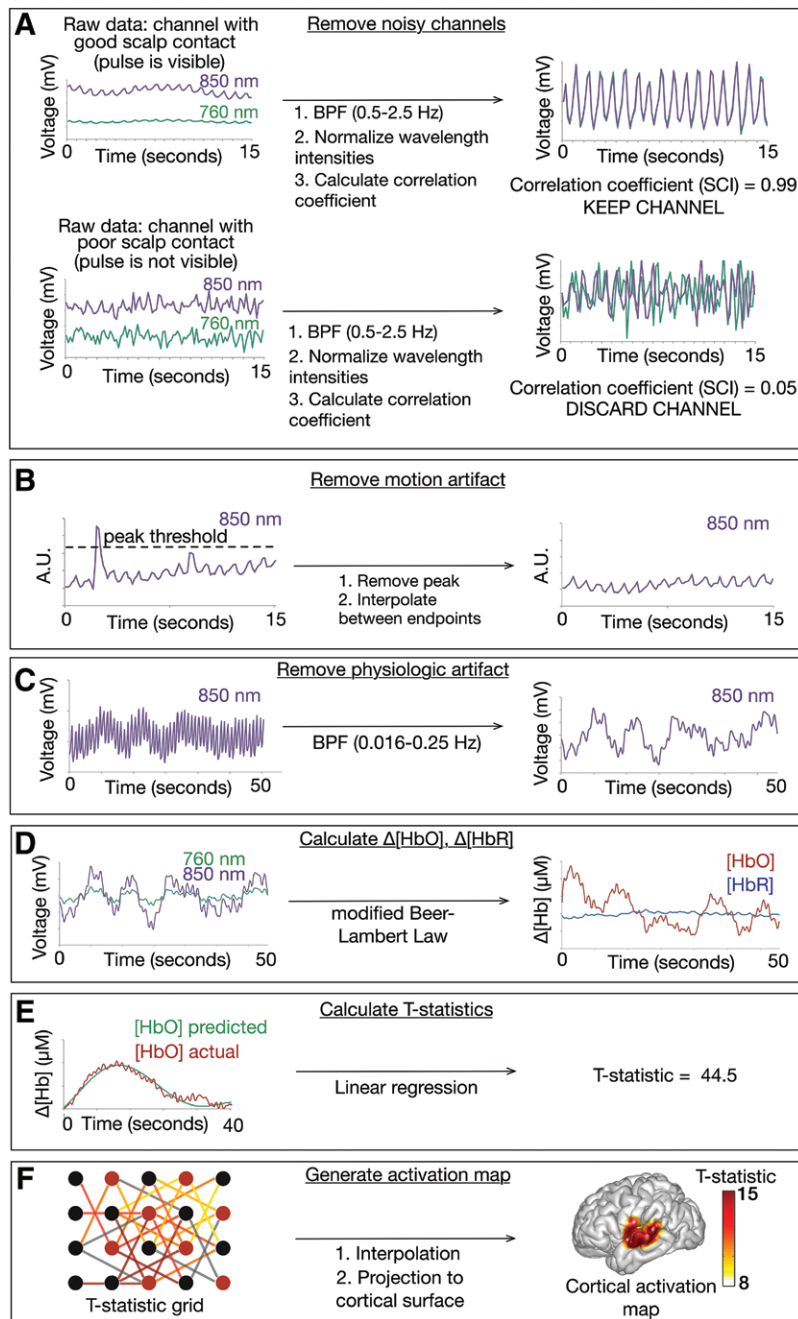
Fig. 3. Data filtering and analysis protocol. A, Removal of channels with excessive noise due to poor scalp contact. The goal was to only analyze channels with good cardiac signals, a characteristic of good scalp contact (Pollonini et al. 2014). Two channels of raw data are shown, one with good scalp contact (top) and another with poor scalp contact (bottom). Each channel had a band-pass filter (BPF) applied between 0.5 and 2.5 Hz to isolate the cardiac pulse (for heart rates between 30 and 150 beats/min). Next, the signals from each wavelength were normalized and the correlation coefficient (describing how similar the two wavelengths are to one another) was calculated. The similarity of the signals from each wavelength is reflected by the correlation coefficient, deemed the scalp contact index (SCI). Channels with an SCI ≥0.70 were considered "good" and were kept for further processing and analysis, while those with an SCI <0.70 were discarded. B, Motion artifacts (due to jaw movement, swallowing, or head movement) were removed from each channel. To identify the start and stop times of motion artifacts, each channel was band-pass filtered between 0.1 and 3.0 Hz to remove slow signal drift and normalized to the intensity of the highest peak of the entire time course. Peaks in the signal exceeding 20% of the maximum peak intensity (dotted line labeled "peak threshold") were identified as motion artifact. The start and stop times of the peaks were determined. We then went back to the raw data and removed the artifacts by linearly interpolating between the start and stop time points. C, After removing motion artifacts, physiologic artifacts (due to respiration and cardiac pulse) were removed from each channel by band-pass filtering the raw data with motion artifacts removed between 0.016 and 0.25 Hz. Data from only one wavelength is shown for simplicity. D, For each channel, the voltage data from both wavelengths were used to calculate Δ[HbO] and Δ[HbR] via the modified Beer–Lambert Law. E, The change in HbO and HbR for each channel was compared with the predicted curves via linear regression, resulting in T-statistics for each channel. A high T-statistic indicates a good fit between the actual and predicted hemoglobin curves. For this example, only HbO is shown for simplicity. F, T-statistics calculated for each channel were arranged in a grid representing their position in the source–detector array. To generate cortical activation maps from the sparse, yet overlapping, array of T-statistics, we performed two-dimensional interpolation and then projected the matrix onto a standard cortical surface for visualization. HbO indicates oxyhemoglobin; HbR, deoxyhemoglobin; A.U., arbitrary units.

Beer–Lambert law was used to calculate the relative concentrations of oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) (Fig. 3D) (Sassaroli & Fantini 2004). For each channel, all five trials of each stimulus type were block averaged, and linear regression was carried out between the block-averaged hemodynamic response and a predicted hemodynamic response. The predicted hemodynamic response was created in three steps. First, 50 responses from 26 normal-hearing individuals listening to normal speech were selected by eye as being obvious stimulus-evoked responses. Second, these responses were then normalized and averaged. Finally, the average response was smoothed by fitting it with a fifth-order polynomial that provided the best least-squares fit. We chose this approach to defining a predicted hemodynamic response rather than using a previously published fMRI blood-oxygenation-level-dependent (BOLD) response model because the popular fMRI models we reviewed did not accurately predict obvious fNIRS responses noted by eye. We presume this is related to the fact that fNIRS analysis does not have the spatial sensitivity of fMRI and so the responses we measure are spatiotemporally filtered along the optical path through the biological tissues. Finally, the regressions between the predicted response and the measured response resulted in the calculation of a T-statistic for each channel, quantifying the quality of fit between the predicted and actual hemodynamic response curves (Fig. 3E) (Friston et al. 1995; Monti 2011). The magnitudes of the hemodynamic responses (e.g., the beta or parametric values) were not analyzed as many unmeasurable factors vary between participants that affect these parameters (such as skull thickness, scalp contact, etc.). Therefore, we used the T-statistic because even though the relative changes have meaning, the absolute magnitudes do not (Sato et al. 2005).

