

# NIH Public Access

**Author Manuscript**

Neuroimage. Author manuscript; available in PMC 2014 November 15.

Published in final edited form as:

Neuroimage. 2013 November 15; 82: 500–509. doi:10.1016/j.neuroimage.2013.06.001.

## **NEURAL CORRELATES OF ADAPTATION IN NORMAL HEARING SUBJECTS TO FREE LEARNING WITH COCHLEAR IMPLANT ACOUSTIC SIMULATIONS**

**Christopher J. Smalt**1, **Javier Gonzalez-Castillo**2, **Thomas M. Talavage**1, **David B. Pisoni**3, and **Mario A. Svirsky**<sup>4</sup>

<sup>1</sup>School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN

<sup>2</sup>Section on Functional Imaging Methods, Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda, MD

<sup>3</sup>Department of Psychology, Indiana University, Bloomington, IN

<sup>4</sup>Department of Otolaryngology, New York University School of Medicine, New York, NY

### **Abstract**

Neurobiological correlates of adaptation to spectrally degraded speech were investigated with fMRI before and after exposure to a portable real-time speech processor that implements an acoustic simulation model of a cochlear implant (CI). The speech processor, in conjunction with isolating insert earphones and a microphone to capture environment sounds, was worn by participants over a two week chronic exposure period. fMRI and behavioral speech comprehension testing was conducted before and after this two week period. After using the simulator each day for 2 hours, participants significantly improved in word and sentence recognition scores. fMRI shows that these improvements came accompanied by changes in patterns of neuronal activation. In particular, we found additional recruitment of visual, motor, and working memory areas after the perceptual training period. These findings suggest that the human brain is able to adapt in a short period of time to a degraded auditory signal under a natural learning environment, and gives insight on how a CI might interact with the central nervous system. This paradigm can be furthered to investigate neural correlates of new rehabilitation, training, and signal processing strategies non-invasively in normal hearing listeners to improve CI patient outcomes.

### **1. INTRODUCTION**

Cochlear implants (CIs) can restore hearing to profoundly deaf individuals through direct stimulation of auditory neurons in the VIII cranial nerve. Patient performance on language tasks after implantation varies substantially, both in terms of asymptotic performance level, and in terms of the speed with which this level is achieved (Bassim et al., 2005). In the immediate postoperative period, some patients exhibit poor performance at speech comprehension and are sometimes only capable of detecting the presence of sounds. Improvement of CI user performance typically involves months of rehabilitation (sometimes

<sup>© 2013</sup> Elsevier Inc. All rights reserved.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

as long as one year) that culminates in asymptotic performance levels that vary among different patients; even those who share a common model of implant and electrical stimulation strategy (Dorman, 1993).

Cochlear implants use a microelectrode array surgically implanted into the inner ear, which results in a fixed number of sites where the nerve population can be effectively stimulated see (Wilson, 2000) for a description of basic functioning principles of CIs. This results in limited spectral resolution for CI users; with typical performance shown to be equivalent to having 8 spectral bands or channels of information (Shannon et al., 2004). Studies that utilize neuroimaging techniques have demonstrated that the processing of acoustic signals in the brain differs between CI users and normal hearing subjects. Some of the first positron emission tomography (PET) studies proved that auditory stimulation with CIs produces bilateral activations in the primary and secondary auditory cortex (Herzog et al., 1991). Subsequent studies found that activations in these areas differed between normal hearing subjects and different groups of CI users (Giraud et al., 2001a; Giraud et al., 2000; Naito et al., 1997; Wong et al., 1999). Moreover, differences were also found in the inferior frontal cortex (Wong et al., 1999), as well as in regions outside classical language areas such as dorsal occipital cortex (Giraud et al., 2001a), precuneus (Giraud et al., 2001a) and subcortical auditory regions (Giraud et al., 2000). Additionally, studies of speech perception with CI users have reported recruitment of cortical regions such as visual cortex (Giraud et al., 2001b; Green et al., 2008; Kang et al., 2004) or the posterior cingulate (Kang et al., 2004), regions which are not usually reported with normal hearing subjects. These combined data provide evidence that CI users and normal hearing subjects use differing neural strategies to achieve speech comprehension. However, in these studies it is difficult to distinguish if these neuronal differences result only from adaptation to processing the degraded signal or from other factors such as recovery from long periods of deafness or maturation—the latter being a particular confound in the case of studies involving young children. In addition, each CI user has a unique deafness, with the onset, mechanism, and nature of deafness being patient-specific. Therefore, it is impractical to control the perception of external stimuli across individual CI users. The second limitation is that almost all CI devices are incompatible with functional magnetic resonance imaging (fMRI), the most powerful of the present ensemble of functional neuroimaging technologies with respect to localization and tracking of changes over time. Therefore, it is impractical to study this intuitively desirable population for the purposes envisioned.

Rather, the larger population of normally hearing listeners represents the ideal target population. Acoustic simulation models of CI's have become a common means to test processing strategies (Shannon et al., 2004) and evaluate training (Harnsberger et al., 2001; Li and Fu, 2007). These models manipulate the acoustic signal in the same way as a typical CI, simulating for a normal hearing listener what a real CI user might hear. Davis and Johnsrude (2003) and Narain et al. 2003 observed a left-lateralized response to vocoded speech primarily in the superior temporal gyrus, inferior frontal and premotor as compared to normal speech. Other studies looked at adaptation specifically (e.g., Eisner et al. 2010, Hervais-Adelman et al. 2012, Erb et al. 2012). Eisner et al. (2010) observed the effect of short term training in the scanner by interleaving blocks of training and testing with vocoded sentences, finding that the left inferior frontal gyrus (IFG) and left superior temporal sulcus (STS) were sensitive to the learnability of the sentences, while the left angular gyrus (AG) reflected the most significant activation directly related to the learning process. Hervais-Adelman et al. (2012) attempted to exclude the confound of sentence level processing by focusing on single words, and reported differences between clear speech and vocoded speech in the left insula, and prefrontal and motor cortices, but found no interaction between test sessions (before and after training) and speech type. Instead of training inside the scanner, Erb et al. (2012) used sought to predict individual differences in vocoded speech

adaptation ability by correlating with various metrics including amplitude modulation rate discrimination and gray matter volume (left thalamus).

