Frontotemporal activation differs between perception of simulated cochlear implant speech and speech in background noise: An image-based fNIRS study

Jessica Defenderfer[a,*], **Samuel Forbes**[b], **Sobanawartiny Wijeakumar**[c], **Mark Hedrick**[a], **Patrick Plyler**[a], **Aaron T. Buss**[d]

[a]Speech and Hearing Science, University of Tennessee Health Science Center, Knoxville, TN, United States

[b]Psychology, University of East Anglia, Norwich, England

[c]Psychology, University of Nottingham, Nottingham, UK

[d]Psychology, University of Tennessee, Knoxville, TN, United States

Abstract

In this study we used functional near-infrared spectroscopy (fNIRS) to investigate neural responses in normal-hearing adults as a function of speech recognition accuracy, intelligibility of the speech stimulus, and the manner in which speech is distorted. Participants listened to sentences and reported aloud what they heard. Speech quality was distorted artificially by vocoding (simulated cochlear implant speech) or naturally by adding background noise. Each type of distortion included high and low-intelligibility conditions. Sentences in quiet were used as baseline comparison. fNIRS data were analyzed using a newly developed image reconstruction approach. First, elevated cortical responses in the middle temporal gyrus (MTG) and middle frontal gyrus (MFG) were associated with speech recognition during the low-intelligibility conditions. Second, activation in the MTG was associated with recognition of vocoded speech with low intelligibility, whereas MFG activity was largely driven by recognition of speech in background noise, suggesting that the cortical response varies as a function of distortion type. Lastly, an accuracy effect in the MFG demonstrated significantly higher activation during correct perception relative to incorrect perception of speech. These results suggest that normal-hearing adults (i.e., untrained listeners of vocoded stimuli) do not exploit the same attentional mechanisms of the frontal cortex used to resolve naturally degraded speech and may instead rely on segmental and phonetic analyses in the temporal lobe to discriminate vocoded speech.

*Corresponding author. jdefende@uthsc.edu (J. Defenderfer).

Keywords

fNIRS; Image reconstruction; Cochlear implants; Post-lingual deafness; Vocoded speech; Event-related design; Speech recognition; Sentence processing; NeuroDOT

1. Introduction

Despite myriad sources of distraction in daily life, listeners' perception of speech demonstrates surprising resilience. The robustness of speech perception owes to the neural redundancy within the auditory system, whereby subcortical neural firing strongly correlates with stimulus patterns and becomes increasingly discerning to specific feature combinations of speech at the level of the cortex (Gervain and Geffen, 2019; Schnupp, 2006). Likewise, comprehension of speech generally follows a hierarchy of processing such that acoustic sensory analyses begin at the temporal lobe, and higher level, attentional mechanisms of the frontal cortex are recruited to resolve more complicated speech information (Davis and Johnsrude, 2003; Friederici, 2011). When degraded listening conditions complicate speech understanding, additional brain regions become activated beyond those recruited during favorable listening conditions (Defenderfer et al., 2017; Du et al., 2014; Mattys et al., 2012).

The neural response can vary based on the manner in which speech is compromised. For example, brain activity in some regions may exhibit a diminished response as intelligibility is reduced (Billings et al., 2009), while in other regions, a heightened response suggests specific neural mechanisms are activated to optimize speech understanding (Davis and Johnsrude, 2003; Davis et al., 2011). Neural processing of common external distortions (e.g.; multi-talker babble, background noise) have appeared in frontal regions, whereas speaker-related distortion (i.e. accented speech, voice quality) appear in temporal regions (Adank et al., 2012; Davis and Johnsrude, 2003; Kozou et al., 2005). Many studies attribute higher-order linguistic processes such as switching attention, inference-making, and response selection from competing stimuli to the frontal cortex (Friederici et al., 2003; Obleser et al., 2007; Rodd et al., 2005). Temporal regions are recruited to perform auditory analyses and early speech decoding processes (Hickok and Poeppel, 2007). Thus, the speech perception network uses multiple mechanisms to enhance perception in unfavorable listening conditions.

Cochlear implant (CI) users face unique challenges when listening to speech amid background noise due to the compounding effects of having a compromised auditory system in addition to dealing with the inherent signal distortion of the processor (Macherey and Carlyon, 2014). Despite widespread success with restoring access to speech, the use of CIs continues to exhibit huge variability in post-implantation outcomes (Blamey et al., 2013; Lazard et al., 2012). The CI speech processor inherently degrades all auditory input by stripping away the fine spectral properties of the speech signal. Post-lingually deafened individuals, who at one point had normal hearing, commonly report that listening through the CI does not resemble their auditory memories prior to their hearing loss (Boëx et al., 2006; James et al., 2001). Thus, there is a period of neural discordance wherein listeners are adapting to the altered input and re-learning the gamut of sounds in daily life (i.e.,

remapping neural pathways). In some listeners who continue to struggle using the CI, the attentional mechanisms within the neural systems of speech perception may not be flexible enough to enhance processing of speech when attempting to listen amid background noise.

CI speech simulations have long been used to examine how the NH auditory system treats stimuli that lack the perceptual properties it otherwise is accustomed to processing (Goupell et al., 2020; Pals et al., 2012; Sheldon et al., 2008). The process of vocoding is an artificial manipulation that results in speech stimuli that are similar to the output of speech processors worn by CI listeners. Fine spectral information is stripped from the speech input while preserving temporal properties of the speech envelope (Shannon et al., 1995), effectively removing the properties that make speech sound natural. The use of vocoded speech with NH listeners allows us to simulate variability of speech recognition performance observed in the CI population and also examine the impact of spectral degradation on the neural response. Prior to losing their hearing, post-lingually deafened CI recipients had normal auditory function, indicating that the neural infrastructure associated with typical hearing function was, at one point, intact. This may help explain why speech-related activity in post-lingually deafened CI users resemble that of NH listeners (Hirano et al., 2000; Olds et al., 2016; Petersen et al., 2013). Additionally, experienced CI users have demonstrated use of speech perception mechanisms also employed by NH listeners (Moberly et al., 2014; Moberly et al., 2016). It's important to note that use of vocoded stimuli with NH subjects is not expected to mimic how the neural system of CI listeners process auditory stimuli, as there are fundamental differences between the peripheral/central auditory systems of NH and CI users (L. Chen et al., 2016; Sandmann et al., 2015; Zhou et al., 2018). In the present study, CI speech simulations are expected to influence neural and behavioral responses in NH listeners, revealing effects unique to the spectral degradation of a CI. Thus, in the current project we assessed neural activity of NH adults to better understand how the frontotemporal response to CI simulations (i.e., artificial distortion) differs from processing speech in noise (i.e., natural distortion).

Frontotemporal activation has been cited in a number of studies that have manipulated speech intelligibility with vocoding. Temporal lobe engagement, specifically in the superior temporal gyrus (STG) and/or superior temporal sulcus (STS) (Giraud et al., 2004; Pollonini et al., 2014), underscore neural sensitivity to temporal speech features preserved in the vocoded speech. Other studies have found neural correlations with intelligibility along the STG and effort-related processing associated with prefrontal cortex (PFC) activity (Davis and Johnsrude, 2003; Eisner et al., 2010; Lawrence et al., 2018). PFC activation has also been observed during comprehension of vocoded speech stimuli, relative to speech in quiet (Hervais-Adelman et al., 2012). Similarly, results of an fMRI examination, which were later replicated using fNIRS (Wijayasiri et al., 2017), revealed significant PFC activation, specifically in the inferior frontal gyrus (IFG), while listeners attended to vocoded speech, relative to speech in quiet. Importantly, simply hearing the vocoded stimuli was not associated with IFG activity. Rather, activation depended on whether listeners were attending to the speech (Wild et al., 2012). These studies, however, used other attentional manipulations in the context of speech perception suggesting that PFC activity is not specific to processing vocoded speech and may be associated with the higher-level processes such as inhibition (Hazeltine et al., 2000), performance monitoring (Ridderinkhof et al., 2004),

working memory (Braver et al., 1997; J. D. Cohen et al., 1994), and attention (Godefroy and Rousseaux, 1996). Even so, a large body of evidence indicates that PFC activation plays an important role in optimizing speech recognition during difficult listening conditions (Demb et al., 1995; Obleser and Kotz, 2010; Poldrack et al., 2001; Wong et al., 2008). Thus, the specific role of PFC regions in processing vocoded speech remains to be demonstrated.

One way to better understand the neural mechanisms that give rise to speech perception is to examine the differences in cortical activation related to correct and incorrect speech recognition. The few neuroimaging studies that have made this direct comparison have reported elevated activation in different frontotemporal regions to both accurate (Dimitrijevic et al., 2019; Lawrence et al., 2018) and inaccurate perception (Vaden et al., 2013). Additionally, a recent fNIRS examination of temporal lobe activity in NH adults reported increased temporal cortex activation during accurate recognition of sentences in noise when compared to incorrect trials, highly intelligible vocoded speech stimuli, and speech-in-quiet stimuli (Defenderfer et al., 2017). What did not emerge from this study were differences in activation between natural speech and vocoded speech stimuli. Notably, the interpretation of these results was limited, first, by the regions measured, as the fNIRS probe only covered bilateral temporal lobes. Additionally, the vocoded sentences were highly intelligible and participants achieved near perfect performance in this condition. Incorporation of a vocoded speech condition with low intelligibility could reveal important cortical differences associated with how the brain optimizes recognition of degraded speech.

1.1. Current study

The aim of this study was to investigate the effects that simulated CI speech and speech in background noise have on the brain response. We recorded cortical activity using functional near-infrared spectroscopy (fNIRS), a non-invasive, portable, cost-effective imaging tool that utilizes the interaction between hemoglobin (Hb) and near infrared light to estimate cortical activation (Villringer et al., 1993). Unlike functional magnetic resonance imaging (fMRI), fNIRS generates very little noise and is compatible with hearing aid devices such as cochlear implants (CI) or hearing aids (Lawler et al., 2015; Saliba et al., 2016) making it an ideal tool for studying the neural basis of speech perception processes.

We address the limitations of previous research in two ways. First, we designed an fNIRS probe to cover left frontal and temporal regions and conducted volumetric analyses which can provide better alignment of data across participants and localize activation to cortical regions. This analysis method has been validated through comparison with concurrently-measured fMRI data (Eggebrecht et al., 2012; Wijeakumar et al., 2017) and has been applied in studies using only fNIRS (Forbes et al., 2021; Wijeakumar et al., 2019). This method will allow us to assess the degree to which these compensatory mechanisms activate across temporal and frontal cortices, and further, determine how they might differ during recognition of artificial versus natural forms of distorted speech. Second, many studies that use CI speech simulations parametrically vary speech intelligibility by altering the number of frequency channels (Hervais-Adelman et al., 2012; Miles et al., 2017; Obleser et al., 2008). Instead, we sought to create a realistic, low-intelligibility vocoded condition in which sentences amid background noise were vocoded. These stimuli are likely better

approximations of a CI user's daily listening experience and will allow us to investigate the neural mechanisms that are engaged when attention is needed to focus on speech lacking the fine spectral features usually characteristic of natural speech.

We used an event-related design to compare cortical activity associated with accuracy (correct, incorrect), intelligibility (high, low), and type of speech distortion (background noise, vocoding). The task included high- and low- intelligibility conditions for both vocoded speech (artificial distortion) and speech in background noise (natural distortion), using sentences in quiet for comparison. Intelligibility (as measured by averaged speech recognition score) was approximately equivalent between degraded speech types. However, acoustic composition between speech-shaped background noise and vocoded stimuli vary drastically. Unlike vocoding speech, incorporating background noise does not eliminate any component of the speech signal. Instead, the added noise acts as an energetic masker, blending acoustic signals and decreasing intelligibility of salient acoustic features of speech (Mattys et al., 2009). It is likely that neural mechanisms associated with extracting meaning from speech may differ according to the manner in which the speech is distorted. Thus, while behavioral performance is equal between these two speech conditions, we expect variations in the way the cortex resolves each form of distortion. For instance, we expect the auditory systems of NH listeners to be more familiar (and thus better prepared) to process speech in noise relative to vocoded speech. Top-down attentional mechanisms associated with activity in frontal regions should be available to deploy during speech-in-noise conditions but may not be flexible enough to optimize recognition of simulated CI speech. In typical, noisy settings, CI listeners face a number of complicating factors. First, ambient noise adds auditory input that is irrelevant to the targeted speech signal. Second, speech recognition is further compounded by the inherent signal distortion from the speech processor. Therefore, the low-intelligibility vocoded condition was created to reflect an ecologically-valid listening environment experienced by CI users by adding low level background noise to sentences in quiet prior to applying the vocoding process (detailed in section 2.2). The neural responses to these vocoded stimuli might help us better understand the cortical mechanisms used by post-lingually deafened CI listeners to resolve spectrally-degraded speech. By studying the interaction of these two factors in this unique way, we are hoping to increase our understanding of the mechanisms mediating accurate speech perception. Such an understanding may help guide future work to improve speech perception after CI implantation.