Visualization and Calculation of Cortical Activation Areas • At this point in analysis process, each emitter–detector combination (channel) could be represented by a single number, which was the T-statistic representing the goodness of fit between the measured response and the predicted response. We then selected only those channels with a 3.0- to 3.3-cm interoptode distance to calculate the overall cortical activation in that participant for that stimulus. Because many of these channels crisscross each other, they provide spatial oversampling. We averaged the T-statistics from overlapping channels to reduce noise and then performed spline interpolation with a step size of 0.5 mm to translate the relatively sparse T-statistic distribution to a dense mesh of values suitable for visualization (Fig. 3F). A threshold for statistical significance was set, and the interpolated area with T-statistics exceeding the threshold was recorded as the activation area for each stimulus condition. We selected the threshold T-statistic to be 8 because we found that it effectively eliminated most activations in subjects who had their CI turned off, which presumably were artifacts. In addition, it enhanced the differences in brain activation to the different speech stimuli in normal-hearing subjects.

As opposed to simply plotting a 2D square of the fNIRS data (i.e., a topographic map), we chose to plot the representative responses we present in this article on a standard brain image to make them easier to visualize. A cortical surface was first generated in MATLAB from a GifTI formatted surface obtained from a representative adult's anatomical MRI data. The cortical surface was then smoothed by using a weighted spherical harmonic representation (Chung et al. 2008). The 2D interpolated T-statistic map was projected to the cortical surface, and the cortex surface was colored based on the projected T-statistic values. This strategy is reasonable given our previously published work demonstrating that the responses measured with fNIRS derive from this region of the brain (Sevy et al. 2010; Pollonini et al. 2014). However, it is important to note that this plotting scheme has no impact on the individual or group statistical analyses performed in this article.

Group-average activation maps were generated by comparing the list of T-statistics for all participants on a channel-by-channel basis versus no activation (a flat line with no activation during the stimulus presentation, i.e., the null hypothesis) using paired t tests. Only channels passing the SCI threshold were included in this analysis, and each stimulus was analyzed independently. Each analysis thus resulted in the generation of a 2D map of p values that indicated the statistically significant channels over the cohort of participants for that stimulus. Each map was then interpolated and projected to the cortical surface after a threshold for significance ($p < 0.05$) was set. In a similar manner, cohort difference maps were generated by comparing T-statistic on a channel-by-channel basis for each stimulus. Because in these analyses two different cohorts of participants were being compared, the unpaired t test was used rather than the paired t test. Otherwise, the analyses and plotting were performed in an identical manner.

Statistical Analysis

All representative examples were followed up with statistical analyses of averaged data. Unless otherwise stated, categorical variables are presented as frequencies and percentages, and continuous variables as means and SEs after ensuring that such variables had normal distributions. For comparisons with more than two groups, analysis of variance was first performed followed by the Student's t test with Bonferroni correction. The Mann–Whitney U test was used to compare data with skewed distributions when indicated. Paired t tests were used to assess the differences in cortical activation areas within groups, including the difference in activation area with CI on and off, and differences in cortical activation area for normal, channelized, and environmental sounds compared with scrambled speech for left and right HbO and HbR. Between groups, nonpaired t tests were used. All reported p values are two tailed, with a $p < 0.05$ indicating statistical significance when Bonferroni correction was not used (i.e., only one comparison was performed).

Multiple linear regression was used to characterize the association between word and speech perception scores and the normal:scrambled activation area ratio for HbO and HbR in both hemispheres. Age, side of CI implantation, duration of deafness, and time from implantation were included as covariates based on their relation to postimplantation performance in the literature (van Dijk et al. 1999; Gifford et al. 2008; Roditi et al. 2009; Budenz et al. 2011; Roberts et al. 2013). These variables were only included after univariate analysis of each variable yielded a p value of 0.25 or less (Bendel & Afifi 1977). Missing variables were eliminated from the analysis in a pairwise manner, and independent variables were assessed for colinearity before inclusion in the model, with a variance inflation factor of 10 being the threshold for unacceptable colinearity. Beta coefficients are reported for a 1-SD increase for independent variables. All statistical analyses were performed in SPSS (Version 21; IBM, Armonk, NY).

RESULTS

Our fNIRS imaging setup consisted of a headset holding an array of infrared light sources and detectors against the scalp overlying the LTL/STG (Fig. 4A, left) (Klem et al. 1999; Sevy et al. 2010). During a listening task, source–detector pairs (channels) measured increases in oxyhemoglobin concentration ([HbO]) and decreases in deoxyhemoglobin ([HbR]) that synchronize with the time course of the auditory stimulus (Fig. 4A, right). We compared the HbO and HbR curves for each channel to the predicted response using linear regression; only channels with a resulting T-statistic >8 were used to define the cortical activation area. This T-statistic threshold was chosen because after studying how well the raw data tracings fit the predicted curves well by eye. It excluded obviously noisy data without excluding data that fit the predicted response reasonably well. These channels were spatially interpolated according to the layout of the headset to construct a color-coded activation map, which was then projected onto a standardized three-dimensional cortical model to visualize the responses.