Typically, acoustic models of CI speech are used for offline processing in a laboratory, a drawback which prevents interaction with the environment and scenarios truly realistic of what CI users face in the real world. Further, it is impossible for participants to hear their own voice in real time or during conversation. The goal of this study was to investigate the neural correlates of adaptation to degraded speech using a free-learning paradigm that permits our subjects to adapt in natural everyday environments linking perception and sensory motor systems. In addition to measuring changes in speech recognition performance, functional magnetic resonance imaging (fMRI) was conducted for each subject before and after a two-week exposure period, which involved wearing a real-time portable CI simulator for 2 hours each day. The portable CI simulator, implemented on the iPod Touch (Smalt et al., 2011), was based on work by Kaiser et al. (2000) and can be used to evaluate different CI training protocols.

#### **2. METHODS**

#### **2.1 Subjects**

Fifteen right-handed self-reported normal hearing subjects (9 male, 6 female) were recruited from the Purdue University student body to participate in the experiment involving a twoweek exposure period with vocoded speech. In addition, eight control subjects were recruited (5 male, 3 female) were recruited who participated in the behavioral testing only with no training involved in between. Participants were closely matched in age (mean age:  $24 \pm 3$  years) with no history of auditory disorders. Only two subjects had ever been exposed to noise vocoded-speech (exposure group), while none of the other participants had any experience with that type of auditory stimulus. All subjects gave informed consent in compliance with a protocol approved by the Institutional Review Board of Purdue University.

#### **2.2 Experimental Paradigm**

This experiment consisted of three different components: (1) an un-supervised learning period with a real-time portable cochlear implant simulator to improve performance in degraded speech comprehension; (2) behavioral testing sessions conducted in a sound treated room before and after the learning period to measure any changes in word comprehension of degraded speech; and (3) two fMRI imaging sessions performed prior and subsequent to each behavioral sessions. Fig. 1 provides a schematic of this experiment protocol.

#### **2.3 Portable CI Simulator**

A CI speech processor uses an array of non-overlapping filters which divide an acoustic signal it up into frequency bands corresponding to the number of electrodes in a CI. The amplitude and rate of electrical stimulation in each channel is determined based on the information in each frequency band. In a noise-vocoded simulation of a CI, the carrier frequency in each band is modulated with band-pass noise with the resulting acoustic signal being the sum of each channel (Shannon et al., 1995).

A portable implementation of CI simulation on the iPod Touch (Smalt et al., 2011) was developed based upon on the vocoder from Kaiser et al. (2000), and was used in this experiment to degrade naturally produced speech in real-time. The program can be used on the the Apple iPod Touch and iPhone; devices that carry an ARM processor equipped with a vector floating point unit. The Apple RemoteIO audio libraries enabled the real-time

simulation to run at a latency of 10 ms from audio input to output. The portability of the iPod provided a number of benefits for training purposes, allowing participants to not only hear sounds in their environment, but to hear their own speech. Butterworth infinite impulse response (IIR) filters of order N=4 were used in each of eight un-shifted (i.e. ideally inserted, no basalward shift) noise band channels, with the filter bank logarithmically spaced using (Greenwood, 1990) in the range from 250 Hz to 8 kHz.

#### **2.4 Behavioral Training and Testing Procedures**

To assess perceptual learning as a result of exposure to the real-time CI simulator, the device was worn by all 15 normal-hearing participants over a two week chronic exposure period for two hours per day. Participants were instructed to actively listen to speech or music throughout each two hour period. The device recorded the total run time for the program to verify that the participants were using the device for the required duration. The apparatus consisted of a lapel microphone and insert earphones, rated at 35–42 dB noise isolation (Etymotic, 2010).

Participants wore their portable simulator during testing which consisted of a word and sentence recognition task carried out in a sound treated room before and after exposure under free field conditions (Smalt et al., 2011). Three tests to measure degraded speech performance were used: phonetically-balanced (PB) words, PRESTO multi-talker sentences, and semantically anomalous sentences (Park et al., 2010). Participants wore their portable simulator during the entire testing session in a sound treated room before and after exposure under free field conditions. Although all stimuli were presented in their original form through the loudspeaker, participants only heard degraded speech through their simulator. The number of presentations of each stimulus type was 50 (words), 90, and 50 (sentences) respectively, where no stimuli were repeated across sessions. Subjects were asked to repeat out loud the word or sentence they heard while wearing the portable CI simulator. Each stimulus was presented twice, and included two corresponding responses. Scoring was done for each participant based on the percent correct of key words, as obtained from a recording of their verbal responses. The overall score for a given test was taken as the mean score of all of that subjects' trials.

Overall differences in speech perception before and after exposure were evaluated with a three way ANOVA for each test type (PB, PRESTO, anomalous sentences). Factors in the analysis included group (exposure, control), session (pre/post), and subject, with planned comparisons across each session to look for a learning effect. Only the first response to each stimulus was used in the ANOVA.

#### **2.5 fMRI Experimental Paradigm**

Patterns of neural activation as measured by the fMRI BOLD effect were evaluated before and after chronic exposure to the portable CI simulator using an auditory sentence comprehension blocked experimental paradigm previously demonstrated to be highly reliable across imaging sessions at the group level (Gonzalez-Castillo and Talavage, 2011). In the current study, two functional runs (duration=380 s) of the above-mentioned paradigm were conducted each for two run types: normal and degraded (CI simulated) speech. Each functional run had the same structure of blocks (Fig. 2), including 5 repetitions of the following sequence of blocks: visual instructions (5 s); auditory stimulus block (25 s); visual instructions (5 s); response/attention control task (15 s); and rest (20 s). During the visual instruction presentation and response/ attention control task blocks, no auditory stimulus was presented to the subjects. In the auditory stimulus blocks of the run, participants were asked to listen to six sentences from the City University of New York (CUNY) (Boothroyd et al., 1988) presented binaurally via pneumatic headphones while looking at a fixation

crosshair. Following this listening period, subjects responded to 3 fill-in-the-blank multiple choice questions displayed on the screen regarding one of the previously presented sentences, selected at random.Subjects were familiarized with the task by exposure to a practice run before entering the scanner during each session. This experimental paradigm was performed twice for each participant approximately two weeks apart (sentences were never repeated within or across sessions). Comparing across sessions (pre/post exposure), changes in brain activation exhibited by the CI sentences that differ from the normal condition would likely be a result of changes in processing strategies due to adaptation to transformed inputs.