2. Methods and analyses

2.1. Participants

The Institutional Review Board of the University of Tennessee Knoxville approved the experimental protocol and plan of research. Based on our previous fNIRS study (Defenderfer et al., 2017), power analyses of a two-factor, within-subjects design suggests a minimum of 38 subjects to achieve a power of 80% with an effect size of 0.14. Thirty-nine adults (mean age 24.76 years, 21 females) participated in the study. All participants completed a consent form, handedness questionnaire, and demographic inventory prior to the experiment. Participants were between the ages of 18 to 30 years old, right-handed,

native English-speakers, and passed a hearing screening with auditory thresholds better than or equal to 25 dB HL at 500, 1000, 2000 and 4000 Hz. Participants received monetary compensation for their time. It is possible that the NIR wavelengths of interest are susceptible to absorption characteristics of hair color and density; however, subjects were not selected with regard to hair or skin color (Strangman et al., 2002). One participant was later discovered to have had a brain tumor which was removed 2 years prior to the experimental session; this dataset was excluded from the group analyses. The study results are based on 38 adults (20 females).

2.2. Speech material

Stimuli were created using sentences from the Hearing in Noise Test (HINT) (Nilsson et al., 1994), which are male-spoken and phonemically-balanced. A total of five listening conditions were created using Adobe Audition (v. 7) and Audacity (Audacity Team, 2017) software. Speech in quiet (SQ) was used as a baseline comparison to the distorted conditions. Two of the conditions were designed with high intelligibility (H) where ceiling performance was expected: vocoded speech (HV) and speech in low-level noise (HN). Two conditions were designed to be of low intelligibility (L) where performance was expected to be 50% on average across subjects: speech in high-level noise (LN) and sentences with low-level noise that were then vocoded (LV). Pilot data were collected from a sample of 40 NH individuals to determine the appropriate signal-to-noise ratios (SNRs) that would yield an average score of 50% correct for each low-intelligibility condition (these individuals were not participants in the current study).

HINT sentences were digitally isolated from their original lists and sampled at 44,100 Hz into 3-second tracks. For noise manipulations, a 3-second clip of the original HINT speech-shaped noise track was mixed with the isolated sentences. This noise is composed of the spectral components of all HINT sentences which are converted into a broadband spectrum identical to that of the HINT corpus. The measured total RMS value of each sentence utterance was modified to reflect target SNRs such that the level of the utterance changed, while the level of noise remained constant. This way, participants would not perceive noticeable changes in noise levels from trial to trial. Sentences were mixed with noise to reflect a +10 dB SNR for the HN condition and –4 dB SNR for the LN condition.

Speech stimuli for the HV and LV conditions were vocoded with AngelSim™ (TigerCIS) Cochlear Implant and Hearing Loss Simulator software. The HV condition contained 8-channel vocoded sentences. Isolated sentence files were band-passed into eight frequency channels, and temporal envelopes were extracted in each frequency band by half-rectification and low-pass filtering. The extracted envelope was used to modulate wide-band white noise and lastly, filtered with a bandpass filter. Trials for the LV condition received one additional step prior to vocoding. Sentences were first mixed with the HINT noise track at a +7 dB SNR, and were then 8-channel vocoded to simulate a realistic listening condition that CI recipients experience in a day-to-day environment.

Condition information is summarized in Fig. 1A. The presentation level was determined by measuring the full acoustic stimulus of 5 sentences from each listening condition with a sound level meter and 2 cc coupler (standard ANSI coupler to approximate residual ear

canal volume while wearing inserts), equaling approximately 65 dB SPL on average. SQ, HV, and HN conditions contained 30 sentence trials each. LN and LV contained 40 sentence trials each. Average performance in the low-intelligibility conditions were targeted at 50% resulting in approximately 20 correct trials and 20 incorrect trials per condition. Relative to the high-intelligibility conditions, the number of trials for the low-intelligibility conditions were reduced to avoid participant fatigue, while still maintaining a sufficient number of trials for statistical comparison. Participants received familiarization trials at the beginning of each block of trials for a condition (three for high-intelligibility conditions, six for low-intelligibility conditions). While NH participants are unfamiliar with vocoded stimuli, the familiarization trials were not intended to train performance with vocoded stimuli, rather orient participants to the nature of the stimuli. These trials were not included in the analyses. In total, there were 170 trials per participant.

2.3. Procedure

The current study implemented a speech recognition task in an event-related experimental design previously reported (Defenderfer et al., 2017). A research assistant placed the insert earphones and positioned the custom-made NIRS headband over designated regions of interest. The headpiece was adjusted to meet the participant's comfort level while also remaining adequately secure to ensure good contact between the optodes and scalp. Next, spatial coordinates for five scalp landmarks (right and left preauricular points, vertex (CZ), nasion, and inion) and the position of every source light and detector on each participant's head were recorded using Polhemus digitizing system.

Condition blocks were randomized to rule out any effect of order, and all sentence trials of one condition were presented together. Participants received breaks at the end of each condition block. In an attempt to reduce the introduction of signal artifacts, participants were asked to sit still and reserve large body movements for breaks between conditions. The following description of the trial paradigm can be seen in Fig. 1 from Defenderfer et al. (2017): Each trial began with a silent period (500ms) prior to onset of sentence presentation (3000ms), followed by a second silent period (jittered at 500, 1500, or 2000ms in a 2:1:1 ratio, respectively). A click sound (250ms) was played after each trial, cueing the participant to repeat the sentence during the repetition phase (3000ms). This is followed by another silent pause (jittered at 1000, 1500 or 2000 ms in a 2:1:1 ratio, respectively) before the beginning of the next trial. Timing of trial presentation was jittered to avoid collinearity between trial columns in the design matrix (Dale, 1999). Jittering reduces the occurrence of a deterministic pattern in the neural response and allows us to use deconvolution methods to parse rapid event-related activity from its associated trial type (Aarabi et al., 2017).

Participants were asked to listen to each sentence, wait for a "click," and then repeat as much of the sentence out loud to the best of their abilities. Participants were encouraged to guess any part of the sentence, and if not able to provide a response at all, they were told to say "I don't know." Instructions were also displayed on a computer monitor prior to the beginning of each listening condition, and each block of trials began at the participant's discretion by pressing the spacebar on the keyboard. Performance was scored as a percentage of correct trials within each condition. Using the HINT scoring criteria, a correct

response was defined by correctly repeating the entire HINT sentence (allowing 'article' exceptions). Participants did not receive feedback on performance accuracy. Sessions were audio-recorded and later scored by two research assistants.

2.4. fNIRS Methods

2.4.1. Hardware and probe design—The original data files used in the current study comply with the requirements of the institute, comply with IRB guidelines, and are available in the public domain (http://dx.doi.org/10.17632/4cjgvyg5p2.1). This study was conducted using a Techen continuous-wave 7 (CW7) NIRS system including 8 detectors and 4 source lights. The Techen CW7 simultaneously measures hemodynamic changes using 690 and 830 nm wavelengths. The experimental task was implemented in E-Prime (v. 3.0) and fNIRS data was synchronized to stimulus presentation with time-stamps at trial onsets. Given the limited number of source lights and detectors, we opted to focus the probe configuration over the left hemisphere owing to its dominant role in speech and language processing (Belin et al., 1998; Hickok and Poeppel, 2007). A headpiece was custom-made to record fNIRS data from left frontal and temporal cortices. The design accommodated a range of head sizes and comprised of thirteen 30 mm long channels and one 10 mm short separation (SS) channel (Fig. 1B). Channels were conFig.d to record data over T3 (STG), F3 and F7 (IFG) scalp locations of the 10:20 Electrode System. Incorporating short-distance channels has been shown to reasonably identify extracerebral hemodynamic changes (Gagnon et al., 2011; Sato et al., 2016). Due to the limited number of available sources/detectors, only one SS channel was included (Fig. 1B, channel 5). Noise within the head volume measured with fNIRS is spatially inhomogeneous across the scalp (Huppert, 2016). Therefore, it is possible that the single SS channel did not effectively remove artifact caused by superficial blood flow on the more distant long channels if the scalp blood flow patterns were different than what was measured on the SS channel. In order to optimize the effect of the SS channel, it was positioned over the temporal muscle and near the center of the probe design to target the most robust source of noise and capture superficial artifact associated with temporal muscle activity during vocalization (Schecklmann et al., 2017; Scholkmann et al., 2013).

2.4.2. Pre-processing of NIRS data and creation of light model for NeuroDOT —fNIRS data were analyzed in MATLAB with functions provided in HOMER2 (Huppert et al., 2009) and NeuroDOT (Eggebrecht and Culver, 2019). First, data were pre-processed in HOMER2. The raw signal intensity was de-meaned and converted to an optical intensity measure. Due to the potential motion/muscle artifact associated with speaking tasks, a more liberal correction approach was selected to counteract signal contamination. First we applied the hybrid method of combining spline interpolation and Savitzky-Golay filtering techniques (p = 0.99, frame size = 10s) to correct large spikes and baseline shifts in the data (Jahani et al., 2018; Savitzky and Golay, 1964; Scholkmann, Gerber, Wolf and Wolf, 2013; Scholkmann et al., 2010). Second, we used the modified wavelet-filtering technique (implemented with Homer2 *hmr-MotionCorrectWavelet*) (Molavi and Dumont, 2012) using an IQR threshold of 0.72. This method has been shown to effectively diminish motion artifact during experiments with speech tasks (Brigadoi et al., 2014). Channel-wise time series data at this stage of the processing are plotted from representative channels in the frontal and temporal lobe in Fig. 2A.

The first step before reconstructing the fNIRS data into image space is to prepare the atlas that will be used to create a structural image that is aligned to the digitized anatomical landmarks for each participant. Here, we used Colin's atlas. Next, a light model was created using the digitized spatial coordinates for the source and detector positions. Using *AtlasViewer*, photon migration simulations were performed to create sensitivity profiles by estimating the path of light for each channel using parameters for absorption and scattering coefficients for the scalp, CSF, gray and white matter (Bevilacqua et al., 1999; Custo et al., 2006). Sensitivity profiles were created with Monte-Carlo simulations of 10,000,000 photons for each channel (Fang and Boas, 2009). An example of the combined sensitivity profile for the entire probe is shown on a representative head volume in Fig. 1B. Sensitivity profiles for each channel were thresholded at 0.0001 and combined together to create a mask for each participant that reflected the cortical volume from which all NIRS channels were recording. A group mask was then created which included voxels in which at least 75% of participants contributed data. Since the fNIRS probe spanned lobes that were discontinuous in tissue, this group mask was divided into two separate masks corresponding to frontal and temporal lobes which allowed for activation in the two lobes to be analyzed separately.

2.4.3.   Image reconstruction with NeuroDOT—NIRS data were bandpass filtered to retain frequencies between 0.02 Hz and 0.5 Hz, removing high and low frequency noise that are often motion-based. Systemic physiology (pulse and respiration) was then removed by regressing the short separation data from the other channels. Finally, data were converted to hemoglobin concentration values using a differential path-length factor of 6 for both wavelengths. Volumetric timeseries data were constructed from these cleaned channel data following the procedure outlined by Forbes et al. (2021).

Image reconstruction in NeuroDOT integrates the simulated light model created in *AtlasViewer* with the pre-processed channel-space data. Measurements from the sensitivity profiles for each source-detector pair are organized into a 2-D matrix (measurements X voxels). NIRS files are converted to NeuroDOT format, in which SD information (source, detector, wavelength, separation) and stimulus paradigm timing information are extracted into reformatted variables. Channel data, originally sampled at 25 Hz, was down-sampled to 10 Hz to mitigate costly computational demands. A challenge unique to optical imaging is proper estimation of near infrared light diffusion in biological tissue, as image reconstruction of the NIRS data is subject to rounding errors and may lead to an under-determined solution (Calvetti et al., 2000). Therefore, the Moore-Penrose generalized inverse (Eggebrecht et al., 2014; Tikhonov, 1963; Wheelock et al., 2019) is used to invert the sensitivity matrix for each wavelength using a Tikhonov regularization parameter of $\lambda_1 = 0.01$ and spatially variant parameter of $\lambda_2 = 0.01$. Optical data are then reconstructed into the voxelated space for each chromophore (NeuroDOT function *reconstruct-img*). Relative changes in HbO and HbR are obtained each wavelength's respective absorption and extinction coefficients (NeuroDOT function *spectroscopy_img*) (Bluestone et al., 2001).

After reconstruction, general linear modeling is used to estimate the amplitude of HbO and HbR for each condition and for each subject across the measured voxels.We used an HRF derived from diffuse optical tomography (DOT) data for both HbO and HbR responses because it has shown to be a better fit than HRFs derived from fMRI (Forbes et al.,

2021; Hassanpour et al., 2014). The GLM comprised of eight regressors, including (1) speech in quiet (SQ), (2) speech in noise with high-intelligibility (HN), (3) vocoded speech with high-intelligibility (HV), (4) correct speech in noise with low-intelligibility (LN_c), (5) correct vocoded speech with low-intelligibility (LV_c), (6) incorrect speech in noise with low-intelligibility (LN_i), (7) incorrect vocoded speech with low-intelligibility (LV_i), and (8) time stamps associated with the vocal responses after each trial. Each event was modelled with a 3 second box-car function (corresponding to the duration of the sentence stimuli) that was convolved with a hemodynamic response function defined as a mixture of gamma functions (created using *spm_Gpdf*; $h_1 = 4$, $l_1 = 1 = 0.0625$; $h_2 = 12$, $l_2 = 0.0625$).