We performed behavioral measures of speech perception and fNIRS imaging in 35 normal-hearing adult controls and 32 adults with CIs. These tests included the monosyllabic CNC words (Peterson & Lehiste 1962) and the AzBio sentence recognition (Spahr et al. 2012) protocols. All normal-hearing control participants score 95 to 100% on these tests. However, the deaf participants hearing through a CI demonstrated wide variability in the behavioral tests (Fig. 4B and Table 3). CNC word scores ranged from 20 to 94% (mean 58% and SD 21%), which is consistent with the larger populations with mean CNC words scores in the range of 54 to 61% (Bassim et al. 2005; Gifford et al. 2008). AzBio sentence recognition scores ranged from 28 to 97% (mean 60% and SD 25%), which is also consistent with larger study populations with mean AzBio scores ranging from 58 to 72% (Gifford et al. 2008; Spahr et al. 2014; Massa & Ruckenstein 2014). We considered those participants with CNC words scores ≥70% and ≤40% (the top 10 and bottom 10 participants from our cohort) to have good and poor speech perception, respectively.
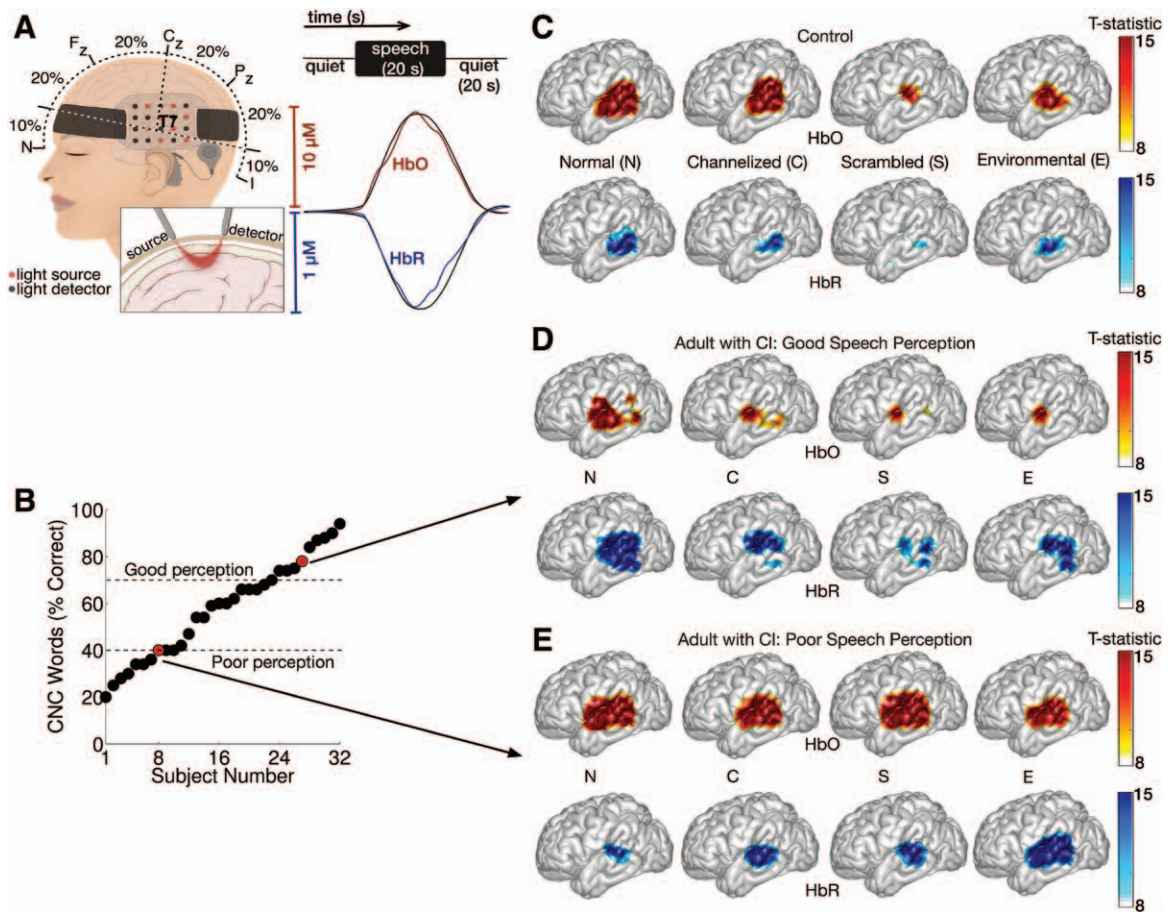

Fig. 4. Functional near-infrared spectroscopy experimental protocol and cortical hemodynamic responses. A, Left: Illustration of the optode array placement. The cochlear implants (CI)'s external processor is also shown. Nasion (N), inion (I), frontal zero (F$_z$), center zero (C$_z$), and parietal zero (P$_z$). Inset: near-infrared light emitted from the light source travels through skull and superficial tissues to reach the cortex. The amount of light that is backscattered and detected at the scalp's surface varies as the amount of light absorbed by oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) varies. Right: Representative [HbO] (red line) and [HbR] (blue line) tracings from one channel superimposed on the predicted hemodynamic response functions (black lines) in a normal-hearing adult in response to a 20-sec speech stimulus. B, Consonant-nucleus-consonant (CNC) word scores from the cohort of adults with CIs. Subcohorts of adults with CIs having good and poor speech perception are defined (dotted lines). C, Left cortical activation maps to the speech stimuli of different intelligibility levels and environmental sounds for a representative control participant. D, Cortical activation maps to the stimuli for a representative adult with CI with good speech perception (red dot). E, Cortical hemodynamic response maps to the stimuli for a representative adult with CI with poor speech perception (red dot).

Given the known responsiveness of the STG to syllables and natural speech in normal-hearing participants (Belin et al. 2000, 2002; Fecteau et al. 2004; Grodzinsky & Nelken 2014; Mesgarani et al. 2014), we expected that intelligible speech would evoke larger cortical responses than unintelligible speech in implanted adults. To test this, we used four previously validated acoustic stimuli that demonstrate this activation pattern within the LTL/STG of normal-hearing controls when measured by fMRI and fNIRS (Pollonini et al. 2014) (in order of intelligibility): normal speech, channelized speech, scrambled speech, and environmental sounds. Normal-hearing controls had strong oxyhemoglobin and deoxyhemoglobin cortical responses to normal speech (Fig. 4C; Movie S1, Supplemental Digital Content 5, http://links.lww.com/EANDH/A242). The response areas were about the same with channelized speech, but smaller responses were found with scrambled speech and environmental sounds. Cortical activation maps averaged over the normal-hearing controls confirmed this qualitative finding (Fig. 5).

As expected, implanted adults with good speech perception showed a cortical activation pattern similar to that of controls; as the speech stimuli became less intelligible, there was less cortical activation (Fig. 4D). However, participants with poor speech perception had similarly large areas of cortical activation for all four stimuli (Fig. 4E). Average difference maps were then generated between normal-hearing participants, adults with CIs with good speech perception, and adults with CIs with poor speech perception (Fig. 6). There were strong similarities between deaf adults hearing through a CI with good speech perception and normal-hearing controls. However, in deaf adults hearing through a CI, there was a fundamental difference between good and poor speech perceivers in that scrambled speech and environmental sounds created overly large responses within the LTL/STG of poor speech perceivers. This finding links the quality of speech perception to the amount of LTL/STG activation.

To further explore and quantify this association, we divided the activation area for each auditory stimulus by the activation area to scrambled speech, henceforth referred to as the "activation area ratio." This normalization reduced variability across participants caused by differences in scalp–brain distance and differences in the number of channels with adequate scalp contact (Sato et al. 2005; Cui et al. 2011). The activation area ratios were determined for each participant, for each hemisphere, and for both HbO and HbR. Some participants demonstrated no activation to scrambled speech. In these situations where calculating the activation area ratio would lead to a division by zero error, we used 1 cm^2 (the minimum detectable area of our experimental setup) as the area of scrambled speech. If a participant had no activation for both normal and scrambled speech, that participant was deemed a nonresponder for that hemoglobin species in that hemisphere (for instance, left HbO) and was not included in that particular statistical analysis.