#### **2.6 fMRI Data Acquisition**

Imaging was conducted at the Purdue MRI Facility (West Lafayette, IN) on a General Electric (Waukesha, WI) 3T Signa HDx scanner, using an 16-channel brain-array coil (Invivo), matching scan parameters to that of (Gonzalez-Castillo and Talavage, 2011). During the two imaging sessions, a single high resolution axial fast spoiled gradient-echo sequence acquisition (38 slices, slice thickness=3.8 mm, spacing=0 mm, FOV=24 cm, inplane resolution=256×256) was followed by four functional runs. Functional runs were obtained using a multi-slice echo planar imaging sequence (TR=2.5 s, TE=22 ms, 38 slices, slice thickness=3.8 mm, spacing=0 mm, in-plane resolution= 64×64, FOV=24 cm, flip angle=77°). Additionally, at the end of each imaging session, sagittal high resolution T1 weighted images (number of slices=190; slice thickness=1.0 mm; FOV=24 cm; in-plane resolution=256×224) were also acquired for alignment and presentation purposes. In addition to imaging data, behavioral data was collected while subjects were in the scanner. Subjects indicated their selection of the missing word during Response blocks via a 4-button response box (CURDES Fiber Optic Response Box Model No: HH-2×4-C). Both responses and response times were recorded.

#### **2.7 fMRI Data Analysis**

Data analysis was conducted using SPM8 software ([http://www.fil.ion.ucl.ac.uk/spm/\)](http://www.fil.ion.ucl.ac.uk/spm/). Preprocessing steps for each subject included: (1) discarding three initial volumes to account for T1 saturation, (2) motion correction, (3) registration of each functional run to the subjects segmented gray matter maps from the T1 weighted anatomical scan, (4) segmentation and transformation of the anatomy to the standard MNI, (5) applying the MNI transformation to the EPI data, (6) spatial smoothing with a Gaussian kernel (full width at half-maximum  $=$ 6mm). Individual subject analyses were first performed in all subjects to compare fMRI activation in the Listen block as compared to the Rest block (Listen > Rest) using the canonical HRF. This analysis was performed for both CI simulated and normal sentences. Statistical maps of activation from the (Listen > Rest) condition were compared across subject using a paired  $t$ -test across session (pre / post exposure), and across sentence type (normal / vocoded speech). Two thresholds were used when analyzing fMRI results: one for significance with multiple comparison correction ( $p_{FDR}$ <0.05) and one slightly relaxed to look for trends  $(p<0.0005)$ .

#### **2.8 Correlation of fMRI and Behavioral Analysis**

A stepwise regression model (entrance threshold  $p=0.05$ , exit threshold  $p=0.10$ ) was used to ascertain if any correlation existed between the mean BOLD t-statistic over 116 predefined MarsBaR (Brett et al., 2002) regions of interest (ROIs) and behavioral measurements (Breedlove et al., 2012). The algorithm finds a model that best captures the variation in the data, limiting the problem of over-fitting as in simple multiple regression models. The stepwise regression models the change in average fMRI t-statistic for each ROI to predict the out of scanner percent words correct scores combined across both sessions.The goal of such an analysis is to investigate correlations between bold fMRI data and the behavioral

metrics. The five variables (from table 1) collected from out of scanner behavioral data were chosen as the prediction variables since they had the greatest number of trials and are most likely to give a more accurate assessment of behavioral performance (Foster and Zatorre, 2010), while the constraints of fMRI limit accuracy of estimates (multiple choice vs. % words correct). The stepwise regression was done for both types of functional runs, with the hypothesis that there should be less significant correlation between using the portable CI simulator compared with the normal functional runs as opposed to the cochlear implant simulated runs. Data from each session (pre/post) are combined into the same regression, while fMRI sentence type was processed separately (normal, CI simulated).

#### **3. RESULTS**

#### **3.1 Behavioral Results**

Figure 3 summarizes the speech recognition performance pre/post training during testing sessions in a sound treated booth while subjects wore their portable CI simulator (Fig 3.A), as well as during the fMRI experiment (Fig 3.B). Using a linear regression on behavioral performance (% key words correct) for degraded speech combined across sessions, the outof-scanner PRESTO sentences were found to be correlated with the in-scanner responses to CUNY CI simulated sentences (% responses correct) ( $p < 0.05$ ), but not with the normal sentences ( $p < 0.35$ ). Table 1 indicates the correlation between the in-scanner and out-ofscanner behavioral metrics, where the PRESTO sentences showed the highest degree of correlation.

**3.1.1 Out of Scanner Results—**Fifteen subjects participated in the 2-week exposure period and fMRI study and did show a significant improvement in percentage words correct in all of the tests conducted when using there portable CI simulator at the group level. In stark contrast to the controls, 2 out of the 15 subjects scored lower in the follow-up session for the PB words, 0/15 for the PRESTO sentences, and 0/15 for the anomalous sentences. The range of improvement on the PRESTO sentences across session ranged from 6.5%– 35%.

An omnibus 3-way ANOVA (A×B×C; A=Test Type [PB, PRESTO, Anomalous], fixed; B=session [Pre, Post], fixed; C=subject [1..15], random) on percent key words correct revealed a significant main effect of session ( $F_{1,78} = 17.51$ ,  $p < 0.0001$ ) and test-type ( $F_{1,78}$ )  $= 60.37, p < 0.0001$ ), with no interaction ( $F_{1,78} = 0.97, p = 0.38$ ). A Priori Tukey-Kramer adjusted multiple comparisons between the session for each test-type comparing pre / post training were all significant ( $p < 0.01$ ). Based on the ANOVA, it can be concluded that the group of subjects who underwent the exposure period improved significantly between sessions on all tests with the portable CI simulator.