2.4.4. Validating image-reconstruction of fNIRS data—Image-based analyses of fNIRS data is a method that continues to be developed. Therefore, it is important to check for consistency after the image reconstruction process. Following the procedures described in Forbes et al. (2021) (see section 6.2), we correlated the channel-based time series data with the image-reconstructed time series for all subjects in this study. The mean amplitude of HbO and HbR were extracted from a 2 cm size sphere of voxels around the voxel with maximum sensitivity for each channel. Correlations were carried out between the average image-reconstructed time series and the channel-wise time series. In total, 988 correlations were performed for 38 subjects, 13 channels each (channel 5, the short separation channel, was excluded in this analysis). The histogram in Fig. 2B plots frequency of correlational values between channel and image-based time series data. Of the 988 correlations, 922 were greater than 0.25 (minimal acceptable threshold reported in Forbes et al., 2021). Within this subset which exceeded the criterion, the mean *r* value was 0.7. Thus, from these analyses we can conclude that the image-based reconstruction was an accurate reproduction of the channel-base data.

2.5. Statistical analyses

2.5.1. Analyses of variance between conditions—Group analyses were carried out using 3dMVM in AFNI (G. Chen et al., 2014). A summary of each statistical test can be reviewed in Table 1. fNIRS estimates cortical activation by tracking changes in the hemodynamic response that follows neural activity (Steinbrink et al., 2006). The process of neurovascular coupling suggests that neural activation results in a net increase of oxygenated hemoglobin (HbO) and concurrent net decrease of deoxygenated hemoglobin (HbR) (Buxton et al., 1998). For this reason, we included hemoglobin as a factor with measures of HbO and HbR. The first two repeated-measures ANOVAS examines how noise (Table 1A) and the process of vocoding (Table 1B) affect the neural response relative to the baseline response to speech in quiet. Table 1C details an ANOVA that examines whether cortical activity interacts with distortion type (noise versus vocoding) and/or intelligibility (high versus low). The ANOVA in Table 1D examines whether trial accuracy has an effect on the cortical response and whether this interacts with distortion type.

Unlike fMRI data in which noise is relatively uniform within the brain volume, noise in fNIRS data is heteroscedastic such that 1) temporal noise artifacts (i.e., motion, speaking) cause the artifact distribution to be heavy-tailed (yielding non-normal distribution) and 2) spatial noise is inherently different from channel to channel (Huppert, 2016). Therefore, we

conducted an omnibus 2 (Hemoglobin) X 5 (Condition) preliminary ANOVA to generate the voxel-wise residuals from each condition. These residuals were used to generate spatial autocorrelation parameters. AFNI's 3dClustSim uses these parameters to estimate minimum cluster size need to achieve a family-wise error of $\alpha < 0.05$ (in the case of multiple comparisons, alpha represents the probability of making at least one type I error) with a voxel-wise threshold of $p < 0.05$ (Cox et al., 2017). This process indicated a minimum cluster threshold of 83 voxels for the frontal lobe mask and 43 voxels for the temporal lobe mask. Voxel-wise HbO and HbR beta estimates were averaged for each participant from the clusters that satisfied threshold requirements and used to carry out follow-up tests (SPSS, IBM, version 25). The Greenhouse-Geisser correction for violations to sphericity were applied where necessary, and Bonferroni corrections were used to account for multiple comparisons in follow-up analyses.

2.5.2. Correlational analyses between performance score and cortical activation—Correlational analyses were carried out with AFNI's 3dttest++. Using a p threshold of 0.05, the LN_c voxel-wise HbO map was tested against zero using subject behavioral scores for the LN condition as a covariate to identify cortical regions where variance of LN_c voxel-wise betas covaried with performance. The same analysis was performed between the LV_c voxel-wise HbO map and behavioral scores from the LV condition. The same cluster size thresholds were applied as described above.

3. Results

Study results were based on data from 38 participants (20 females). Due to a task programming error, a small number of LV trials were unintentionally excluded from the experimental task for some of the participants: four participants received 32 LV trials and seven participants received 39 LV trials instead of the intended total of 40 LV trials. Relative to the total number of trials, it is unlikely that the absence of these trials affected the statistical analyses.

3.1. Behavioral data: speech recognition performance

Behavioral performance and comparisons between each condition are detailed in Table 2. Participants achieved, on average, 99.7% ($SD +/- 1\%$) in the HN condition, 92.5% ($SD +/- 5.6\%$) in the HV condition, and perfect scores in the SQ condition. LN performance varied between subjects ranging 27.5% to 75% correct; LV scores varied between subjects from 27.5% to 72.5% correct. As expected, average scores in LN and LV were 48% ($SD +/- 12.4\%$) and 49.9% ($SD +/- 10.8\%$), respectively. Two research assistants independently scored each speech perception measure with high interrater reliability ($\alpha = 0.9817$, Krippendorff alpha). Paired samples t-tests between performance scores in each condition can be viewed in Table 2. Performance in the LV and LN conditions were significantly worse than the other conditions but were not significantly different from each other.

3.2. fNIRS data

3.2.1. Image-based data—Results of the ANOVAs are listed in Table 3. The MNI coordinates of the center of mass for each cluster denote the cluster location. Note that these coordinates and the number of voxels is not a quantitative measure of each cluster, as these image-based analyses are a projection of the two-dimensional fNIRS data into three-dimensional space. Rather, the MNI coordinates and cluster size provide an enhanced description of activation localization and extent of the response, respectively. Significant main effects and interaction effects appeared in portions of the temporal and frontal cortices for all ANOVAs, suggesting that our ROIs were sensitive to the experimental task.

Table 3A summarizes the results of the first ANOVA which examined how added noise affected the baseline response to speech in quiet. A main effect of Hemoglobin was found in the IFG ($F(37) = 11.62$, $p = .002$) and two clusters in the middle temporal gyrus (MTG) ($F(37) = 6.81$, $p = .013$ and $F(37) = 10.34$, $p = .003$, respectively). Further inspection of the first cluster in the MTG (anterior to the second MTG cluster) revealed an inversed response, in which change in HbO was negative and change in HbR was positive. An interaction between Hemoglobin and Noise Level appeared in the middle frontal gyrus (MFG) ($F(37) = 5.87$, $p = .004$) and IFG ($F(37) = 6.41$, $p = .006$). The first cluster revealed significantly higher activity for speech recognition in high level background noise (LN_c) relative to the easier SQ and HN conditions (see Fig. 3). In the second cluster, changes in both HbO and HbR were found to be negative for all three conditions, where the most negative changes occurred in the LN_c condition. In a similar manner, the ANOVA in Table 3B examined how simulated CI speech affected the baseline response to speech in quiet. Main of effects of Hemoglobin were observed in the MFG ($F(37) = 10.61$, $p = .002$) and MTG ($F(37) = 7.77$, $p = .008$). The second hemoglobin response was in-versed, showing negative changes in HbO and positive changes in HbR. An interaction between Hemoglobin and Vocoding Level was observed in the MTG ($F(37) = 6.59$, $p = .002$), where activation during the low-intelligibility vocoded (HV_c) and speech in quiet conditions showed significantly greater activation relative to the high intelligibility vocoded condition (HV) (see Fig. 3). Results of follow-up paired samples t-tests for the Hemoglobin X Noise Level (ANOVA A) and Hemoglobin X Vocoding Level (ANOVA B) interactions are listed in Table 4.

Table 3C summarizes the ANOVA which tested whether cortical response was affected by or interacted with distortion type (noise, vocoding) and intelligibility (high, low). A main effect of Hemoglobin was observed in the MFG ($F(37) = 12.32$, $p = .001$) and in the MTG ($F(37) = 7.06$, $p = .012$). The first demonstrated the conventional activation response (increase in HbO, decrease in HbR) while the second demonstrated an inversed response. Hemoglobin X Distortion interactions were observed in the IFG ($F(37) = 6.08$, $p = .018$) and MTG ($F(37) = 9.06$, $p = .005$) (see Fig. 4). The interaction in the IFG revealed changes in hemoglobin to be negative for both oxy- and deoxyhemoglobin, and no significant difference between HbO and HbR for either distortion condition. The MTG cluster showed significant activation for speech in noise conditions (HN, LN_c) relative to a lack thereof during the vocoded speech conditions (HV, LV_c). Hemoglobin X Intelligibility interactions were observed in the MTG ($F(37) = 18.59$, $p < .001$) and MFG ($F(37) = 5.99$, $p = .019$), both of which revealed

significantly more activation during low-intelligibility conditions (LN_c, LV_c) relative to high-intelligibility conditions (HN, HV).

Table 3D summarizes the results of the second ANOVA which examined the effect of distortion type (background noise and vocoding) and trial accuracy (correct and incorrect). This ANOVA analyzed responses between LN_c, LN_i, LV_c and LV_i. Two effects of Hemoglobin were observed in separate clusters in the IFG ($F(37) = 10.35$, $p = .003$ and $F(37) = 9.74$, $p = .003$, respectively). The first demonstrated the conventional hemodynamic response, while the second was inverted. A third effect of Hemoglobin was found in the MTG ($F(37) = 8.26$, $p = .007$). An interaction between Hemoglobin and Accuracy appeared in the MFG ($F(37) = 9.72$, $p = .004$) and the MTG ($F(37) = 7.32$, $p = .010$). The MFG cluster demonstrated a significant increase in activation during correct responses relative to incorrect responses (see Fig. 5A). Alternatively, the second cluster in the MTG revealed negative changes in both HbO and HbR.

Lastly, we examined brain-behavior associations by running correlational analyses between accuracy in the low-intelligibility conditions and HbO change during these conditions. In the LN_c condition, HbO change in the MFG was negatively associated with performance ($r = -.458$; $p = .004$; see Fig. 6). HbO change in the IFG was positively associated with LN performance ($r = .393$; $p = .015$), but the majority of HbO measures were less than zero. Change of HbO from LV_c trials in the MFG was positively associated with performance ($r = .430$; $p = .007$), whereas HbO change was negatively associated with performance in the MTG ($r = -.398$; $p = .013$).

4. Discussion

4.1. Effects of distortion type and speech intelligibility on the cortical response

In the current report, we examined the effect of distortion, speech intelligibility, and performance outcome on the neural responses in left frontal and temporal cortices. Consistent with existing literature on speech processing, activation was found across large swaths of these regions (Golestani et al., 2013; Mattys et al., 2012; Peelle, 2018). This study sheds light on how the NH brain reacts to decreasing intelligibility and how compensatory mechanisms differ between distortion types. LN and LV conditions were designed to decrease average speech perception performance by approximately 50%, whereas their corresponding conditions with high intelligibility, HN and HV, yielded ceiling effects in behavioral performance (Table 2A). To assess activation during correct perception of speech in noise, HN and LN_c conditions were contrasted with SQ (Table 1A). The noise effect in the MFG (Fig. 3, top) was driven by stronger activation to the LN_c trials, relative to both HN and SQ. This is consistent with previous reports (Golestani et al., 2013; Wong et al., 2008) and would suggest that the elevated activity is associated with neural mechanisms that support speech understanding during degraded listening conditions, and its absence amid highly intelligible speech in low-level noise (HN) suggests it's not an obligatory response to the presence of noise. The second interaction seen in the IFG showed a nonconventional pattern of activity and was inverse of the response seen in the MFG cluster. That is, this region showed a significant decrease in HbO low-intelligibility noise (LN) condition. Decreases in HbO levels are not well understood (discussed in more detail in Section 4.4),

but one possibility is that MFG and IFG are vascularly coupled. More evidence of a possible relationship between the MFG and IFG during the processing of LN stimuli can be observed in the brain-behavior correlational analyses (see Section 4.4).

The corresponding analysis of vocoding speech effects (ANOVA in Table 1B) reveal an interaction in the MTG (Fig. 3), in which vocoded speech with low intelligibility (LV$_c$) was associated with stronger activity relative to highly intelligible vocoded condition (HV). This region of the temporal lobe has been associated with combining phonetic and semantic cues, allowing for the recognition of sounds as words and comprehension of a word's syntactic properties (Gow Jr., 2012; Graves et al., 2008; Majerus et al., 2005). Conditions where the signal information is highly compromised would force the listener to rely more heavily on this mechanism. Hence, elevated MTG activity may reflect compensatory neural engagement associated with enhancing lexical interface between sound and meaning (Hickok and Poeppel, 2004). Interestingly, the Hb X Vocoding Level cluster also revealed that speech in quiet (SQ), or natural speech evoked significantly stronger activation relative to the HV condition as well. The integrity of spectral and temporal information is uncompromised in the SQ stimuli, and therefore, this activity may reflect the unrestricted lexical representation of phonemic and syllabic speech information in the temporal cortex (Poeppel et al., 2008). This finding is consistent with the Ease of Language Understanding (ELU) model, which suggests that activation associated with natural speech processing will be represented by mechanisms in the STG and MTG (Rönnberg et al., 2013). If incoming speech information fails to rapidly and immediately map onto known phonemic/lexical representations in the temporal cortex, higher level linguistic mechanisms might then be recruited to exploit other available features of the speech, not unlike the pattern of activation seen in the Hemoglobin X Noise Level interaction in the MFG. It's surprising to find that HV lacked activation relative to both SQ and LV. Even though HV sentences were highly intelligible, resulting in near ceiling performance, the speech was still compromised due to the vocoding process. Hence, we might expect that this degradation might interfere with matching phonemic/syllabic representations. It's possible that the highly intelligible vocoded sentences-in-quiet aren't sufficiently degraded to trigger compensatory strategies, but also lack the full perceptual qualities of natural speech to evoke typical speech processing mechanisms, as well.