In normal-hearing controls, the activation area ratio increased with increasing intelligibility of the speech stimulus (Fig. 7A). This was found for both HbO and HbR in both hemispheres. Implanted adults with good speech perception demonstrated a similar pattern, with increasing activation area ratios corresponding to increasing intelligibility of the speech stimulus. In contrast, implanted adults with poor speech perception had similar activation area ratios across all four stimuli. The activation area ratios for normal to scrambled speech (N:S) versus the behavioral measures of speech perception were plotted for each
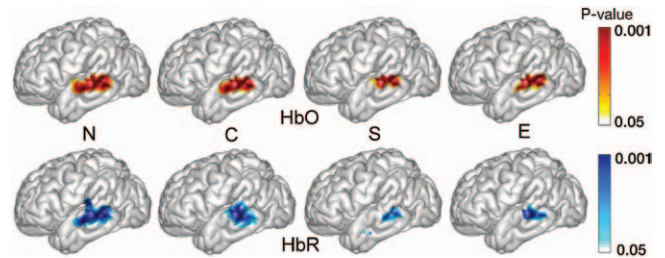

Fig. 5. Group-average activation maps for normal-hearing controls. Statistically significant activations across the cohort are shown for oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) for normal speech (N), channelized speech (C), scrambled speech (S), and environmental sounds (E), n = 35.

implanted participant who had these tests performed (Fig. 7B, C). Participants with higher CNC words and AzBio sentence test scores tended to have larger N:S activation areas.

There are many factors that are known to affect performance when hearing through a CI, for example, age correlates both with acoustical hearing abilities and with central processing abilities (Gates & Mills 2005; Humes et al. 2012; Jin et al. 2014; Park et al. 2015). Although there was a significant overlap between the ages of the participants in the implanted and normal-hearing cohorts (Table 1), their mean ages were different (Mann–Whitney U test, $p < 0.01$). To account for age and other factors as a potential covariates underlying the differences in cortical activation patterns, we performed multiple linear regression analyses using the N:S activation area ratio as the dependent variable. The independent factors included age, side of implantation, duration of deafness, time of implant, and either CNC words scores or AzBio sentence scores. Both the CNC words and AzBio scores independently correlated with the N:S activation area ratio for HbO and HbR in each hemisphere. The fact that significance was found across multiple
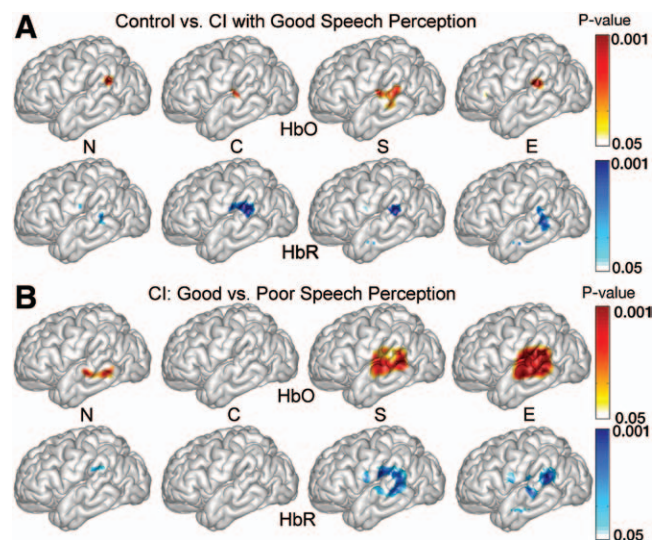

Fig. 6. Cortical activation difference maps for controls and adults with cochlear implants (CIs) with good and poor speech perception. A, When T-statistic maps were compared between controls and adults with CIs with good speech perception, there were only minimal differences in the activation areas for each stimulus. B, The cortical responses to scrambled speech and environmental sounds exhibited significantly different patterns between adults with CIs with good and poor speech perception.

TABLE 3. Behavioral measures of speech perception

	CI On		CI Off	
	Mean Score (SD)	n	Mean Score (SD)	n
SRT, dB HL	21 (4.5)	24	—	—
CNC words, %	56 (21)	29	16 (15)	28
AzBio, %	58 (20)	23	20 (19)	15

Depending on their clinical needs, not every test was performed in every participant.
CNC, consonant-nucleus-consonant; SRT, speech recognition threshold.

independent comparisons strongly supports the concept that cortical activation correlates with speech perception. While the duration of deafness was also found to be inversely related to the N:S activation area ratio for HbO and HbR in both hemispheres when AzBio was analyzed, but not for CNC scores (Tables 4 and 5). This finding is difficult to interpret but may represent an effect of the cortical reorganization that can occur in association with hearing loss (Campbell & Sharma 2014). Most importantly, however, none of the other factors (age in particular) were related to the differences in cortical activation patterns.

To exclude the possibility that differences in the activation area ratio were linked to differences in auditory thresholds, the activation area ratio was plotted against the SRT (Fig. 8). Significant correlations were not observed, indicating that the N:S activation area ratio was not simply providing an assessment of a participant's perception of sound intensity. Lastly, we also performed three more multiple linear regression analyses using the SRT, CNC words, and AzBio scores as dependent variables and age, side of implantation, duration of deafness, and time of implant as dependent variables. Both age and time of implant correlated with SRT (standardized beta = 0.802 and −0.228,

respectively, $p < 0.001$ for both). This makes sense because our CI programing strategy is to gradually increase the power (and thus lower the auditory threshold) over time as patients acclimate to using the device. However, there was not a significant relationship between CNC words or AzBio scores and any of the independent variables. Thus, these data rule out the most likely alternative causes of the cortical responses we measured. Therefore, taken together, these data indicate that the LTL/STG activation pattern provides an objective measure of speech perception by a deaf participant listening through a CI.

To determine the impact of reduced auditory stimulation, we repeated the measures after turning off the CI. Most of the implanted participants had a unilateral CI and the other ear was left unaided for fNIRS testing. This meant that some degree of residual acoustical hearing persisted in the unimplanted ears and perhaps even at minimal level in the implanted ears. Behavioral measures of speech perception revealed speech understanding in many participants even with the CI turned off although the results significantly improved when the CI was turned on (Fig. 9A, B and Table 3). Deaf adults demonstrated strong cortical responses when their CI was on, but only minimal responses with the CI turned off (Fig. 9C). On average, the cortical activation area in response to speech was greatest for controls, followed by implanted adults with their CI on; implanted adults with their CI off had the smallest areas of cortical activation (Fig. 9D). These responses were bilateral for all conditions. As a reference, the cortical surface area of the STG has been measured to be just slightly larger than what we found in normal-hearing adults (left: $19.83 \pm 3.52\,cm^2$; right: $17.28 \pm 2.62\,cm^2$; Meyer et al. 2013). In general, patients with residual hearing maintained small cortical responses to speech when the implant was turned off, whereas those without residual hearing did not. The N:S activation area ratio dropped when the CI was turned off, consistent with the drop in speech perception (Fig. 9E).
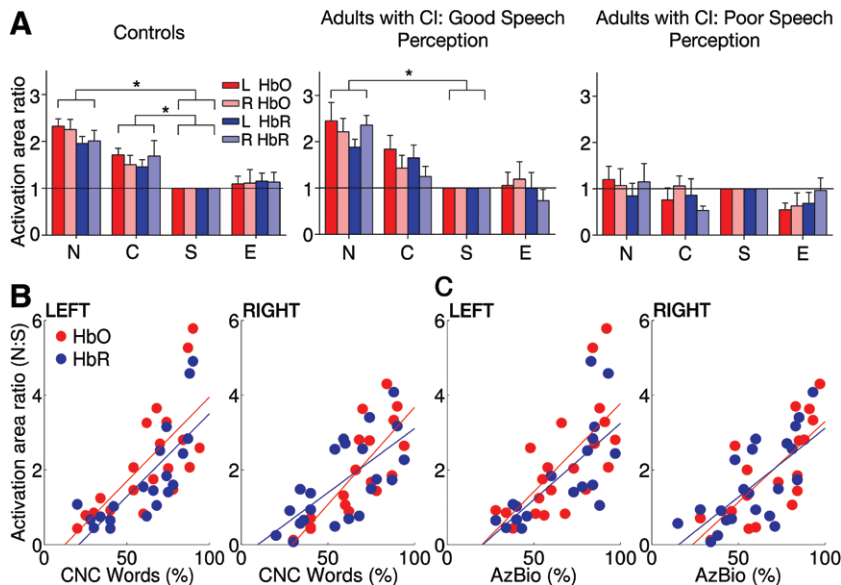