**3.1.2 In Scanner Results—**An omnibus 3-way ANOVA (A×B×C; A=Sentence Type [Normal, CI Simulated], fixed; B=session [Pre, Post], fixed; C=subject [1..15], random) demonstrated a significant main effect of session ( $F_{1,28} = 4.93$ ,  $p < 0.03$ ) and test-type ( $F_{1,28}$ ) = 33.02,  $p$  < 0.0001), with no interaction ( $F_{1,28}$  = 2.29,  $p$  = 0.14). A Priori Tukey-Kramer adjusted multiple comparisons between the session for each test-type comparing pre / post training showed a significant increase in performance in response to CI simulated sentences  $(p < 0.01)$ , but not for normal sentences  $(p = 0.62)$ .

**3.1.3 Behavioral Controls—**A second group of eight control subjects did not undergo the exposure period, and showed no statistically significant change in percentage words correct for any of the tests conducted across session while using the portable CI simulator (Fig. 4). Furthermore, 5 out of the 8 subjects scored lower in the follow-up session for the

PB words, 2/8 for the PRESTO sentences, and 3/8 for the anomalous sentences. The range of improvement on the PRESTO sentences across session ranged from −8.5%−6.3%, far less than that of the group who underwent the free-learning period.

An omnibus 3-way ANOVA (A×B×C; A=Test Type [PB, PRESTO, Anomalous], fixed; B=session [Pre, Post], fixed; C=subject [1..8], random) on percent key words correct revealed no main effect of session ( $F_{1,36} = 0.1$ ,  $p = 0.75$ ). A significant effect of test-type was observed ( $F_{2,36} = 45.44$ ,  $p < 0.0001$ ), with no interaction ( $F_{2,36} = 0.18$ ,  $p = 0.84$ ). Priori Tukey-Kramer adjusted multiple comparisons between the session for each test-type comparing pre / post training found no statistical significance ( $p = 0.65$ ). Based on the ANOVA results it can be concluded that the control group did not improve significantly based on repeating the same behavioral testing protocol.

#### **3.2 fMRI Results**

**3.2.1 Reproducibility—**A paired *t*-test at the group level was used to compare activation across the two sessions for the normal sentences during the Listen block. By looking at the effect of session, the amount of significant change in activation should be negligible if the task is consistent. This statistical analysis was performed on data collected for this study, as well as on the first 2 runs of 15 additional subjects previously acquired on the same scanner using the an identical paradigm (Gonzalez-Castillo and Talavage, 2011), and processing steps. No activation ( $p_{FDR}$ <0.05) was found for the separate paired t-tests between session 1&2 or 1&5 for the Gonzalez-Castillo 2011 data. Similarly, no activation changes  $(p_{FDR}$ <0.05) were found in the current study for normal sentences.

**3.2.2 Group Comparisons across Session—**In order to adequately understand the changes in activation related to the learning period with the CI Simulator, it is necessary to make individual comparisons across session for both types of functional runs i.e. sentence type (normal, CI simulated). Unlike the normal sentences, CI simulated sentences showed statistically significant ( $p_{FDR}$ <0.05) changes in activation across session; regions activated included superior temporal gyrus (STG), insula, cuneus, precentral gyrus, and cingulate gyrus, as shown in Fig. 5. A summary of the significantly activated clusters across session for CI simulated sentences can be found in Table 2. The relaxed threshold of ( $p<0.0005$ ) found no new clusters of activation for this condition.

**3.2.3 Group Comparisons across Sentence Type—**Comparing the normal and CI sentence conditions in the session following the exposure period using a paired  $t$ -test revealed activation differences in the bilateral superior and middle temporal gyri (BA 21, 22) at a significance level of ( $p_{FDR}$ <0.05). At a more relaxed significance level of  $(p<0.0005)$ , the two sentences types showed activation differences in several regions after the free learning period, as shown in Fig. 6. Activated regions included the STG, insula, precuneus, occipital gyrus, and SFG. A summary of the significantly activated clusters when comparing sentence types after the learning period can be found in Table 3.

#### **3.3 Correlation of fMRI and Behavioral Analysis**

Table 4 summarizes the correlations between the PRESTO behavioral data and 116 predefined regions of interest (ROIs) for both the functional runs using CI simulated and normal sentences. Behavioral data collected outside the scanner produced a much larger number of ROIs as well as high R^2 values, indicating that CI functional runs are a good predictor of behavioral performance in these regions. The regions implicated in the statistical analysis also match functional maps when comparing across session for the CI simulated runs, including the right Heschl's gyrus, precentral gyrus, postcentral gyrus, supplemental motor area, cingular gyrus, angular gyrus, and parietal lobe. Fig. 7 highlights

each region for both the normal (Fig 7.A) and CI simulated (Fig 7.B) sentences, where the color scale indicates the resulting R^2 value after the stepwise regression.

### **4. DISCUSSION**

After chronic exposure to real-time acoustic cochlear implant simulation, our results show a correlation between changes in behavior and functional measures, suggesting that short-term adaptation takes place and is physiologically measureable. In addition to improved sentence and word comprehension, we observed increased brain activation using fMRI in traditional and nontraditional speech brain areas, including working memory (Insula, inferior frontal cortex), motor (supplemental motor area), and visual (cuneus) regions. These changes suggest that training translated into widespread recruitment of areas beyond auditory cortex.

The paradigm used in this study allows participants to hear their own voice while engaging in speech activities and brings the learning process to the more naturalistic environment of everyday life. This translated into significant changes within the dorsal stream of speech processing, believed to rely on motoric representations of speech to aid in speech comprehension. An additional advantage of the current setup is that utilizing degraded speech in normal hearing listeners avoids confounds such as differences in surgical implantation and in length of deafness prior to implantation.