The interesting difference between the Noise Level ANOVA and the Vocoding Level ANOVA is where these compensatory strategies are recruited. That is, directly comparing speech in noise with the baseline response to speech in quiet reveals that listeners, at a group level, tend to rely on top-down frontal speech processing to resolve noise-degraded speech. Understanding speech in background noise is made easier by recruiting linguistic mechanisms such as inference-making, inhibition, and switching attention. Consistent with previous imaging studies (Davis et al., 2011; Mattys et al., 2012; Scott et al., 2004; Wong et al., 2008), the present study shows elevated frontal activation in the MFG associated with recognizing speech degraded by noise. On the other hand, directly comparing vocoded speech with speech in quiet conditions indicate that listeners rely on initial cortical processing in the temporal lobe to resolve highly-degraded vocoded speech.

Both types of distortion were contrasted with each other in ANOVA C, which was designed to evaluate whether cortical activation during accurate speech recognition interacted with the type of distortion and/or its level of intelligibility. Two interactions between Hemoglobin and Intelligibility appeared in the MTG and MFG, both demonstrating that regardless of the manner that the speech is distorted, conditions with low intelligibility are associated with significantly higher activity relative to conditions with high intelligibility. This is consistent with Davis and Johnsrude's seminal fMRI investigation of hierarchical speech comprehension, which reported 'form-independent' (form of distortion) activation in the left middle and frontal gyri (Davis and Johnsrude, 2003). This means that the typical auditory system is able to resolve degraded speech (regardless of the type), by recruiting higher-level linguistic mechanisms when they are available. The contrast between noise-degraded speech and vocoded speech (regardless of intelligibility) reveals a Hemoglobin X Distortion interaction in the MTG, in which a stronger cortical response is observed during speech in noise relative to vocoded speech. This could be because the cortical responses of NH listeners are attuned to processing speech in noise, as this is a common experience in everyday life. Listeners show that they are able to exploit top-down mechanisms to optimize speech understanding even when the speech is vocoded, as evidenced by the Hemoglobin X Intelligibility interactions in the frontal lobe. However, due to subtractive nature of the vocoded speech combined with their lack of experience with vocoded stimuli, the neural pathways to access these top-down strategies are not stabilized, and therefore less reliable (explaining the lack of significant frontal activity when contrasting vocoded speech with baseline speech in quiet).

Overall, these findings reveal important differences in how the temporal lobe and frontal lobe resolve these two types of distortion. Previous research indicates that neural mechanisms of speech recognition adapt with task demands, the listeners' motivation/attention, and semantic knowledge from previous experience (Leonard et al., 2016; Rutten et al., 2019). Results of the current study are consistent with this account. Given a lifetime of conversations riding in the car, talking on the phone, eating at restaurants, or listening to the television over the hum of an air conditioner or vacuum cleaner, listeners with normal hearing have extensive, well-established neural representations and pathways associated with listening to speech in background noise. If incoming auditory information is compromised, listeners are able to pull from multiple cortical networks to optimize speech understanding. This explains the robust frontal response during low-intelligibility speech in noise when compared to high-intelligibility conditions (SQ, HN) (Hb X Noise Level interaction from ANOVA A), in addition to the increased temporal sensitivity to speech in noise when directly compared to vocoded speech (Hb X Distortion interaction from ANOVA C). However, when speech is simulated to reflect a more realistic listening condition experienced by CI listeners, NH listeners show less reliance on experience-driven, top-down pathways and more reliance on bottom-up auditory analysis and word meaning processing.

4.2. Effects of distortion type as a function of behavioral outcome

Consistent with previous reports that compare the neural response of correct and incorrect perception (Dimitrijevic et al., 2019; Lawrence et al., 2018), we observed a significant

interaction between Hemoglobin and Accuracy in the MFG (Fig. 5A), such that significant activity was observed during accurate speech recognition trials. This suggests that recruitment mechanisms in the frontal lobe are not obligatory responses that come online during more complex tasks, but instead, directly relate to whether subjects are doing the task successfully. This interaction is collapsed across type of distortion, suggesting that listeners are able to exploit similar MFG mechanisms when sufficient speech information is preserved in artificially distorted speech.

Given its domain-general functionality, the role of PFC activation has been associated with experimental tasks involving response conflict and error-detection (Carter et al., 1998; Rushworth et al., 2007). An elevated response during accurate performance, however, aligns with many neuroimaging accounts that associate left PFC activity with performance monitoring during tasks where attentional control is needed to optimize performance when the task is challenging but doable (M. X. Cohen et al., 2008; Dosenbach et al., 2008; Eckert et al., 2016; Kerns, 2006). The FUEL model (framework for understanding effortful listening) would further suggest that this activation is modulated by the listener's motivation to perform the task (Pichora-Fuller, 2016). The cost of exerting attentional control is related to the reward-potential associated with the task (be it external or intrinsic) (Shenhav et al., 2013); therefore, activation increases in the frontal lobe during challenging cognitive tasks insomuch that the participant is sufficiently motivated and able to perform the task. It's important to note that neither motivation nor effort was measured in the current study. Additionally, measures of effort have been shown to vary significantly between listening conditions where behavioral performance is otherwise equivalent, indicating that the negative impact of increasing cognitive demands can go unnoticed if simply assessing a performance score (Francis et al., 2016; Zekveld and Kramer, 2014). However, the FUEL account could, in part, explain the lack of activation during incorrect perception if listeners are disengaged during incorrect trials. Several studies have already documented the impact of decreasing intelligibility on measures of effort (Ohlenforst et al., 2018; Winn et al., 2015). Ongoing work (Defenderfer et al., 2020; Zekveld et al., 2014) using independent measures of effort, such as pupillometry, concurrently with neural measures may help to resolve the role of effort in the relationship between brain and behavior reported here.

It is interesting to note that we did not find a comparable result to the accuracy effect reported in a previous study (Defenderfer et al., 2017), where significantly greater activity in the temporal lobe was associated with correct speech-in-noise trials. Rather, the Hb X Accuracy cluster in the MTG shows a non-canonical response pattern with decreased HbO. Previous imaging studies of speech perception have reported left antero-lateral temporal activation to be associated with speech intelligibility (Evans et al., 2014; Narain et al., 2003; Obleser et al., 2007; Scott et al., 2000). While we found an accuracy effect (correct > incorrect, i.e., intelligible > unintelligible) in the frontal lobe, we did not find this effect in the STG. First, it's possible that the noise/artifact associated with speaking may have been so pronounced, that after SS channel signals were corrected from the dataset, no meaningful effects could be recovered. Secondly, while the conditions used in Defenderfer et al. (2017) were nearly identical to speech-in-quiet (SQ), vocoded (HV), and speech-in-noise (LN) conditions of the current study, Defenderfer et al. (2017) used loudspeakers for stimulus presentation, whereas the current study used insert earphones. Thus, the manner of stimulus

presentation could have altered the quality of the stimulus and/or attentional strategies in this task.

4.3 Evidence of brain-behavior relationship with speech recognition

By examining the relationship between neural activation and behavioral performance in the low-intelligibility conditions, we were also able to gain insight into how individual differences in activation were associated with success on the low-intelligibility speech perception conditions (LN_c and LV_c). Average change in HbO was negatively correlated with speech scores in the MFG (Fig. 6A). This negative association suggests that listeners' ability to recognize speech in noise is inversely related to the degree to which the region is engaged. Recall that group-level MFG activity is associated with processing of correct LN trials relative to the higher intelligibility conditions. Together, these results suggest that MFG activity supports speech recognition more strongly for individuals that perform more poorly in the LN condition. Neuroimaging evidence indicates that some listeners may exhibit neural adaptation such that neural responses decrease as listeners become accustomed to novel stimuli (Blanco-Elorrieta et al., 2021). Within the current task, response variability from this region may reflect cortical efficiency with which subjects are able to resolve the speech in noise, such that the poorest performers rely more heavily on MFG activation (with little to no adaptation), and better performers that have more efficient frontal mechanisms exhibit relatively lower activation. Second, a positive association between change HbO during LN_c trials and LN performance was observed in the IFG. As shown in Fig. 6B, however, change in HbO tended to be negative and increased closer to zero with higher speech scores. In contrast to MFG, the Hb X Noise Level interactions from ANOVA A revealed a significant decrease in HbO on the LN condition which was inverse to the pattern of MFG. Thus, the coupling between these two regions is evident at the individual subject level, and the relationships with speech score and activation at the individual level is consistent with group level activation.

Average HbO change in the IFG during LV_c trials was positively correlated with performance during the LV condition, while an inversed relationship was observed in the MTG (Fig. 6B). Previous work examining the learnability of vocoded speech stimuli report a similar correlation between comprehension of CI simulations and activity in the IFG (Eisner et al., 2010). Many methodological differences exist between the vocoded stimuli of the current study and that used in Eisner et al. (2010). However, it remains possible that this cluster in the IFG could be demonstrating variability in activation based on individual differences in learning capacity across the sample, as higher activity is associated with better performance. The negative correlation in the MTG suggests that better performers in the LV condition need not rely as heavily on the initial cortical processes of the temporal lobe, and are more readily able to recruit frontal mechanisms to resolve heavily degraded vocoded speech.

4.4 Non-canonical (inverted) hemodynamic responses

In this study, we observed inverted hemodynamic responses, demonstrating changes in hemoglobin opposite of the typical canonical response (e.g., negative HbO, positive HbR). The hemodynamic response that we measure with fNIRS is a secondary measure of

the neuronal activity taking place in the cortex. The complex nature of neurovascular coupling often requires careful interpretation of results, as the neurophysiological basis of a negative or inverted response is not completely understood. Inverted responses are commonly reported in infants, and studies have suggested multiple possible explanations for such responses including changes in hemocrit during transition from fetal to adult hemoglobin (Zimmermann et al., 2012) or the interaction between ongoing developmental changes during infancy and the influences of stimulus complexity and experimental design (Issard and Gervain, 2018). However, the research on the relationship between inverted NIRS responses and cortical activity in adults is limited. Evidence from fMRI (Christoffels et al., 2007), magnetoencephalography (MEG) (Ventura et al., 2009), fNIRS (Defenderfer et al., 2017), and electroencephalography (EEG) (Chang et al., 2013) studies indicate that inverted response functions could reflect cortical suppression related to speech production processing. Given that the current study had participants vocalize their responses during the task, it's possible that the inverted/non-canonical responses observed here could reflect such speech-related suppression. However, any influence of speaking-related artifact on the activity from responses during speech perception should have been mitigated, as we modeled the responses phase in the GLM, and every condition trial was followed by a vocal response (therefore, contrasting conditions should cancel out any effect of this). Alternatively, NIRS methodological studies demonstrate that muscle activity can cause increases in both HbR and HbO (Volkening et al., 2016; Zimeo Morais et al., 2017) and has been shown to influence NIRS data during tasks which involve overt speaking (Schecklmann et al., 2010). The inverted/non-canonical responses observed near the temporal muscle could reflect the influence of muscle-related activity. Channel-wise time series data demonstrating non-canonical and/or inverted hemodynamic responses are plotted at the bottom of Fig. 2A. Additionally, the physical act of speaking can cause respiration-induced fluctuations of carbon dioxide (CO_2) in the vascular system (Scholkmann et al., 2013). Decreases in CO_2 are associated with cerebral vasoconstriction and can result in a relative increase in HbR (Tisdall et al., 2009). All things considered, the inverted responses observed in the present study should be cautiously interpreted, as the neurophysiological mechanisms underlying non-canonical responses are not fully understood.

5. Limitations and future directions

While fNIRS presents numerous advantages relative to other imaging techniques, the nature of near infrared light poses inherent limitations to the technique. Generally, the measurement depth is limited to regions within 1.5 – 2 cm of the scalp; therefore, our interpretation of fNIRS recordings are limited to the cortical surface (Chance et al., 1988). Additionally, the changes in optical density measure by fNIRS is a cumulative result, reflecting possible contributions from superficial blood flow, skin circulation, and cardiovascular effects (Quaresima et al., 2012). These limitations are addressed in the current study by experimental methods we implemented and application of rigorous artifact correction techniques prior to extracting hemodynamic estimates. Muscle artifact may have contaminated certain signals closer to the temporal muscle, and it's possible that this contributed to the inversed responses observed in this study, as muscle artifact can lead to an inversed response of either HbO and/or HbR (Volkening et al., 2016). Incorporation

of a short distance probe over the temporal muscle mitigated the possible effects of muscle artifact.

Comparison of accurate and inaccurate trials during sentence recognition revealed neural response differences between behavioral outcomes. The criteria for accuracy measurement was uncompromising, as the entire sentence had to be repeated correctly to qualify as a correct response. This indicates that incorrect responses comprised a wide range of possible answers (e.g.; confidently but incorrectly repeating the sentence, with the presumption that an accurate response was given; repeating most of the sentence correctly and missing one word; simply saying, "I don't know", etc.). This method of coding neglects potential neural variations that may exist in such a wide variety of response types. Furthermore, vocalization of trial responses implicates other potential sources of artifact to fNIRS recordings (discussed in Section 4.4). In the future, we plan to explore use of silent response methods to report perception, such as closed set, forced choice methods, signal detection, or typing response on a keyboard (Faulkner et al., 2015).

Previous research shows evidence of both behavioral and neural adaptation during cognitive tasks (Guediche et al., 2015; Samuel and Kraljic, 2009). Specifically, behavioral adaptation to vocoded speech stimuli has shown significant improvements in perception accuracy after exposure and training with 30 vocoded sentences (Davis et al., 2005). Therefore, it is possible that activation to correct vocoded speech trials could be related either to participants' learning of the stimuli and/or could have adapted over time (Eisner et al., 2010).