Fig. 7. Cortical activation correlates with behavioral measures of speech perception. A, Activation areas ratios were normalized to the activation area for scrambled speech (represented by solid black line at activation area ratio = 1). Normal speech (N), channelized speech (C), scrambled speech (S), and environmental sounds (E). Controls (n = 35); cochlear implant (CI) with good speech perception (n = 10); CI with poor speech perception (n = 10). Error bars represent the standard error of the mean. *$p < 0.01$. B, The normal:scrambled (N:S) activation area ratio vs. consonant-nucleus-consonant (CNC) word scores for both hemispheres in adults with their CIs turned on (n = 32). C, The normal:scrambled (N:S) activation area ratio vs. AzBio scores for each hemisphere in adults with their CIs turned on (n = 23). Error bars represent the standard error of the mean. HbO indicates oxyhemoglobin; HbR, deoxyhemoglobin.

TABLE 4. Multiple linear regression analysis demonstrates that the CNC words score correlates with the N:S activation area ratio

	L HbO		R HbO		L HbR		R HbR	
	Standardized β	*p*	Standardized β	*p*	Standardized β	*p*	Standardized β	*p*
Age (yr)	−0.08	0.76	−0.37	0.20	−0.18	0.59	−0.05	0.88
CI side (R = 0, L = 1, both = 2)	−0.08	0.78	−0.02	0.94	0.12	0.74	0.30	0.38
Duration of deafness (yr)	−0.26	0.11	−0.27	0.11	−0.22	0.26	−0.30	0.11
Time from implant (mo)	0.20	0.13	−0.02	0.86	−0.21	0.21	0.00	0.99
CNC words (CI on)	1.23	*<0.01*	1.33	*<0.01*	1.06	*<0.01*	0.76	*0.01*
R²	0.68		0.66		0.53		0.57	
F	13.37	*<0.01*	12.44	*<0.01*	7.58	*<0.01*	8.73	*<0.01*

Standardized β is the β-coefficient for 1-SD increase in CNC words. Significant values (p ≤ 0.05) indicated by italic print.
CI, cochlear implant; CNC, consonant-nucleus-consonant; HbO, oxyhemoglobin; HbR, deoxyhemoglobin; L, left; N:S, normal:scrambled; R, right.

Together, these data confirm that the CI produced the responses measured by fNIRS, further supporting the notion that the cortical activation pattern correlates with speech perception.

DISCUSSION

The artificial electrical stimulation produced by a modern CI provides less information about the auditory environment than that provided by a normal-functioning cochlea. The fact that many deaf patients hearing through a CI can decipher speech is a testament to the ability of the human brain to adapt and to process poor-quality sensory information (Clark 2015). However, this process does not work for every implanted patient, and we currently have no clinical test to explain the pathophysiology underlying the difficulty that some patients have with speech understanding. Here, we show that activation patterns within the LTL/STG in the implanted population correlate with behavioral measures of speech perception. Furthermore, we have found that the neural basis of poor speech perception in implanted participants is not that sound information does not reach the LTL/STG but that these participants are unable to discriminate speech from the signal that does reach it.

While many neuroimaging methods exist, they have been limited to a research setting in the implanted population due to a lack of CI compatibility and suitability for routine use in the outpatient setting. For our purposes, some of these issues can be overcome using fNIRS. For example, fNIRS has several advantages over fMRI in that it is not affected by the magnet within the CI and it is easy to use in a routine clinical environment because it is small and portable. However, it has a lower spatial resolution than fMRI (approximately 5 to 10 mm versus

approximately 1 mm) (Cui et al. 2011). A benefit over PET scan studies is that it does not utilize ionizing radiation and thus can be used safely for repeated testing over multiple clinic visits. A benefit over EEG is that long-speech stimuli can be used, as opposed to the brief stimulus pulses designed to isolate the cortical response from the electrical artifact of the CI in time. Thus, the stimuli reflect normal listening situations. Lastly, fNIRS does not require a lengthy setup or experimental protocol; the time between when a participant entered and left the testing room was 20 to 30 min for these experiments.

Despite these positive attributes, fNIRS is not without its drawbacks. fNIRS is limited to imaging superficial regions of cortex, as the light cannot penetrate the tissue to measure deeper structures. In addition, scalp thickness has a significant impact on the ability of fNIRS to image cortical activity (Beauchamp et al. 2011; Cui et al. 2011). There is no doubt that fNIRS has poor spatial localization when compared against fMRI. In addition, the exact position of the optodes certainly might vary slightly (<1 cm between participants). Because of these limitations, we have been extremely conservative in analyzing the responses we measured. First, we have been careful not make any claims regarding the spatial pattern of cortical activations. Second, we calculated the total area of activation to several different stimuli and then normalized the responses of each participant individually. In this way, each participant could serve as his/her own control. In future studies, the spatial patterns of activation could be explored by localizing optode positions using the preimplant MRI most patients receive as part of their medical evaluation of hearing loss and modeling the light travel paths (Cristobal & Oghalai 2008; Lin et al. 2010, 2011; Jerry

TABLE 5. Multiple linear regression analysis demonstrates that the AzBio score and the duration of deafness independently correlate with the N:S activation area ratio

	L HbO		R HbO		L HbR		R HbR	
	Standardized β	*p*	Standardized β	*p*	Standardized β	*p*	Standardized β	*p*
Age (yr)	−0.12	0.77	−0.49	0.22	−0.43	0.28	−0.32	0.38
CI side (R = 0, L = 1, both = 2)	0.18	0.65	0.23	0.55	0.25	0.50	0.36	0.30
Duration of deafness (yr)	−0.49	*0.05*	−0.55	*0.03*	−0.50	*0.04*	−0.54	*0.02*
Time from implant (mo)	−0.24	0.20	−0.07	0.67	−0.26	0.16	−0.04	0.79
AzBio (CI on)	1.19	*<0.01*	1.44	*<0.01*	1.42	*<0.01*	1.19	*<0.01*
R²	0.55		0.59		0.59		0.66	
	6.12	*<0.01*	6.99	*<0.01*	7.16	*<0.01*	9.00	*<0.01*

Standardized β is the β for 1 SD increase in consonant-nucleus-consonant words. Significant values (p ≤ 0.05) indicated by italic print.
CI, cochlear implant; HbO, oxyhemoglobin; HbR, deoxyhemoglobin; L, left; N:S, normal:scrambled; R, right.
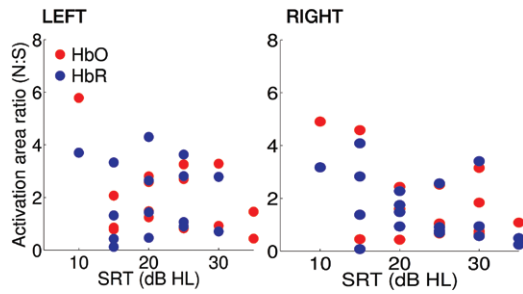
Fig. 8. Cortical activation patterns do not correlate with auditory thresholds. The normal:scrambled (N:S) activation area ratio vs. the speech reception threshold (SRT) for oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) in left and right hemispheres in adults with their cochlear implants turned on (n = 26). $p > 0.10$ for HbO and HbR in both hemispheres.