#### **4.1 Behavioral Adaptation to Degraded speech**

Behavioral performance was assessed outside the scanner before and after the exposure period while users wore a portable CI simulator for both sentence tests (PRESTO, anomalous), and an open set word test (PB). Participants showed improvement in percent correct of key words after two weeks of exposure to the portable CI simulator. In addition, a statistically significant increase in comprehension of CUNY sentences was observed in the scanner after the learning period for CI simulated sentences but not for normal sentences. Across both sessions, the PRESTO sentence scores obtained with the portable CI simulator outside the scanner were found to be correlated the CUNY CI simulated tasks collected during the fMRI scan. Normal CUNY sentences as presented during the fMRI sessions were not found to be correlated with behavioral measures collected with the portable CI simulator, an expected result since performance improved only for the CI simulated tasks. These results taken together with the behavioral control data provide strong evidence that the performance increase was actually due to the exposure to degraded speech, not merely familiarity with the task or other factors related to the experimental environment (Loubinoux et al., 2001; Specht et al., 2003). The same previously used in-scanner testing strategy (Gonzalez-Castillo and Talavage, 2011) also revealed no performance change across session for normal sentences. Our behavioral results demonstrate that exposure to degraded speech can transfer to novel stimuli in a short time period, even when "training" occurs in an uncontrolled environment, and are consistent with previous perceptual learning studies with vocoded speech (Davis et al., 2005; Hervais-Adelman et al., 2008), and even non-words (Hervais-Adelman et al., 2011).

#### **4.2 Neural Correlates of Adaptation to Degraded speech**

In addition to studying behavioral change due to the exposure to degraded speech in normal hearing individuals, we sought to investigate the neural correlates of this processing using fMRI. By comparing brain activations that occur across session for CI simulated sentences but not normal speech, we aimed to isolate changes related to adaptation of the degraded speech. We also compared our fMRI results with behavioral measures to examine which changes in activation are correlated with changes in performance in predefined regions of interest.

**4.2.1 Reproducibility—**Our fMRI protocol was identical to that used previously for data collected under the same paradigm, scanner, and analysis (Gonzalez-Castillo and Talavage, 2011). One difference between studies is that S01, S02 and S05 were collected, on average, within 7.8 days for the Gonzalez-Castillo data set, as opposed to just over two weeks for the current study. Also in the original study five runs per condition were available, while here there only two. The original study reported 86% voxel consistency over the five sessions at the group level, and no statistically significant activation change across session for normal speech. Our reproducibility study based on two sessions confirmed lack of statistically significant activation for the normal speech under the current modified experimental setup. This result permit us presume that changes exhibited across session in the CI simulated sentences and not in the normal sentences are likely due to changes in neural processing due to the exposure period.

#### **4.2.2 Changes in Degraded Speech Processing across Session—**When

comparing across sessions, activation patterns show significant changes in a set of distributed brain regions, including non-auditory areas, for CI simulated sentences as compared to normal sentences (See Table 2). Our hypothesis is that after a learning period, in addition to changes in primary and secondary auditory areas, new functional pathways are also established in both the dorsal (mapping sound into articulatory-based representations) and ventral (mapping sound into meaning) streams of speech processing (Hickok and Poeppel, 2004, 2007) to accommodate changes in behavioral performance.

**Primary and secondary auditory:** Activation changes across sessions were found for spectrally degraded speech in both the left and right superior temporal gyrus and sulcus (STG/STS). More activation after training was observed in the right hemisphere near Heschl's gyrus for CI sentences, perhaps indicating that more spectral information is being recovered from the signal. This is consistent with previous studies suggesting that right auditory cortex is involved in fine spectral analysis of auditory input (Schönwiesner et al., 2005; Zatorre et al., 2002). Additionally, Wernicke's area (BA 22 and BA 39) also showed strong increases in activation with training. This area is associated with a post-sensory interface that accesses long-term semantic knowledge, and is not typically activated by degraded speech when an abstract representation is not available (Obleser and Eisner, 2009; Obleser et al., 2007). Activation of this region only after training suggests that subjects were able to construct these associations during the training.

**Ventral Speech Processing Stream:** Our results show strong activation changes after the learning period with degraded speech in portions of the ventral stream, namely bilateral anterior and posterior Middle temporal gyrus (MTG, BA 21–22). Consistent with the Hickock and Poeppel model, left hemisphere activations were slightly dominant in this stream and most likely indicate an improvement in lexical comprehension. Since the stimuli for the fMRI portion of the experiment were sentences, it is likely that increased intelligibility upstream or elsewhere in the system might lead to improved lexical processing.

**Dorsal Speech Processing Stream:** Training with degraded speech also came accompanied by changes in areas part of the dorsal processing stream, namely insula and left inferior frontal cortex (BA47). In particular, changes in the insular region (BA 47, BA 13) suggest that working memory may be enhanced after training, and that stored information regarding the degraded speech is being utilized to improve performance. In addition to insula activation in vocoded speech (Hervais-Adelman et al., 2012), changes in the left inferior frontal cortex (BA 47) have been implicated in music (Levitin and Menon, 2003) and speech processing associated with rehearsal and maintenance of verbal information (Petrides et al.,

Smalt et al. Page 10

1993). Our results could suggest that after two weeks subjects have a better feature representation of the CI simulated speech to access and aid in comprehension. Other previous studies have suggested that both normal and degraded speech stimuli can recruit anterior insula when contrasting any active task with rest in both normal hearing subjects (Binder et al., 1999; Binder et al., 1997) and CI users (Giraud et al., 2001a). Activations in anterior insula could then be a result of task demands, not necessarily as a result of exposure to acoustic stimuli. Binder et al. (2004) suggested that the anterior insula may be recruited when additional decision-making processing is required for participants over some auditory signals. Our results could also support this hypothesis, and suggest that training may serve to enable this additional recruitment, and that additional decision-making machinery is required for degraded speech as compared to normal speech.