Future research should investigate cortical associations with listening effort by using physiological measurements, such as pupillometry, to characterize fNIRS recordings. Simultaneous eye-tracking with fNIRS is exceptionally advantageous, as they both offer unrestrictive and convenient means of investigating cognitive function in typical and special populations. Neural measures that correlate with task performance are likely revealing cortical areas associated with behavioral outcome. However, performance measurements alone do not fully depict listening effort and emphasize the need for physiological markers to describe the cognitive demand encountered during effortful speech recognition. Preliminary results (Defenderfer et al., 2020) indicate that concurrent measurement of fNIRS and pupil data is feasible and reveal the potential to deepen our understanding of listening effort associated with simulated CI speech.

6. Conclusions

Overall, the current findings suggest that frontal and temporal cortices are differentially sensitive to the way speech signals are distorted. When speech is degraded with more natural forms of distortion (background noise), established neural channels in the frontal lobe enact top-down, attentional mechanisms to optimize speech recognition. However, this can be disrupted when the speech quality is deteriorated to the point where accurate perception is less likely, as cortical activation is significantly diminished during incorrect trials. Despite equivalent behavioral performance between speech-in-noise and vocoded speech conditions, cortical response patterns in NH adults suggest heavier reliance on temporal lobe function

during vocoded speech conditions. Diminished frontal cortex activity during vocoded speech conditions suggests that untrained listeners of vocoded stimuli do not as reliably recruit the same attentional mechanisms employed to resolve more natural forms of degraded speech. Finally, the correlations between speech perception scores and cortical activity motivate future research to examine individual differences more closely as the participants that performed better on the low-intelligibility condition differed in their reliance on cortical mechanisms from those indicated by group level activation.

Acknowledgements

References

Aarabi A, Osharina V, Wallois F, 2017. Effect of confounding variables on hemodynamic response function estimation using averaging and deconvolution analysis: an event-related NIRS study. Neuroimage 155, 25–49. doi:10.1016/J.NEUROIMAGE.2017.04.048. [PubMed: 28450140]

Adank P, Davis MH, Hagoort P, 2012. Neural dissociation in processing noise and accent in spoken language comprehension. Neuropsychologia 50 (1), 77–84. doi:10.1016/j.neuropsychologia.2011.10.024. [PubMed: 22085863]

Team Audacity, 2017. Audacity. The Name Audacity(R) Is a Registered Trademark of Dominic Mazzoni Retrieved from http://audacity.sourceforge.net/.

Belin P, Zilbovicius M, Crozier S, Thivard L, Fontaine A, Masure MC, Samson Y, 1998. Lateralization of speech and auditory temporal processing. J. Cogn. Neurosci 10 (4), 536–540. [PubMed: 9712682]

Bevilacqua F, Piguet D, Marquet P, Gross JD, Tromberg BJ, Depeursinge C, 1999. In vivo local determination of tissue optical properties: applications to human brain. Appl. Opt 38 (22), 4939. doi:10.1364/ao.38.004939. [PubMed: 18323984]

Billings CJ, Tremblay KL, Stecker GC, Tolin WM, 2009. Human evoked cortical activity to signal-to-noise ratio and absolute signal level. Hear. Res 254 (1–2), 15–24. 10.1016/j.heares.2009.04.002. [PubMed: 19364526]

Blamey PJ, Artieres F, Baskent D, Bergeron F, Beynon A, Burke E, … Lazard DS, 2013. Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 patients. Audiol. Neurootol 18 (1), 36–47. doi:10.1159/000343189. [PubMed: 23095305]

Blanco-Elorrieta E, Gwilliams L, Marantz A, Pylkkänen L, 2021. Adaptation to mispronounced speech: evidence for a prefrontal-cortex repair mechanism. Sci. Rep 11 (1), 1–12. doi:10.1038/s41598-020-79640-0. [PubMed: 33414495]

Bluestone AY, Abdoulaev G, Schmitz CH, Barbour RL, Hielscher AH, 2001. Threedimensional optical tomography of hemodynamics in the human head. Opt. Express 9 (6), 272. doi:10.1364/oe.9.000272. [PubMed: 19421298]

Boäx C, Baud L, Cosendai G, Sigrist A, Kós MI, Pelizzone M, 2006. Acoustic to electric pitch comparisons in cochlear implant subjects with residual hearing. JARO 7 (2), 110–124. doi:10.1007/s10162-005-0027-2. [PubMed: 16450213]

Braver TS, Cohen JD, Nystrom LE, Jonides J, Smith EE, Noll DC, 1997. A parametric study of prefrontal cortex involvement in human working memory. Neuroimage 62 (5), 49–62.

Brigadoi S, Ceccherini L, Cutini S, Scarpa F, Scatturin P, Selb J, … Cooper RJ, 2014. Motion artifacts in functional near-infrared spectroscopy: A comparison of motion correction techniques applied to real cognitive data. Neuroimage 85, 181–191. doi:10.1016/J.NEUROIMAGE.2013.04.082. [PubMed: 23639260]

Buxton RB, Wong EC, Frank LR, 1998. Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. Magn. Reson. Med 39 (6), 855–864. [PubMed: 9621908]

Calvetti D, Morigi S, Reichel L, Sgallari F, 2000. Tikhonov regularization and the L-curve for large discrete ill-posed problems. J. Comput. Appl. Math 123 (1–2), 423–446. doi:10.1016/S0377-0427(00)00414-3.

Carter CS, Braver TS, Barch DM, Botvinick MM, Noll DC, Cohen JD, 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 280 (5364), 747–749. doi:10.1126/science.280.5364.747. [PubMed: 9563953]

Chance B, Leigh JS, Miyake H, Smith DS, Nioka S, Greenfeld R, … Young M, 1988. Comparison of time-resolved and -unresolved measurements of deoxyhemoglobin in brain. Proc. Natl. Acad. Sci 85 (14), 4971–4975. doi:10.1073/pnas.85.14.4971. [PubMed: 3393526]

Chang EF, Niziolek CA, Knight RT, Nagarajan SS, Houde JF, 2013. Human cortical sensorimotor network underlying feedback control of vocal pitch. Proc. Natl. Acad. Sci. U S A 110 (7), 2653–2658. doi:10.1073/pnas.1216827110. [PubMed: 23345447]

Chen G, Adleman NE, Saad ZS, Leibenluft E, and Cox RW (2014). Applications of multivariate modeling to neuroimaging group analysis: a comprehensive alternative to univariate general linear model.

Chen L, Sandmann P, Thorne JD, Bleichner MG, Debener S, 2016. Cross-modal functional reorganization of visual and auditory cortex in adult cochlear implant users identified with fNIRS. Neural Plast. 2016. doi:10.1155/2016/4382656, Research.

Christoffels IK, Formisano E, Schiller NO, 2007. Neural correlates of verbal feedback processing: an fMRI study employing overt speech. Hum. Brain Mapp 28 (9), 868–879. doi:10.1002/hbm.20315. [PubMed: 17266104]

Cohen JD, Forman SD, Braver TS, Casey BJ, Servan-Schreiber D, Noll DC, 1994. Activation of the prefrontal cortex in a nonspatial working memory task with functional MRI. Hum. Brain Mapp 1 (4), 293–304. doi:10.1002/hbm.460010407. [PubMed: 24591198]

Cohen MX, Ridderinkhof KR, Haupt S, Elger CE, Fell J, 2008. Medial frontal cortex and response conflict: evidence from human intracranial EEG and medial frontal cortex lesion. Brain Res. 1238, 127–142. doi:10.1016/j.brainres.2008.07.114. [PubMed: 18760262]

Cox RW, Chen G, Glen DR, Reynolds RC, Taylor PA, 2017. FMRI clustering in AFNI: false-positive rates redux. Brain Connectivity 7 (3), 152–171. doi:10.1089/brain.2016.0475. [PubMed: 28398812]

Custo A, Wells WM, Barnett AH, Hillman EMC, Boas DA, 2006. Effective scattering coefficient of the cerebral spinal fluid in adult head models for diffuse optical imaging. Appl. Opt. 45 (19), 4747–4755. doi:10.1364/AO.45.004747. [PubMed: 16799690]

Dale AM, 1999. Optimal experimental design for event-related fMRI. Hum. Brain Mapp 8 (2–3), 109–114. [PubMed: 10524601]

Davis MH, Ford MA, Kherif F, Johnsrude IS, 2011. Does semantic context benefit speech understanding through "top-down" processes? Evidence from time-resolved sparse fMRI. J. Cogn. Neurosci 23 (12), 3914–3932. doi:10.1162/jocn_a_00084. [PubMed: 21745006]

Davis MH, Johnsrude IS, 2003. Hierarchical processing in spoken language comprehension. J. Neurosci 23 (8), 3423. Retrieved from http://www.jneurosci.org/content/23/8/3423.abstract. [PubMed: 12716950]

Davis MH, Johnsrude IS, Hervais-Adelman A, Taylor K, McGettigan C, 2005. Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. J. Experim. Psychol 134 (2), 222–241. doi:10.1037/0096-3445.134.2.222.

Defenderfer J, Kerr-German A, Hedrick M, Buss AT, 2017. Investigating the role of temporal lobe activation in speech perception accuracy with normal hearing adults: an event-related fNIRS study. Neuropsychologia 106 (September), 31–41. doi:10.1016/j.neuropsychologia.2017.09.004. [PubMed: 28888891]

Defenderfer J, McGarr M, Tas AC, 2020. Change in pupil size reveals impact of simulated-cochlear implant speech on listening effort. J. Vision 20 (1409). doi:10.1167/jov.20.11.1409.

Demb JB, Desmond JE, Wagner AD, Vaidya CJ, Glover GH, Gabrieli JDE, 1995. Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. J. Neurosci 15 (9), 5870–5878. doi:10.1523/jneurosci.15-09-05870.1995. [PubMed: 7666172]

Dimitrijevic A, Smith ML, Kadis DS, Moore DR, 2019. Neural indices of listening effort in noisy environments. Sci. Rep 9 (1), 11278. doi:10.1038/s41598-019-47643-1. [PubMed: 31375712]

Dosenbach NU, Fair DA, Cohen AL, Schlaggar BL, Petersen SE, 2008. A dualnetworks architecture of top-down control. Trends Cogn. Sci 12 (3), 99–105. doi:10.1016/j.tics.2008.01.001. [PubMed: 18262825]

Du Y, Buchsbaum BR, Grady CL, Alain C, 2014. Noise differentially impacts phoneme representations in the auditory and speech motor systems. Proc. Natl. Acad. Sci. U S A 111 (19), 7126–7131. doi:10.1073/pnas.1318738111. [PubMed: 24778251]

Eckert MA, Teubner-Rhodes S, Vaden KIJ, 2016. Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. Ear Hear. 37, 101S–110S. doi:10.1097/aud.0000000000000300. [PubMed: 27355759]

Eggebrecht AT, Culver JP, 2019. Neurodot: an extensible Matlab toolbox for streamlined optical functional mapping. In: Optics InfoBase Conference Papers, Part F142 - (7 2019), pp. 2019–2022. doi:10.1117/12.2527164.

Eggebrecht AT, Ferradal SL, Robichaux-Viehoever A, Hassanpour MS, Dehghani H, Snyder AZ, … Culver JP, 2014. Mapping distributed brain function and networks with diffuse optical tomography. Nat. Photonics 8 (6), 448–454. doi:10.1038/nphoton.2014.107.Mapping. [PubMed: 25083161]

Eggebrecht AT, White BR, Ferradal SL, Chen C, Zhan Y, Snyder AZ, … Culver JP, 2012. A quantitative spatial comparison of high-density diffuse optical tomography and fMRI cortical mapping. Neuroimage 61 (4), 1120–1128. doi:10.1016/j.neuroimage.2012.01.124. [PubMed: 22330315]

Eisner F, McGettigan C, Faulkner A, Rosen S, Scott SK, 2010. Inferior frontal gyrus activation predicts individual differences in perceptual learning of cochlear-implant simulations. J. Neurosci 30 (21), 7179. Retrieved from http://www.jneurosci.org/content/30/21/7179.abstract. [PubMed: 20505085]

Evans S, Kyong JS, Rosen S, Golestani N, Warren JE, McGettigan C, … Scott SK, 2014. The pathways for intelligible speech: multivariate and univariate perspectives. Cereb. Cortex 24 (9), 2350–2361. doi:10.1093/cercor/bht083. [PubMed: 23585519]

Fang Q, Boas DA, 2009. Monte carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units. Opt. Express 17 (22), 20178. doi:10.1364/OE.17.020178. [PubMed: 19997242]

Faulkner KF, Tamati TN, Gilbert JL, and Pisoni DB (2015). List Equivalency of PRESTO for the Evaluation of Speech Recognition, 26(6), 582–594. 10.1002/stem.1868.Human

Forbes SH, Wijeakumar S, Eggebrecht AT, Magnotta VA, Spencer JP, 2021. A processing pipeline for image reconstructed fNIRS analysis using both MRI templates and individual anatomy. BioRxiv doi:10.1101/2021.01.14.426719, 2021.01.14.426719.