& Oghalai 2011; Oghalai et al. 2012). Despite these limitations, fNIRS is a powerful technique that is well suited for use in an outpatient clinical setting and could be easily implemented in clinical studies of CI function.

Our finding of larger cortical responses to speech than environmental sounds in controls and implanted adults with good speech perception aligns well with previously published fNIRS and fMRI data (Belin et al. 2002; Fecteau et al. 2004; Pollonini et al. 2014). Vocoded speech is almost entirely intelligible to adults with normal hearing, which explains the similar areas of cortical response to normal and channelized stimuli we found in the control cohort (Shannon et al. 1995). Our finding that implanted adults with high speech perception scores exhibit a pattern of cortical activation similar to controls is consistent with PET data that the degree of cortical response to speech correlates with the quality of speech perception among adults with CIs and that there is a stronger cortical response to speech

than to white noise and environmental sounds in adults with CIs and in normal-hearing adults (Naito et al. 2000; Green et al. 2005).

Implanted adults with low speech perception scores demonstrated large cortical activation areas to all stimuli, without preferential responses to speech stimuli. Our studies are unable to differentiate where in the auditory pathway this abnormal pattern of responses is initiated. It could begin as early as the auditory nerve, due to less-than-optimal stimulation of the auditory nerve fibers by the electrical contacts along the CI electrode. Alternatively, it could occur at a later stage in the auditory processing hierarchy, with neurons failing to process the key characteristics of speech (frequency, intensity, and timing) necessary to render speech intelligible. Consistent with our results, abnormal electrically evoked cortical potentials have been found in implanted children with poor speech perception despite the ability to hear sound (Gordon et al. 2005). Our findings suggest that humans may attempt to compensate for poor-quality sensory information by increasing the activity within the LTL/STG.

This may be due to an increased listening effort. For example, EEG studies have demonstrated that the differences in task difficulty, for example, with sound localization, can alter the amount of effort that CI users invest in performing the tasks (Senkowski et al. 2014). Thus, the attention of the subject to the stimulus alters his/her cortical response. In addition, hearing-impaired individuals have increased listening effort to noise with no meaningful content compared with normal-hearing controls (Larsby et al. 2005; Zekveld et al. 2010). Correspondingly, the listening effort for CI recipients and normal-hearing controls is similar if the two groups have similar word recognition scores (Hughes & Galvin 2013). Our findings of larger areas of brain activation to scrambled speech in CI users with
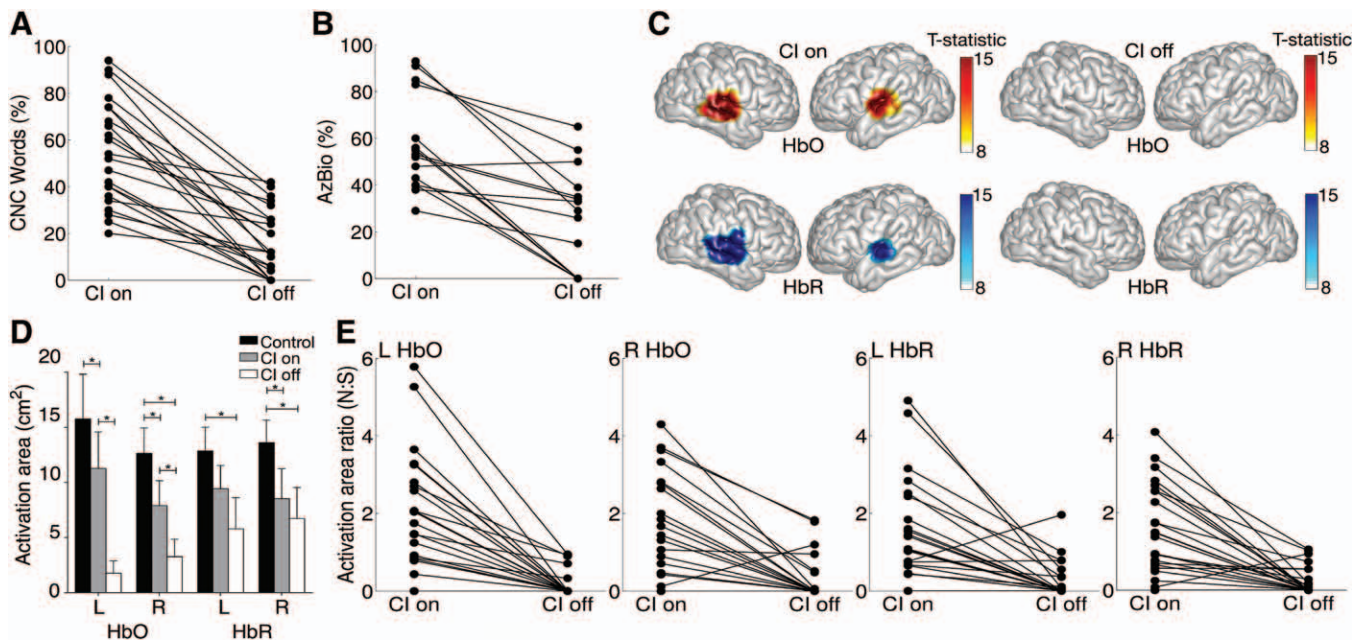

Fig. 9. Cortical responses with cochlear implant (CI) on and off. A, Consonant-nucleus-consonant (CNC) word scores with CI on and off. B, AzBio sentence scores with CI on and off. C, Cortical hemodynamic response maps for a representative adult with a CI listening to normal speech with the device turned on and turned off. D, Average areas of cortical activation to normal speech among normal-hearing controls and deaf adults with CIs with their devices on and off. Error bars represent the standard error of the mean. *$p < 0.01$. E, The normal:scrambled (N:S) activation area ratios decreased upon turning off the CI ($p \leq 0.01$ for each subplot, two-tailed paired t test). HbO indicates oxyhemoglobin; HbR, deoxyhemoglobin.

poor word recognition scores strongly support these studies. It is even possible that the increased brain activity we measured in poor CI performers is the neural correlate of the increased listening effort.