**Areas outside the streams of speech processing:** Other areas of activation found to change after the learning period included the cuneus, and the cingulate gyrus, both of which have been observed in PET imaging studies of CI users (Giraud et al., 2001a). Activation in the cuneus (secondary visual cortex) is not unexpected as it has been previously observed in response to auditory stimulation in CI users (Giraud et al., 2001a; Kang et al., 2004), and it is well established that auditory and visual cortices receive polysensory information (Benevento et al., 1977). This secondary visual cortex is often thought of as an association area, and its recruitment in both CI users and the normal hearing listeners from this study indicate additional processing capabilities that develop after exposure, possibly enabling an improved abstract representation of degraded speech. Also, activation was found to be stronger in the precuneus for CI simulated speech as compared to regular speech after training. Kang (Kang et al., 2004) agreed with this interpretation suggesting a visual representation of acoustic sources might be possible and could improve task performance. Alternatively, the change in activation may simply be related to task demand. Sensiorimotor and visual posterior regions of the precuneus have demonstrated connections in resting state studies to both working and premotor areas (Margulies et al., 2009), suggesting that better integration of sensory systems with stored representations may be occurring after exposure to the CI simulator. Finally, activation in the cingulate cortex has been identified in studies involving attention (Maddock, 1999; Maddock et al., 2003), memory (Maddock et al., 2001), and both normal (Xu et al., 2005) and degraded speech. Obleser et al. (2007) also found activation in this region while comparing high and low predictability vocoded sentences. These evidences together suggest that the cingulate cortex may be involved in the semantic or meaning processing of speech recognition. Future Work

One caveat of the current study is that the shift in frequency that occurs in CI users due to the lack of full insertion in CI studies was not included in the study, and could possibly affect the implicated regions of activation. However, the iPod touch application does have this capability implemented, and could be used in future studies. Other potential studies might include varying the number of frequency bands in the simulator to study adaptation across amounts of degradation both with and without upward shifts to model CI insertion – an important consideration as the technology and surgical procedures continues to improve. More detailed knowledge about this adaptation process following cochlear implantation may help improve training protocols and improve speech perception clinical outcomes in the future.

#### **5. CONCLUSION**

Utilizing a portable acoustic simulation of a cochlear implant, we were able to identify neural correlates of adaptation to a distorted input using fMRI. After two weeks of exposure for two hours each day with the degraded speech, increased behavioral performance was accompanied by changes in patterns of neural activation measured with fMRI that suggest

re-organization in both the auditory cortex and non-traditional speech areas. This adaptation could be necessitated by the increased processing power demanded in order to achieve a successful level of comprehension. Since the study was conducted in normal hearing listeners, specific confounds related to studies of CI users (i.e. the unique nature of each deafness) are not a factor in this study.

The findings of this study suggest that shifts in behavioral performance are associated with several changes in brain function. Increased activation in the auditory cortex, most noticeably in the right hemisphere, suggests that more information from the spectral content of the signal is being extracted. Secondly, the observation of increased supplemental motor area and visual cortex activity indicates increased utilization of phonological and abstract representations of the vocoded speech. These changes are accompanied by increased activation in brain areas related to the motor system, speech and working memory, suggesting an enhanced integration of knowledge acquired during the learning period.

#### **Acknowledgments**

This work was partially supported by funding from by NIH R01 (D.B.P.), NIH R01 DC03937 (M.A.S), NIDCD T32 DC00030 (C.J.S).

#### **REFERENCES**

- Bassim MK, Buss E, Clark MS, Kolln KA, Pillsbury CH, Pillsbury HC 3rd, Buchman CA. MED-EL Combi40+ cochlear implantation in adults. Laryngoscope. 2005; 115:1568–1573. [PubMed: 16148696]
- Benevento LA, Fallon J, Davis B, Rezak M. Auditory-visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. Experimental Neurology. 1977; 57:849–872. [PubMed: 411682]
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Rao SM, Cox RW. Conceptual processing during the conscious resting state. A functional MRI study. J Cogn Neurosci. 1999; 11:80–95. [PubMed: 9950716]
- Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T. Human brain language areas identified by functional magnetic resonance imaging. Journal of Neuroscience. 1997; 17:353–362. [PubMed: 8987760]
- Binder JR, Liebenthal E, Possing ET, Medler AD, Ward DB. Neural correlates of sensory and decision processes in auditory object identification. Nature Neuroscience. 2004; 7:295–301.
- Boothroyd A, Hnath-Chisolm T, Hanin L, Kishon-Rabin L. Voice fundamental frequency as an auditory supplement to the speechreading of sentences. Ear and Hearing. 1988; 9:306–312. [PubMed: 2975613]
- Breedlove EL, Robinson M, Talavage TM, Morigaki KE, Yoruk U, O'Keefe K, King J, Leverenz LJ, Gilger JW, Nauman EA. Biomechanical correlates of symptomatic and asymptomatic neurophysiological impairment in high school football. Journal of Biomechanics. 2012
- Brett M, Anton J, Valabregue R, Poline J. Region of interest analysis using an SPM toolbox [abstract/ 497] presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, Sendai, Japan. Available on CD-ROM in. Neuroimage. 2002:16. [PubMed: 11771970]
- Davis MH, Johnsrude IS. Hierarchical processing in spoken language comprehension. Journal of Neuroscience. 2003; 23:3423–3431. [PubMed: 12716950]
- Davis MH, Johnsrude IS, Hervais-Adelman A, Taylor K, McGettigan C. Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. Journal of Experimental Psychology: General. 2005; 134:222. [PubMed: 15869347]
- Dorman, MF. Speech perception by adults. In: Tyler, RS., editor. Cochlear Implants. Singular Publishing: San Diego; 1993. p. 145-190.