Francis AL, MacPherson MK, Chandrasekaran B, Alvar AM, 2016. Autonomic nervous system responses during perception of masked speech may reflect constructs other than subjective listening effort. Front. Psychol 7 (3), 1–15. doi:10.3389/fpsyg.2016.00263. [PubMed: 26858668]

Friederici AD, 2011. The brain basis of language processing: from structure to function. Physiol. Rev 91 (4), 1357–1392. doi:10.1152/physrev.00006.2011. [PubMed: 22013214]

Friederici AD, Rüschemeyer SA, Hahne A, Fiebach CJ, 2003. The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. Cereb. Cortex 13 (2), 170–177. doi:10.1093/cercor/13.2.170. [PubMed: 12507948]

Gagnon L, Perdue K, Greve DN, Goldenholz D, Kaskhedikar G, Boas DA, 2011. Improved recovery of the hemodynamic response in diffuse optical imaging using short optode separations and state-space modeling. Neuroimage 56 (3), 1362–1371. doi:10.1016/j.neuroimage.2011.03.001. [PubMed: 21385616]

Gervain J, Geffen MN, 2019. Efficient neural coding in auditory and speech perception Judit. Physiol. Behav 42 (1), 56–65. doi:10.1016/j.tins.2018.09.004.Efficient.

Giraud AL, Kell C, Thierfelder C, Sterzer P, Russ MO, Preibisch C, Kleinschmidt A, 2004. Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. Cereb. Cortex 14 (3), 247–255. doi:10.1093/cercor/bhg124. [PubMed: 14754865]

Godefroy O, Rousseaux M, 1996. Divided and focused attention in patients with lesion of the prefrontal cortex. Brain Cogn. 30 (2), 155–174. doi:10.1006/brcg.1996.0010. [PubMed: 8811992]

Golestani N, Hervais-Adelman A, Obleser J, Scott SK, 2013. Semantic versus perceptual interactions in neural processing of speech-in-noise. Neuroimage 79, 52–61. doi:10.1016/j.neuroimage.2013.04.049. [PubMed: 23624171]

Goupell MJ, Draves GT, Litovsky RY, 2020. Recognition of vocoded words and sentences in quiet and multi-talker babble with children and adults. PLoS One 15 (12 12), 1–11. doi:10.1371/journal.pone.0244632.

Gow D Jr., 2012. The cortical organization of lexical knowledge: a dual lexicon model of spoken language processing. Brain Lang. 273–288 (3), 1–7. doi:10.1016/j.bandl.2012.03.005.The.

Graves WW, Grabowski TJ, Mehta S, Gupta P, 2008. The left posterior superior temporal gyrus participates specifically in accessing lexical phonology. J. Cogn. Neurosci 20 (9), 1698–1710. doi:10.1162/jocn.2008.20113. [PubMed: 18345989]

Guediche S, Holt LL, Laurent P, Lim SJ, Fiez JA, 2015. Evidence for cerebellar contributions to adaptive plasticity in speech perception. Cereb. Cortex 25 (7), 1867–1877. doi:10.1093/cercor/bht428. [PubMed: 24451660]

Hassanpour MS, White BR, Eggebrecht AT, Ferradal SL, Snyder AZ, Culver JP, 2014. Statistical analysis of high density diffuse optical tomography. Neuroimage 85 (1), 104–116. doi:10.1016/j.neuroimage.2013.05.105.Statistical. [PubMed: 23732886]

Hazeltine E, Poldrack RA, Gabrieli JDE, 2000. Neural Activation During Response Competition 12 (2), 118–129.

Hervais-Adelman A, Carlyon RP, Johnsrude IS, Davis MH, 2012. Brain regions recruited for the effortful comprehension of noise-vocoded words. Lang. Cognit. Processes 27 (7–8), 1145–1166. doi:10.1080/01690965.2012.662280.

Hickok G, Poeppel D, 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92 (1–2), 67–99. doi:10.1016/j.cognition.2003.10.011. [PubMed: 15037127]

Hickok G, Poeppel D, 2007. The cortical organization of speech processing. Nat. Rev. Neurosci 8 (5), 393–402. doi:10.1038/nrn2113. [PubMed: 17431404]

Hirano S, Naito Y, Kojima H, Honjo I, Inoue M, Shoji K, … Konishi J, 2000. Functional differentiation of the auditory association area in prelingually deaf subjects. Auris Nasus Larynx 27 (4), 303–310. [PubMed: 10996488]

Huppert TJ, 2016. Commentary on the statistical properties of noise and its implication on general linear models in functional near-infrared spectroscopy. Neurophotonics 3 (1), 010401. doi:10.1117/1.nph.3.1.010401. [PubMed: 26989756]

Huppert TJ, Diamond SG, Franceschini MA, Boas DA, 2009. HomER: A review of time-series analysis methods for near-infrared spectroscopy of the brain. Appl. Opt 48 (10), D280–D298. [PubMed: 19340120]

Issard C, Gervain J, 2018. Variability of the hemodynamic response in infants: influence of experimental design and stimulus complexity. Develop. Cognit. Neurosci 33 (2 2017), 182–193. doi:10.1016/j.dcn.2018.01.009.

Jahani S, Setarehdan SK, Boas DA, and Yücel MA (2018). Motion artifact detection and correction in functional near-infrared spectroscopy: a new hybrid method based on spline interpolation method and Savitzky–Golay filtering, 5, 15003–15011. Retrieved from 10.1117/1.NPh.5.1.015003

James C, Blamey PJ, Shallop JK, Incerti PV, 2001. Contralateral masking in cochlear implant users with residual hearing in the non-implanted ear. Audiol. Neurotol (6) 87–97. Retrieved from https://search.proquest.com/docview/220297257?accountid=12834.

Kerns JG, 2006. Anterior cingulate and prefrontal cortex activity in an FMRI study of trial-to-trial adjustments on the Simon task. Neuroimage 33 (1), 399–405. doi:10.1016/j.neuroimage.2006.06.012. [PubMed: 16876434]

Kozou H, Kujala T, Shtyrov Y, Toppila E, Starck J, Alku P, Naatanen R, 2005. The effect of different noise types on the speech and non-speech elicited mismatch negativity. Hear. Res 199 (1–2), 31–39. doi:10.1016/j.heares.2004.07.010. [PubMed: 15574298]

Lawler CA, Wiggins IM, Dewey RS, Hartley DEH, 2015. The use of functional near-infrared spectroscopy for measuring cortical reorganisation in cochlear implant users: a possible predictor of variable speech outcomes? Cochlear. Implants Int 16 (S1), S30–S32. doi:10.1179/1467010014Z.000000000230. [PubMed: 25614264]

Lawrence RJ, Wiggins IM, Anderson CA, Davies-Thompson J, Hartley DEH, 2018. Cortical correlates of speech intelligibility measured using functional near-infrared spectroscopy (fNIRS). Hear. Res 370, 53–64. doi:10.1016/j.heares.2018.09.005. [PubMed: 30292959]

Lazard DS, Vincent C, Venail F, Van de Heyning P, Truy E, Sterkers O, … Blamey PJ, 2012. Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: a new conceptual model over time. PLoS One 7 (11), e48739. doi:10.1371/journal.pone.0048739. [PubMed: 23152797]

Leonard MK, Baud MO, Sjerps MJ, Chang EF, 2016. Perceptual restoration of masked speech in human cortex. Nat. Commun 7, 1–9. doi:10.1038/ncomms13619.

Macherey O, Carlyon RP, 2014. Cochlear implants. Curr. Biol 24 (18), R878–R884. doi:10.1016/j.cub.2014.06.053. [PubMed: 25247367]

Majerus S, Van Der Linden M, Collette F, Laureys S, Poncelet M, Degueldre C, … Salmon E, 2005. Modulation of brain activity during phonological familiarization. Brain Lang. 92 (3), 320–331. doi:10.1016/j.bandl.2004.07.003. [PubMed: 15721964]

Mattys SL, Brooks J, Cooke M, 2009. Recognizing speech under a processing load: Dissociating energetic from informational factors. Cognit. Psychol 59 (3), 203–243. doi:10.1016/j.cogpsych.2009.04.001. [PubMed: 19423089]

Mattys SL, Davis MH, Bradlow AR, and Scott SK (2012). Speech recognition in adverse conditions: A review. Language and Cognitive Processes, 27(7–8), 953–978. 10.1080/01690965.2012.705006

Miles K, McMahon C, Boisvert I, Ibrahim R, de Lissa P, Graham P, Lyxell B, 2017. Objective assessment of listening effort: coregistration of pupillometry and EEG. Trends Hearing 21, 1–13. doi:10.1177/2331216517706396.

Moberly AC, Lowenstein JH, and Nittrouer S (2016). Word recognition variability with cochlear implants: "perceptual attention" versus "auditory sensitivity," 37(1), 14–36. 10.1097/AUD.0000000000000204

Moberly AC, Lowenstein JH, Tarr E, Caldwell-Tarr A, Welling DB, Shahin AJ, Nittrouer S, 2014. Do adults with cochlear implants rely on different acoustic cues for phoneme perception than adults with normal hearing? J. Speech Lang. Hear. Res 57, 566–582. doi:10.1044/2014. [PubMed: 24686722]

Molavi B, Dumont GA, 2012. Wavelet-based motion artifact removal for functional near-infrared spectroscopy. Physiol. Meas 33 (2), 259. Retrieved from http://stacks.iop.org/0967-3334/33/i=2/a=259. [PubMed: 22273765]

Narain C, Scott SK, Wise RJS, Rosen S, Leff A, Iversen SD, Matthews PM, 2003. Defining a left-lateralized response specific to intelligible speech using fMRI. Cereb. Cortex 13 (12), 1362–1368. [PubMed: 14615301]

Nilsson M, Soli SD, Sullivan JA, 1994. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J. Acoust. Soc. Am 95 (2), 1085–1099. [PubMed: 8132902]

Obleser J, Eisner F, Kotz SA, 2008. Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. J. Neurosci 28 (32), 8116–8123. doi:10.1523/jneurosci.1290-08.2008. [PubMed: 18685036]

Obleser J, Kotz SA, 2010. Expectancy constraints in degraded speech modulate the language comprehension network. Cereb. Cortex 20 (3), 633–640. doi:10.1093/cercor/bhp128. [PubMed: 19561061]

Obleser J, Wise RJS, Dresner MA, Scott SK, 2007. Functional Integration across Brain Regions Improves Speech Perception under Adverse Listening Conditions. J. Neurosci 27 (9), 2283–2289. doi:10.1523/JNEUROSCI.4663-06.2007. [PubMed: 17329425]

Ohlenforst B, Wendt D, Kramer SE, Naylor G, Zekveld AA, Lunner T, 2018. Impact of SNR, masker type and noise reduction processing on sentence recognition performance and listening effort as

indicated by the pupil dilation response. Hear. Res 365, 90–99. doi:10.1016/j.heares.2018.05.003. [PubMed: 29779607]

Olds C, Pollonini L, Abaya H, Larky J, Loy M, Bortfeld H, … Oghalai JS, 2016. Cortical Activation Patterns Correlate with Speech Understanding After Cochlear Implantation. Ear Hear. 37 (3), e160–e172. doi:10.1097/aud.0000000000000258. [PubMed: 26709749]

Pals C, Sarampalis A, Ba kent D, 2012. Listening Effort With Cochlear Implant Simulations. J. Speech Lang. Hear. Res 56 (4), 1075–1084. doi:10.1044/1092-4388(2012/12-0074. [PubMed: 23275424]

Peelle JE, 2018. Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. Ear Hear. 39 (2), 204–214. doi:10.1097/AUD.0000000000000494. [PubMed: 28938250]

Petersen B, Gjedde A, Wallentin M, Vuust P, 2013. Cortical plasticity after cochlear implantation. Neural Plast. 2013. doi:10.1155/2013/318521.

Pichora-Fuller MK, 2016. How social psychological factors may modulate auditory and cognitive functioning during listening. Ear Hear. 37, 92S–100S. doi:10.1097/aud.0000000000000323. [PubMed: 27355776]

Poeppel D, Idsardi WJ, Van Wassenhove V, 2008. Speech perception at the interface of neurobiology and linguistics. Philosoph. Trans. R. Soc. B 363 (1493), 1071–1086. doi:10.1098/rstb.2007.2160.

Poldrack RA, Temple E, Protopapas A, Nagarajan S, Tallal P, Merzenich M, Gabrieli JD, 2001. Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI. J. Cogn. Neurosci 13 (5), 687–697. doi:10.1162/089892901750363235. [PubMed: 11506664]

Pollonini L, Olds C, Abaya H, Bortfeld H, Beauchamp MS, Oghalai JS, 2014. Auditory cortex activation to natural speech and simulated cochlear implant speech measured with functional near-infrared spectroscopy. Hear. Res 309, 84–93. doi:10.1016/j.heares.2013.11.007. [PubMed: 24342740]

Quaresima V, Bisconti S, Ferrari M, 2012. A brief review on the use of functional near-infrared spectroscopy (fNIRS) for language imaging studies in human newborns and adults. Brain Lang. 121 (2), 79–89. doi:10.1016/j.bandl.2011.03.009. [PubMed: 21507474]

Ridderinkhof KR, Van Den Wildenberg WPM, Segalowitz SJ, Carter CS, 2004. Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. Brain Cogn. 56 (2 SPEC. ISS.), 129–140. doi:10.1016/j.bandc.2004.09.016. [PubMed: 15518930]

Rodd JM, Davis MH, Johnsrude IS, 2005. The neural mechanisms of speech comprehension: FMRI studies of semantic ambiguity. Cereb. Cortex 15 (8), 1261–1269. doi:10.1093/cercor/bhi009. [PubMed: 15635062]

Rönnberg J, Lunner T, Zekveld AA, Sorqvist P, Danielsson H, Lyxell B, … Rudner M, 2013. The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. Front. Syst. Neurosci 7, 31. doi:10.3389/fnsys.2013.00031. [PubMed: 23874273]

Rushworth MF, Buckley MJ, Behrens TE, Walton ME, Bannerman DM, 2007. Functional organization of the medial frontal cortex. Curr. Opin. Neurobiol 17 (2), 220–227. doi:10.1016/j.conb.2007.03.001. [PubMed: 17350820]

Rutten S, Santoro R, Hervais-Adelman A, Formisano E, Golestani N, 2019. Task-relevant acoustic information. Nature Human Behav. 3 (9), 974–987 Retrieved from 10.1038/s41562-019-0648-9. [PubMed: 31285622]

Saliba J, Bortfeld H, Levitin DJ, Oghalai JS, 2016. Functional near-infrared spectroscopy for neuroimaging in cochlear implant recipients. Hear. Res doi:10.1016/j.heares.2016.02.005.