One caveat that should be considered, however, is that the attention of the participant was not quantified or controlled for in our experiments. It is possible that participant attention may modulate the response patterns within the LTL/STG, i.e., perhaps those participants who pay more attention to sound have larger differential responses and also were able to learn to use their CI more effectively. Thus, this study only demonstrates a correlation between cortical responses and behavioral responses, not a causal relationship. Further studies are needed to determine whether imaging with fNIRS can be used as a prognostic indicator after implantation and possibly even to aid CI programing efforts. In addition, the role of contralateral and ipsilateral cortical activation needs to be studied, particularly when consideration of the risks and benefits of unilateral versus bilateral implantation in young children, given our increasing understanding of the consequences of asymmetric hearing during development (Gordon et al. 2015).

REFERENCES

American Speech-Language-Hearing Association. (1988). *Determining Threshold Level for Speech*, 30, 85–89.

Bassim, M. K., Buss, E., Clark, M. S., et al. (2005). MED-EL Combi40+ cochlear implantation in adults. *Laryngoscope*, 115, 1568–1573.

Beauchamp, M. S., Beurlot, M. R., Fava, E., et al. (2011). The developmental trajectory of brain-scalp distance from birth through childhood: Implications for functional neuroimaging. *PLoS One*, 6, e24981.

Belin, P., Fecteau, S., Bédard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends Cogn Sci*, 8, 129–135.

Belin, P., Zatorre, R. J., Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Brain Res Cogn Brain Res*, 13, 17–26.

Belin, P., Zatorre, R. J., Lafaille, P., et al. (2000). Voice-selective areas in human auditory cortex. *Nature*, 403, 309–312.

Bendel, R., & Afifi, A. (1977). Comparison of stopping rules in forward "stepwise" regression. *J Am Stat Assoc 72*:46–53.

Bortfeld, H., Wruck, E., Boas, D. A. (2007). Assessing infants' cortical response to speech using near-infrared spectroscopy. *Neuroimage*, 34, 407–415.

Budenz, C. L., Cosetti, M. K., Coelho, D. H., et al. (2011). The effects of cochlear implantation on speech perception in older adults. *J Am Geriatr Soc*, 59, 446–453.

Campbell, J., & Sharma, A. (2014). Cross-modal re-organization in adults with early stage hearing loss. *PLoS One*, 9, e90594.

Chung, M., Hartley, R., Dalton, K. M., et al. (2008) Encoding cortical surface by spherical harmonics. *Stat Sin*, 18, 1269–1291.

Clark, G. M. (2015). The multi-channel cochlear implant: Multi-disciplinary development of electrical stimulation of the cochlea and the resulting clinical benefit. *Hear Res*, 322, 4–13.

Cristobal, R., & Oghalai, J. S. (2008). Hearing loss in children with very low birth weight: Current review of epidemiology and pathophysiology. *Arch Dis Child Fetal Neonatal Ed*, 93, F462–F468.

Cui, X., Bray, S., Bryant, D. M., et al. (2011). A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. *Neuroimage*, 54, 2808–2821.

Dehghani, H., Brooksby, B., Vishwanath, K., et al. (2003). The effects of internal refractive index variation in near-infrared optical tomography: A finite element modelling approach. *Phys Med Biol*, 48, 2713–2727.

DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proc Natl Acad Sci U S A*, 109, E505–E514.

Fecteau, S., Armony, J. L., Joanette, Y., et al. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study. *Neuroimage*, 23, 840–848.

Firszt, J. B., Chambers And, R. D., Kraus, N. (2002). Neurophysiology of cochlear implant users II: Comparison among speech perception, dynamic range, and physiological measures. *Ear Hear*, 23, 516–531.

Friston, K. J., Holmes, A. P., Poline, J. B., et al. (1995). Analysis of fMRI time-series revisited. *Neuroimage*, 2, 45–53.

Gates, G., Daly, K., Dichtel, W., et al. (1995) Cochlear implants in adults and children. *JAMA*, 274, 1955–1961.

Gates, G. A., & Mills, J. H. (2005). Presbycusis. *Lancet*, 366, 1111–1120.

Gfeller, K., Turner, C., Mehr, M., et al. (2002). Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implants Int*, 3, 29–53.

Gifford, R. H., Shallop, J. K., Peterson, A. M. (2008). Speech recognition materials and ceiling effects: Considerations for cochlear implant programs. *Audiol Neurootol*, 13, 193–205.

Gordon, K., Henkin, Y., Kral, A. (2015). Asymmetric hearing during development: The aural preference syndrome and treatment options. *Pediatrics*, 136, 141–153.

Gordon, K. A., Tanaka, S., Papsin, B. C. (2005). Atypical cortical responses underlie poor speech perception in children using cochlear implants. *Neuroreport*, 16, 2041–2045.

Green, K. M., Julyan, P. J., Hastings, D. L., et al. (2005). Auditory cortical activation and speech perception in cochlear implant users: Effects of implant experience and duration of deafness. *Hear Res*, 205, 184–192.

Grodzinsky, Y., & Nelken, I. (2014). The neural code that makes us human. *Science*, 343, 978–979.

Hall, D. A., & Plack, C. J. (2009). Pitch processing sites in the human auditory brain. *Cereb Cortex*, 19, 576–585.

Holden, L. K., Finley, C. C., Firszt, J. B., et al. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear Hear*, 34, 342–360.

Hughes, K. C., & Galvin, K. L. (2013). Measuring listening effort expended by adolescents and young adults with unilateral or bilateral cochlear implants or normal hearing. *Cochlear Implants Int*, 14, 121–129.

Humes, L. E., Dubno, J. R., Gordon-Salant, S., et al. (2012). Central presbycusis: A review and evaluation of the evidence. *J Am Acad Audiol*, 23, 635–666.

Humphries, C., Liebenthal, E., Binder, J. R. (2010). Tonotopic organization of human auditory cortex. *Neuroimage*, 50, 1202–1211.

Huppert, T. J., Hoge, R. D., Diamond, S. G., et al. (2006). A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *Neuroimage*, 29, 368–382.

Jerry, J., & Oghalai, J. S. (2011). Towards an etiologic diagnosis: Assessing the patient with hearing loss. *Adv Otorhinolaryngol*, 70, 28–36.

Jin, S. H., Liu, C., Sladen, D. P. (2014). The effects of aging on speech perception in noise: Comparison between normal-hearing and cochlear-implant listeners. *J Am Acad Audiol*, 25, 656–665.

Klem, G. H., Lüders, H. O., Jasper, H. H., et al. (1999). The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl*, 52, 3–6.

Lakatos, P., Musacchia, G., O'Connel, M. N., et al. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron*, 77, 750–761.

Larsby, B., Hällgren, M., Lyxell, B., et al. (2005). Cognitive performance and perceived effort in speech processing tasks: Effects of different noise

backgrounds in normal-hearing and hearing-impaired subjects. *Int J Audiol*, *44*, 131–143.

Lin, J. W., Chowdhury, N., Mody, A., et al. (2011). Comprehensive diagnostic battery for evaluating sensorineural hearing loss in children. *Otol Neurotol*, *32*, 259–264.