- Eisner F, McGettigan C, Faulkner A, Rosen S, Scott SK. Inferior frontal gyrus activation predicts individual differences in perceptual learning of cochlear-implant simulations. Journal of Neuroscience. 2010; 30:7179–7186. [PubMed: 20505085]
- Erb J, Henry MJ, Eisner F, Obleser J. Auditory skills and brain morphology predict individual differences in adaptation to degraded speech. Neuropsychologia. 2012; 50:2154–2164. [PubMed: 22609577]
- Etymotic. Earphone Comparison. 2010
- Foster NE, Zatorre RJ. A role for the intraparietal sulcus in transforming musical pitch information. Cerebral Cortex. 2010; 20:1350–1359. [PubMed: 19789184]
- Giraud AL, Price CJ, Graham JM, Frackowiak RS. Functional plasticity of language-related brain areas after cochlear implantation. Brain. 2001a; 124:1307–1316. [PubMed: 11408326]
- Giraud AL, Price CJ, Graham JM, Truy E, Frackowiak RS. Cross-modal plasticity underpins language recovery after cochlear implantation. Neuron. 2001b; 30:657–663. [PubMed: 11430800]
- 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. 2000; 123(Pt 7):1391–1402. [PubMed: 10869051]
- Gonzalez-Castillo J, Talavage TM. Reproducibility of FMRI activations associated with auditory sentence comprehension. Neuroimage. 2011; 54:2138–2155. [PubMed: 20933093]
- Green KM, Ramsden RT, Julyan PJ, Hastings DL. Neural plasticity in blind cochlear implant users. Cochlear Implants Int. 2008; 9:177–185. [PubMed: 18844322]
- Greenwood DD. A cochlear frequency-position function for several species--29 years later. Journal of the Acoustical Society of America. 1990; 87:2592–2605. [PubMed: 2373794]
- Harnsberger JD, Svirsky MA, Kaiser AR, Pisoni DB, Wright R, Meyer TA. Perceptual "vowel spaces" of cochlear implant users: implications for the study of auditory adaptation to spectral shift. J Acoust Soc Am. 2001; 109:2135–2145. [PubMed: 11386565]
- Hervais-Adelman A, Davis MH, Johnsrude IS, Carlyon RP. Perceptual learning of noise vocoded words: Effects of feedback and lexicality. Journal of Experimental Psychology: Human Perception and Performance. 2008; 34:460–474. [PubMed: 18377182]
- Hervais-Adelman AG, Carlyon RP, Johnsrude IS, Davis MH. Brain regions recruited for the effortful comprehension of noise-vocoded words. Language and Cognitive Processes. 2012; 27:1145–1166.
- Hervais-Adelman AG, Davis MH, Johnsrude IS, Taylor KJ, Carlyon RP. Generalization of perceptual learning of vocoded speech. Journal of Experimental Psychology: Human Perception and Performance. 2011; 37:283. [PubMed: 21077718]
- Herzog H, Lamprecht A, Kuhn A, Roden W, Vosteen KH, Feinendegen LE. Cortical activation in profoundly deaf patients during cochlear implant stimulation demonstrated by H2(15)O PET. J Comput Assist Tomogr. 1991; 15:369–375. [PubMed: 1851187]
- Hickok G, Poeppel D. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition. 2004; 92:67–99. [PubMed: 15037127]
- Hickok G, Poeppel D. The cortical organization of speech processing. Nature Reviews Neuroscience. 2007; 8:393–402.
- Kaiser, AR.; Svirsky, MA. Using a personal computer to perform real-time signal processing in cochlear implant research. Hunt, Texas: Proceedings of the Ninth Digital Signal Processing; 2000.
- Kang E, Lee DS, Kang H, Lee JS, Oh SH, Lee MC, Kim CS. Neural changes associated with speech learning in deaf children following cochlear implantation. Neuroimage. 2004; 22:1173–1181. [PubMed: 15219589]
- Levitin DJ, Menon V. Musical structure is processed in "language" areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. Neuroimage. 2003; 20:2142–2152. [PubMed: 14683718]
- Li T, Fu QJ. Perceptual adaptation to spectrally shifted vowels: training with nonlexical labels. J Assoc Res Otolaryngol. 2007; 8:32–41. [PubMed: 17131213]
- Loubinoux I, Carel C, Alary F, Boulanouar K, Viallard G, Manelfe C, Rascol O, Celsis P, Chollet F. Within-session and between-session reproducibility of cerebral sensorimotor activation: a test- retest effect evidenced with functional magnetic resonance imaging. J Cereb Blood Flow Metab. 2001; 21:592–607. [PubMed: 11333370]

- Maddock RJ. The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. Trends Neurosci. 1999; 22:310–316. [PubMed: 10370255]
- Maddock RJ, Garrett AS, Buonocore MH. Remembering familiar people: the posterior cingulate cortex and autobiographical memory retrieval. Neuroscience. 2001; 104:667–676. [PubMed: 11440800]
- Maddock RJ, Garrett AS, Buonocore MH. Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. Hum Brain Mapp. 2003; 18:30–41. [PubMed: 12454910]
- Margulies DS, Vincent JL, Kelly C, Lohmann G, Uddin LQ, Biswal BB, Villringer A, Castellanos FX, Milham MP, Petrides M. Precuneus shares intrinsic functional architecture in humans and monkeys. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:20069–20074. [PubMed: 19903877]
- Naito Y, Hirano S, Honjo I, Okazawa H, Ishizu K, Takahashi H, Fujiki N, Shiomi Y, Yonekura Y, Konishi J. Sound-induced activation of auditory cortices in cochlear implant users with post- and prelingual deafness demonstrated by positron emission tomography. Acta Otolaryngol. 1997; 117:490–496. [PubMed: 9288201]
- Obleser J, Eisner F. Pre-lexical abstraction of speech in the auditory cortex. Trends in Cognitive Sciences. 2009; 13:14–19. [PubMed: 19070534]
- Obleser J, Wise RJS, Dresner MA, Scott SK. Functional integration across brain regions improves speech perception under adverse listening conditions. Journal of Neuroscience. 2007; 27:2283– 2289. [PubMed: 17329425]
- Park H, Felty R, Lormore K, Pisoni D. PRESTO: Perceptually robust English sentence test: Open set —Design, philosophy, and preliminary findings. The Journal of the Acoustical Society of America. 2010; 127:1958.
- Petrides M, Alivisatos B, Meyer E, Evans AC. Functional activation of the human frontal cortex during the performance of verbal working memory tasks. Proceedings of the National Academy of Sciences of the United States of America. 1993; 90:878–882. [PubMed: 8430101]
- Schönwiesner M, Rübsamen R, Von Cramon DY. Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. European Journal of Neuroscience. 2005; 22:1521–1528. [PubMed: 16190905]
- Shannon RV, Fu Q-J, III JG. The Number of Spectral Channels Required for Speech Recognition Depends on the Difficulty of the Listening Situation. Acta Otolaryngol Suppl. 2004; 552:50–54. [PubMed: 15219048]
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. Science. 1995; 270:303–304. [PubMed: 7569981]
- Smalt, CJ.; Talavage, TM.; Pisoni, DB.; Svirsky, MA. Neural Adaptation and Perceptual Learning using a Portable Real-Time Cochlear Implant Simulator in Natural Environments. Boston, MA, USA: IEEE Engineering in Medicine and Biology Society; 2011.
- Specht K, Willmes K, Shah NJ, Jancke L. Assessment of reliability in functional imaging studies. J Magn Reson Imaging. 2003; 17:463–471. [PubMed: 12655586]
- Wilson, BS. Cochear Implant Technology. In: Niparko, JK.; Kirk, IK.; Mellon, NK.; Robbins, AM.; Tucci, DL.; Wilson, BS., editors. Cochlear implants : principles & practices. Philadelphia: Lippincott Williams & Wilkins; 2000. p. 109-127.
- Wong D, Miyamoto RT, Pisoni DB, Sehgal M, Hutchins GD. PET imaging of cochlear-implant and normal-hearing subjects listening to speech and nonspeech. Hear Res. 1999; 132:34–42. [PubMed: 10392545]
- Xu J, Kemeny S, Park G, Frattali C, Braun A. Language in context: emergent features of word, sentence, and narrative comprehension. Neuroimage. 2005; 25:1002–1015. [PubMed: 15809000]
- Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. Trends in Cognitive Sciences. 2002; 6:37–46. [PubMed: 11849614]