Samuel AG, and Kraljic T (2009). Perceptual learning for speech, 71(7), 1207–1218. 10.3758/APP

Sandmann P, Plotz K, Hauthal N, de Vos M, Schonfeld R, Debener S, 2015. Rapid bilateral improvement in auditory cortex activity in postlingually deafened adults following cochlear implantation. Clin. Neurophysiol. 126 (3), 594–607. doi:10.1016/j.clinph.2014.06.029. [PubMed: 25065298]

Sato T, Nambu I, Takeda K, Aihara T, Yamashita O, Isogaya Y, … Osu R, 2016. Reduction of global interference of scalp-hemodynamics in functional near-infrared spectroscopy using short distance probes. Neuroimage 141, 120–132. doi:10.1016/j.neuroimage.2016.06.054. [PubMed: 27374729]

Savitzky A, Golay MJE, 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem 36 (8), 1627–1639. doi:10.1021/ac60214a047.

Schecklmann M, Ehlis A-C, Plichta MM, Fallgatter AJ, 2010. Influence of muscle activity on brain oxygenation during verbal fluency assessed with functional near-infrared spectroscopy. Neuroscience 171 (2), 434–442. doi:10.1016/j.neuroscience.2010.08.072. [PubMed: 20858532]

Schecklmann M, Mann A, Langguth B, Ehlis A-C, Fallgatter AJ, Haeussinger FB, 2017. The temporal muscle of the head can cause artifacts in optical imaging studies with functional near-infrared spectroscopy. Front. Human Neurosci doi:10.3389/fnhum.2017.00456.

Schnupp J, 2006. Auditory filters, features, and redundant representations. Neuron 51 (3), 278–280. doi:10.1016/j.neuron.2006.07.015. [PubMed: 16880121]

Scholkmann F, Gerber U, Wolf M, Wolf U, 2013. End-tidal CO2: an important parameter for a correct interpretation in functional brain studies using speech tasks. Neuroimage 66, 71–79 10.1016/j.neuroimage.2012.10.025. [PubMed: 23099101]

Scholkmann F, Gerber U, Wolf M, Wolf U, 2013. End-tidal CO2: an important parameter for a correct interpretation in functional brain studies using speech tasks. Neuroimage 66, 71–79 10.1016/j.neuroimage.2012.10.025. [PubMed: 23099101]

Scholkmann F, Spichtig S, Muehlemann T, Wolf M, 2010. How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation. Physiol. Meas. 31 (5), 649–662. doi:10.1088/0967-3334/31/5/004. [PubMed: 20308772]

Scholkmann F, Wolf M, Wolf U, 2013. The effect of inner speech on arterial CO2 and cerebral hemodynamics and oxygenation: a functional NIRS study. In: Oxygen Transport to Tissue XXXV. Springer New York, New York, NY, pp. 81–87. doi:10.1007/978-1-4614-7411-1_12.

Scott SK, Blank CC, Rosen S, Wise RJS, 2000. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123 Pt 12, 2400–2406. [PubMed: 11099443]

Scott SK, Rosen S, Wickham L, Wise RJS, 2004. A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. J. Acoust. Soc. Am 115 (2), 813–821. [PubMed: 15000192]

Shannon RV, Zeng F-G, Kamath V, Wygonski J, Ekelid M, 1995. Speech recognition with primarily temporal cues. Science 270 (5234), 303–304. [PubMed: 7569981]

Sheldon S, Pichora-Fuller MK, Schneider BA, 2008. Priming and sentence context support listening to noise-vocoded speech by younger and older adults. J. Acoust. Soc. Am 123 (1), 489–499. doi:10.1121/1.2783762. [PubMed: 18177176]

Shenhav A, Botvinick MM, Cohen JD, 2013. The expected value of control: An integrative theory of anterior cingulate cortex function. Neuron 79 (2), 217–240. doi:10.1016/j.neuron.2013.07.007. [PubMed: 23889930]

Steinbrink J, Villringer A, Kempf F, Haux D, Boden S, Obrig H, 2006. Illuminating the BOLD signal: combined fMRI-fNIRS studies. Magn. Reson. Imaging 24 (4), 495–505. doi:10.1016/j.mri.2005.12.034. [PubMed: 16677956]

Strangman G, Boas DA, Sutton JP, 2002. Non-invasive neuroimaging using near-infrared light. Biol. Psychiatry 52 (7), 679–693. [PubMed: 12372658]

Tikhonov AN, 1963. On the solution of ill-posed problems and the method of regularization. Dokl. Akad. Nauk 151 (3), 501–504.

Tisdall MM, Taylor C, Tachtsidis I, Leung TS, Elwell CE, Smith M, 2009. The effect on cerebral tissue oxygenation index of changes in the concentrations of inspired oxygen and end-tidal carbon dioxide in healthy adult volunteers. Anesth. Analg 109 (3), 906–913. doi:10.1213/ane.0b013e3181aedcdc. [PubMed: 19690266]

Vaden KI, Kuchinsky SE, Cute SL, Ahlstrom JB, Dubno JR, Eckert MA, 2013. The cingulo-opercular network provides word-recognition benefit. J. Neurosci 33 (48), 18979–18986. doi:10.1523/jneurosci.1417-13.2013. [PubMed: 24285902]

Ventura MI, Nagarajan SS, Houde JF, 2009. Speech target modulates speaking induced suppression in auditory cortex. BMC Neurosci. 10, 58. doi:10.1186/1471-2202-10-58. [PubMed: 19523234]

Villringer A, Planck J, Hock C, Schleinkofer L, Dirnagl U, 1993. Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults. Neurosci. Lett 154 (1–2), 101–104. [PubMed: 8361619]

Volkening N, Unni A, Löffler BS, Fudickar S, Rieger JW, Hein A, 2016. Characterizing the influence of muscle activity in fNIRS brain activation measurements. IFAC-PapersOnLine 49 (11), 84–88. 10.1016/j.ifacol.2016.08.013.

Wheelock MD, Culver JP, Eggebrecht AT, 2019. High-density diffuse optical tomography for imaging human brain function. Rev. Sci. Instrum (5) 90. doi:10.1063/1.5086809.

Wijayasiri P, Hartley DEH, Wiggins IM, 2017. Brain activity underlying the recovery of meaning from degraded speech: A functional near-infrared spectroscopy (fNIRS) study. Hear. Res doi:10.1016/j.heares.2017.05.010.

Wijeakumar S, Huppert TJ, Magnotta VA, Buss AT, Spencer JP, 2017. Validating an image-based fNIRS approach with fMRI and a working memory task. Neuroimage 147, 204–218. doi:10.1016/J.NEUROIMAGE.2016.12.007. [PubMed: 27939793]

Wijeakumar S, Kumar A, Delgado Reyes LM, Tiwari M, Spencer JP, 2019. Early adversity in rural India impacts the brain networks underlying visual working memory. Dev. Sci 22 (5), 1–15. doi:10.1111/desc.12822.

Wild CJ, Yusuf A, Wilson DE, Peelle JE, Davis MH, Johnsrude IS, 2012. Effortful listening: the processing of degraded speech depends critically on attention. J. Neurosci 32 (40), 14010–14021. doi:10.1523/jneurosci.1528-12.2012. [PubMed: 23035108]

Winn MB, Edwards JR, Litovsky RY, 2015. The impact of auditory spectral resolution on listening effort revealed by pupil dilation. Ear Hear. (Vol. 36) doi:10.1097/AUD.0000000000000145.The.

Wong PCM, Uppunda AK, Parrish TB, Dhar S, 2008. Cortical mechanisms of speech perception in noise. J. Speech Lang. Hear. Res 51 (4), 1026–1041. doi:10.1044/1092-4388(2008/075. [PubMed: 18658069]

Zekveld AA, Heslenfeld DJ, Johnsrude IS, Versfeld NJ, Kramer SE, 2014. The eye as a window to the listening brain: neural correlates of pupil size as a measure of cognitive listening load. Neuroimage 101, 76–86. doi:10.1016/j.neuroimage.2014.06.069. [PubMed: 24999040]

Zekveld AA, Kramer SE, 2014. Cognitive processing load across a wide range of listening conditions: insights from pupillometry. Psychophysiology 51 (3), 277–284. doi:10.1111/psyp.12151. [PubMed: 24506437]

Zhou X, Seghouane AK, Shah A, Innes-Brown H, Cross W, Litovsky R, McKay CM, 2018. Cortical speech processing in postlingually deaf adult cochlear implant users, as revealed by functional near-infrared spectroscopy. Trends Hear. 22, 1–18. doi:10.1177/2331216518786850.

Morais, Zimeo A,G, Scholkmann F, Balardin JB, Furucho RA, de Paula RCV, Biazoli CE, Sato JR, 2017. Non-neuronal evoked and spontaneous hemodynamic changes in the anterior temporal region of the human head may lead to misinterpretations of functional near-infrared spectroscopy signals. Neurophotonics 5 (01), 1. doi:10.1117/1.nph.5.1.011002.