Lin, J. W., Mody, A., Tonini, R., et al. (2010). Characteristics of malfunctioning channels in pediatric cochlear implants. *Laryngoscope*, *120*, 399–404.

Lloyd-Fox, S., Papademetriou, M., Darboe, M. K., et al. (2014). Functional near infrared spectroscopy (fNIRS) to assess cognitive function in infants in rural Africa. *Sci Rep*, *4*, 4740.

Lu, T., Huang, J., Zeng, F.-G. (2014). Accurate guitar tuning by cochlear implant musicians. *PLoS One*, *9*, e92454.

Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., et al. (2013). Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proc Natl Acad Sci U S A*, *110*, 4846–4851.

Massa, S. T., & Ruckenstein, M. J. (2014). Comparing the performance plateau in adult cochlear implant patients using HINT and AzBio. *Otol Neurotol*, *35*, 598–604.

McLaughlin, M., Lopez Valdes, A., Reilly, R. B., et al. (2013). Cochlear implant artifact attenuation in late auditory evoked potentials: A single channel approach. *Hear Res*, *302*, 84–95.

Mesgarani, N., Cheung, C., Johnson, K., et al. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science*, *343*, 1006–1010.

Meyer, M., Liem, F., Hirsiger, S., et al. (2013). Cortical surface area and cortical thickness demonstrate differential structural asymmetry in auditory-related areas of the human cortex. *Cereb Cortex*, *3*, 1–12

Miller, S., & Zhang, Y. (2014). Validation of the cochlear implant artifact correction tool for auditory electrophysiology. *Neurosci Lett*, *577*, 51–55.

Monti, M. M. (2011). Statistical analysis of fMRI time-series: A critical review of the GLM approach. *Front Hum Neurosci*, *18*, 5–28.

Naito, Y., Tateya, I., Fujiki, N., et al. (2000). Increased cortical activation during hearing of speech in cochlear implant users. *Hear Res*, *143*, 139–146.

Oghalai, J. S., Caudle, S. E., Bentley, B., et al. (2012). Cognitive outcomes and familial stress after cochlear implantation in deaf children with and without developmental delays. *Otol Neurotol*, *33*, 947–956.

Park, M.-H., Won, J. H., Horn, D. L., et al. (2015). Acoustic temporal modulation detection in normal-hearing and cochlear implanted listeners: Effects of hearing mechanism and development. *J Assoc Res Otolaryngol*, *16*, 389–399.

Pasley, B. N., David, S. V., Mesgarani, N., et al. (2012). Reconstructing speech from human auditory cortex. *PLoS Biol*, *10*, e1001251.

Perez, M., & Macias, R. (2004). Criteria of candidacy for unilateral cochlear implantation in postlingually deafened adults III: Prospective evaluation of an actuarial approach to defining a criterion. *Ear Hear*, *25*, 361–374.

Peterson, G. E., & Lehiste, I. (1962). Revised CNC lists for auditory tests. *J Speech Hear Disord*, *27*, 62–70.

Pollonini, L., Olds, C., Abaya, H., et al. (2014). Auditory cortex activation to natural speech and simulated cochlear implant speech measured with functional near-infrared spectroscopy. *Hear Res*, *309*, 84–93.

Roberts, D. S., Lin, H. W., Herrmann, B. S., et al. (2013). Differential cochlear implant outcomes in older adults. *Laryngoscope*, *123*, 1952–1956.

Roditi, R. E., Poissant, S. F., Bero, E. M., et al. (2009). A predictive model of cochlear implant performance in postlingually deafened adults. *Otol Neurotol*, *30*, 449–454.

Sassaroli, A., & Fantini, S. (2004). Comment on the modified Beer-Lambert law for scattering media. *Phys Med Biol*, *49*, N255–N257.

Sato, H., Fuchino, Y., Kiguchi, M., et al. (2005). Intersubject variability of near-infrared spectroscopy signals during sensorimotor cortex activation. *J Biomed Opt*, *10*, 44001.

Scholkmann, F., Kleiser, S., Metz, A. J., et al. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage*, *85*(Pt 1), 6–27.

Senkowski, D., Pomper, U., Fitzner, I., et al. (2014). Beta-band activity in auditory pathways reflects speech localization and recognition in bilateral cochlear implant users. *Hum Brain Mapp*, *35*, 3107–3121.

Sevy, A. B., Bortfeld, H., Huppert, T. J., et al. (2010). Neuroimaging with near-infrared spectroscopy demonstrates speech-evoked activity in the auditory cortex of deaf children following cochlear implantation. *Hear Res*, *270*, 39–47.

Shannon, R. V., Zeng, F. G., Kamath, V., et al. (1995). Speech recognition with primarily temporal cues. *Science*, *270*, 303–304.

Smalt, C. J., Gonzalez-Castillo, J., Talavage, T. M., et al. (2013). Neural correlates of adaptation in freely-moving normal hearing subjects under cochlear implant acoustic simulations. *Neuroimage*, *82*, 500–509.

Spahr, A. J., Dorman, M. F., Litvak, L. M., et al. (2014). Development and validation of the pediatric AzBio sentence lists. *Ear Hear*, *35*, 418–422.

Spahr, A. J., Dorman, M. F., Litvak, L. M., et al. (2012). Development and validation of the AzBio sentence lists. *Ear Hear*, *33*, 112–117.

Steinschneider, M., Nourski, K. V., Rhone, A. E., et al. (2014). Differential activation of human core, non-core and auditory-related cortex during speech categorization tasks as revealed by intracranial recordings. *Front Neurosci*, *8*, 240.

van Dijk, J. E., van Olphen, A. F., Langereis, M. C., et al. (1999). Predictors of cochlear implant performance. *Audiology*, *38*, 109–116.

Villringer, A., Planck, J., Hock, C., et al. (1993). Near infrared spectroscopy (NIRS): A new tool to study hemodynamic changes during activation of brain function in human adults. *Neurosci Lett*, *154*, 101–104.

Viola, F. C., De Vos, M., Hine, J., et al. (2012). Semi-automatic attenuation of cochlear implant artifacts for the evaluation of late auditory evoked potentials. *Hear Res*, *284*, 6–15.

Wang, X. (2000). On cortical coding of vocal communication sounds in primates. *Proc Natl Acad Sci U S A*, *97*, 11843–11849.

Zaramella, P., Freato, F., Amigoni, A., et al. (2001). Brain auditory activation measured by near-infrared spectroscopy (NIRS) in neonates. *Pediatr Res*, *49*, 213–219.

Zekveld, A. A., Kramer, S. E., Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear Hear*, *31*, 480–490.

Zhang, F., Anderson, J., Samy, R., et al. (2010). The adaptive pattern of the late auditory evoked potential elicited by repeated stimuli in cochlear implant users. *Int J Audiol*, *49*, 277–285.

Zhang, F., Hammer, T., Banks, H. L., et al. (2011). Mismatch negativity and adaptation measures of the late auditory evoked potential in cochlear implant users. *Hear Res*, *275*, 17–29.