#### **Highlights**

- **•** Participants were exposed to a portable real-time acoustic cochlear implant simulation
- **•** After two-weeks of free-learning, participants improved in word recognition scores.
- **•** Activation increases with learning in auditory, motor, and working memory areas.
- **•** Behavioral performance changes were correlated with changes in brain function



#### **FIGURE 1.**

Schematic showing relationship between two-week exposure period, functional imaging and behavioral testing sessions.

Smalt et al. Page 16



### I<sup>\*</sup> INSTRUCTIONS

#### **FIGURE 2.**

fMRI auditory comprehension block paradigm, which has previously been shown to be highly reliable across imaging sessions at the group level. Subjects listen to 6 sentences in each block followed by multiple choice questions answered visually. Each block is repeated 5 times per run (Figure reproduced from Gonzalez-Castillo and Talavage 2011).

Smalt et al. Page 17



#### **FIGURE 3.**

Behavioral results across two sessions conducted both in sound booth with participants wearing their portable CI simulator (left panel) and during the fMRI sessions (right panel). Behavioral tasks included perceptually balanced single words (PB), PRESTO sentences, and anomalous sentences (while wearing portable device), and both normal and CI simulated CUNY sentences during the fMRI sessions. Significant training effects are observed both in and out of the scanner for CI simulated sentences after 2 weeks of exposure to the portable CI simulator. Behavioral performance on normal sentences in-scanner remained consistent across session.

Smalt et al. Page 18



#### **FIGURE 4.**

Histogram of behavioral improvement across the two sessions for the control group (gray) and learning group (black). Behavioral performance did not significantly improve for control subjects.



#### **FIGURE 5.**

Activation maps contrasting the two sessions (pre / post) for the CI simulated sentences at  $(p_{FDR}$ <0.05). Normal sentences produced no activation at this significance level. Yellow areas in the figure indicate regions where activation to the degraded speech significantly increased after exposure.



#### **FIGURE 6.**

Activation maps contrasting normal and CI simulated sentences for each session ( $p<0.0005$ ). Red areas in the figure indicate regions where normal sentences produced significantly more activation than degraded sentences, while blue areas indicate regions where degraded sentences produced significantly more activation than normal sentences

Smalt et al. Page 21

 $.25$ 

 $\mathbf{\hat{g}}$ 



**Normal** 

#### **FIGURE 7.**

Map of correlations of a stepwise regression between the BOLD t-statistic over a predefined list of ROIs and behavioral measurements region for both the CI simulated (top) and normal (bottom) sentences with behavioral data collected outside scanner using portable CI simulator. Yellow indicates areas of highest correlation.

R-value correlation between In-Scanner and Out-of-Scanner behavioral metrics. R-value correlation between In-Scanner and Out-of-Scanner behavioral metrics.



Statistically significant correlations ( $p$ <0.05) are indicated with a  $*$ . Statistically significant correlations (p<0.05) are indicated with a \*.

Clusters of significant activation for the contrast across the two fMRI sessions for CI sentences Clusters of significant activation for the contrast across the two fMRI sessions for CI sentences



Significant activations for the pre/post contrast for CI sentences at pFDR<0.05 and a minimum cluster size of 81 mm<sup>3</sup>. Abbreviations: x=Right-Left; y=Anterior-Posterior; z=Inferior-Superior. Significant activations for the pre/post contrast for CI sentences at pFDR<0.05 and a minimum cluster size of 81 mm<sup>3</sup>. Abbreviations: x=Right-Left; y=Anterior-Posterior; z=Inferior-Superior.

Clusters of significant activation for the contrast normal and CI simulated sentences after the free learning period. Clusters of significant activation for the contrast normal and CI simulated sentences after the free learning period.

 $\overline{\phantom{a}}$ 



(indicated with a  $*$ ) and p<0.0005 with a minimum cluster size of 81 mm<sup>3</sup>. Abbreviations: x=Right-Left; u=Anterior-Posterior; z=Inferior-Superior.

Abbreviations: x=Right-Left; u=Anterior-Posterior; z=Inferior-Superior.

Results of stepwise regressions relating change in average t-statistic of BOLD signal for the set of AAL ROIs (Tzourio-Mazoyer et al., 2002) with<br>behavioral data across the two training sessions. Stepwise regressions were behavioral data across the two training sessions. Stepwise regressions were preformed separately for each type of fMRI run (Normal, CI Simulated). Results of stepwise regressions relating change in average t-statistic of BOLD signal for the set of AAL ROIs (Tzourio-Mazoyer et al., 2002) with



 $\mathbf{r}$ 



NIH-PA Author Manuscript

NIH-PA Author Manuscript

Smalt et al. Page 26