Zimmermann BB, Roche-Labarbe N, Surova A, Boas DA, Wolf M, Grant PE, Franceschini MA, 2012. The confounding effect of systemic physiology on the hemodynamic response in newborns. In: Oxygen Transport to Tissue XXXIII. Springer New York, New York, NY, pp. 103–109.
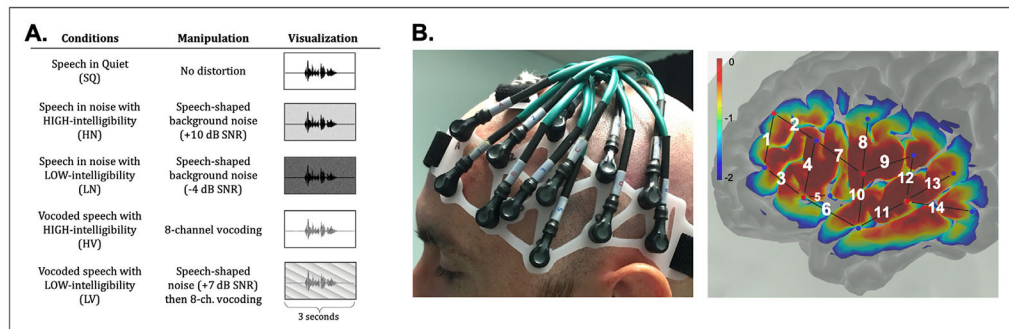
Fig. 1.
A. Abbreviations and descriptions of task conditions. **B.** Custom headpiece positioned on representative participant (left). Sensitivity profile and projection of fNIRS probes onto cortical surface (right). Red and blue dots represent source lights and detectors, respectively. NIRS channels are labeled with white numbers (channel 5 is the short separation channel). The color scale indicates relative sensitivity to neural activation on a logarithmic scale.
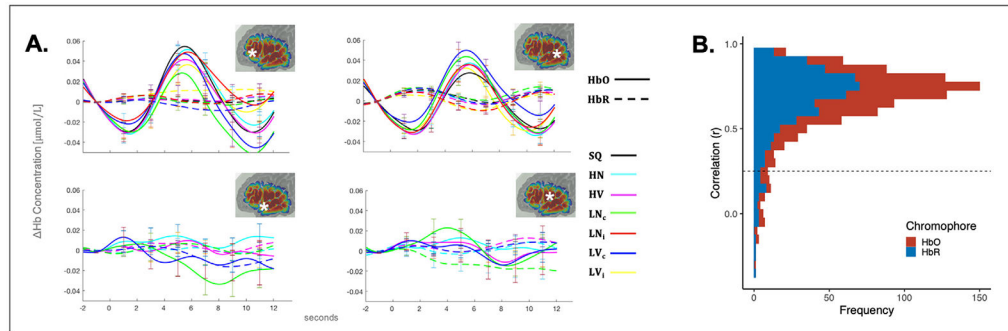
Fig. 2.
A. Line plots of channel-wise time series data for all conditions from a representative frontal channel (channel 3, top left) and a representative temporal channel (channel 13, top right). Examples of non-canonical/inverted responses for conditions from the Intelligibility ANOVA are plotted from channel 6 (bottom left) and channel 9 (bottom right). The approximate location of each channel in relation to the probe configuration are denoted by white asterisks within the insets in the upper right hand corner of each plot. See Fig. 1A for condition abbreviations. Error bars represent the standard error of the mean. **B.** Histogram plotting the frequency of correlations between channel data and image-based data.
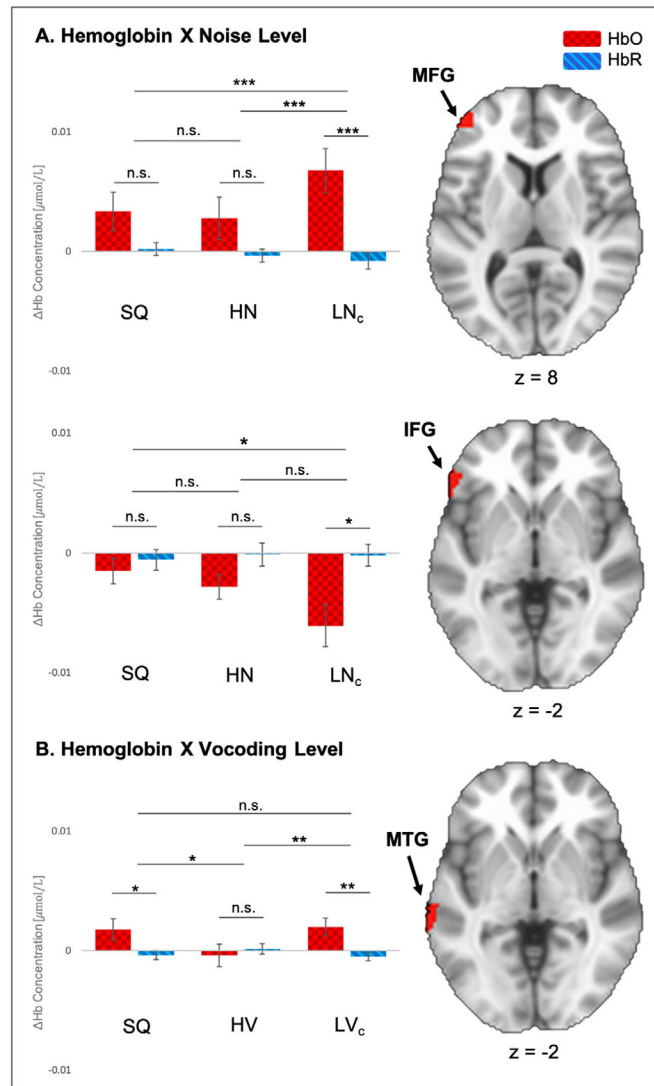
Fig. 3.
Results of ANOVAs A (Hemoglobin (2) X Noise Level (3)) and B (Hemoglobin (2) X Vocoding Level (3)). **A.** *Hemoglobin X Noise Level interaction* – Top bar plot shows average changes in HbO and HbR (ΔHb) during SQ, HN, and LN_c conditions for the interaction in the MFG ($z = 8$); bottom bar plot shows the second interaction of this type in the IFG ($z = -2$). **B.** Hemoglobin *X Vocoding Level interaction* – Bar plot shows average changes in HbO and HbR during SQ, HV, and LV_c conditions for the interaction in the MTG ($z = -2$). HbO, oxyhemoglobin (red squares); HbR, deoxyhemoglobin (blue stripes); IFG, inferior frontal gyrus; MFG, middle frontal gyrus; MTG, middle temporal gyrus. Error bars represent standard error of the mean. Significance was adjusted for multiple comparisons and is marked as follows: * p $\leq$ 0.05, * * p $\leq$ 0.01; * * * p $\leq$ 0.001; n.s., not significant.
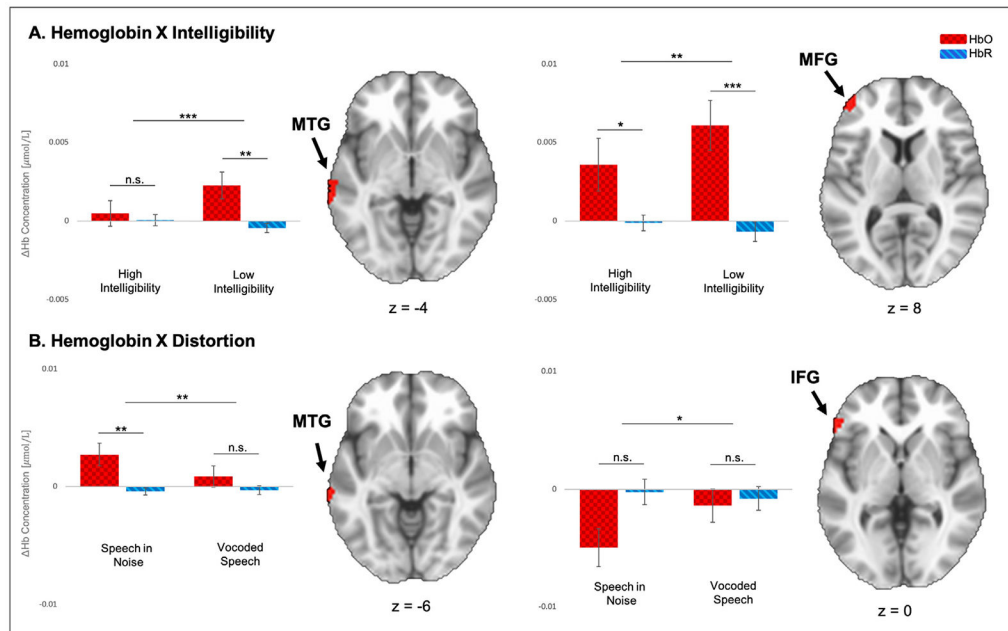
Fig. 4.

Results of ANOVA C (Hemoglobin (2) X Distortion (2) X Intelligibility (2)). **A.**
Hemoglobin X Intelligibility interaction – Left bar plot contrasts average changes in HbO
and HbR (ΔHb) between high- (HN, HV) and low-intelligibility (LN_c, LV_c) trials for the
interaction in the MTG ($z = -4$); right bar plot shows the second interaction of this
type in MFG ($z = 8$). **B.** *Hemoglobin X Distortion interaction* – Left bar plot contrasts
average changes in HbO and HbR between speech-in-noise (HN, LN_c) and vocoded speech
(HV, LV_c) trials for the interaction in the IFG ($z = 0$); right bar plot shows the second
interaction of this type in the MTG ($z = -6$). HbO, oxyhemoglobin (red squares); HbR,
deoxyhemoglobin (blue stripes); IFG, inferior frontal gyrus; MFG, middle frontal gyrus;
MTG, middle temporal gyrus. Error bars represent standard error of the mean. Significance
is marked as follows: *p $\leq$ 0.05, **p $\leq$ 0.01; ***p $\leq$ 0.001; n.s., not significant.
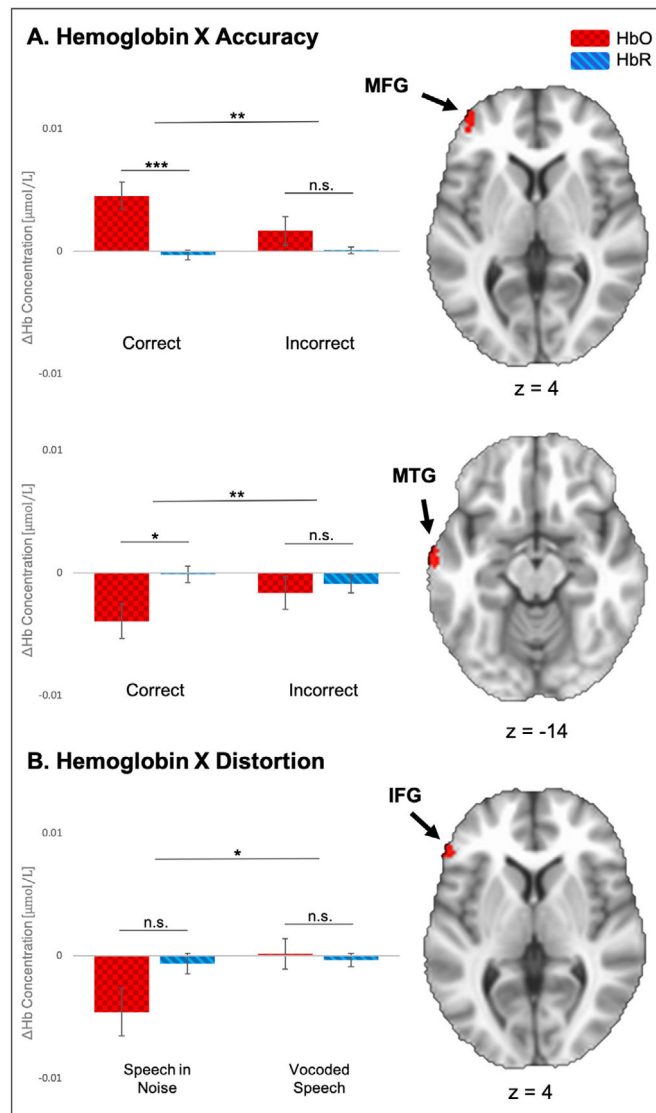
Fig. 5.

Results of ANOVA D, Hemoglobin (2) X Distortion (2) X Accuracy (2) **A.** *Hemoglobin X Accuracy interaction* –Top bar plot contrast average changes in HbO and HbR (Hb) between correct (LN_c, LV_c) and incorrect (LN_i, LV_i) trials for the interaction in the MFG (z = 4); bottom bar plot shows the second interaction of that type in the MTG (z = −14). **B.** *Hemoglobin X Distortion interaction* –bar plot contrasts average changes in HbO and HbR between speech-in-noise (HN, LN_c) and vocoded speech (HV, LV_c) trials for the interaction in the IFG (z = 4). HbO, oxyhemoglobin (red squares); HbR, deoxyhemoglobin (blue stripes); IFG, inferior frontal gyrus; MFG, middle frontal gyrus; MTG, middle temporal gyrus. Error bars represent standard error of the mean. Significance is marked as follows: *p 0.05, **p 0.01; *** p 0.001; n.s., not significant.
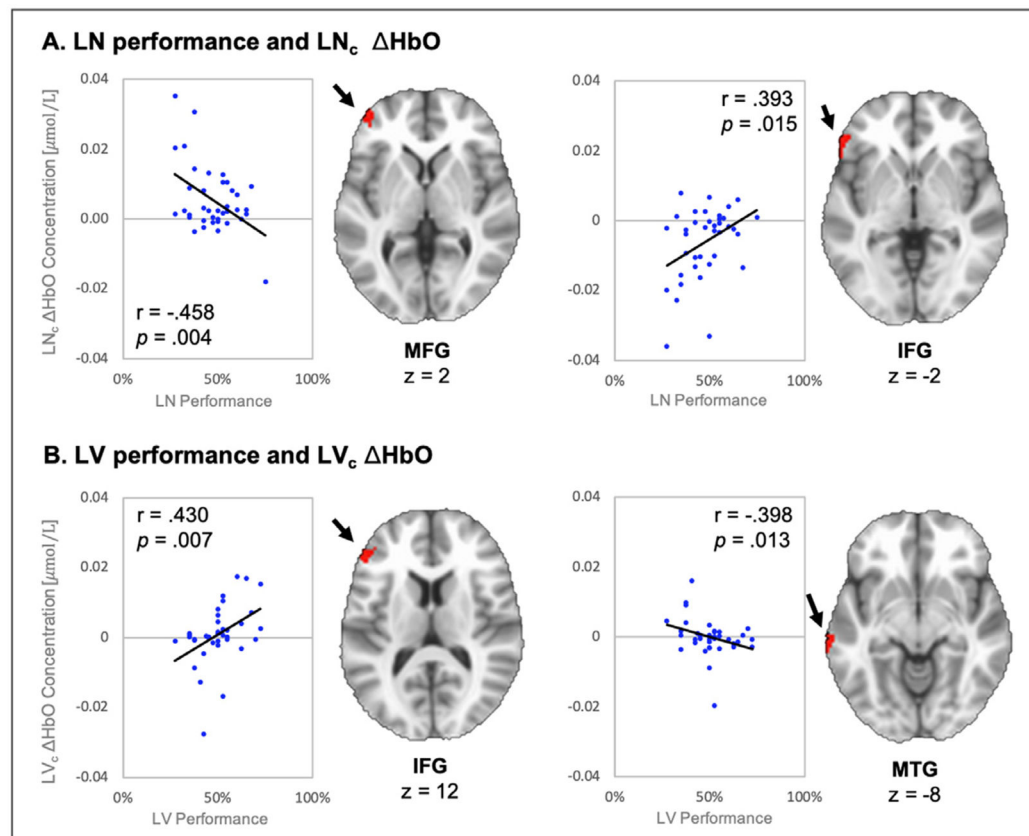
Fig. 6.
Results of brain/behavior correlational analyses between low-intelligibility conditions and their respective performance scores. **A.** Negative correlation in the MFG (left) and positive correlation in the IFG (right) are plotted for LN condition data. **B.** Positive correlation in the MFG (left) and negative correlation in the MTG (right) are plotted for LV condition data. Pearson Correlation (r) and significance shown inside each scatterplot. Linear trendline is in black. Clusters are denoted by black arrows.

Table 1

Breakdown of factors and levels for each repeated-measures ANOVA. The ANOVAs detailed in *A* and *B* test the effects of each type of distortion (speech-shaped background noise and vocoding, respectively) on the baseline response to speech in quiet. ANOVAs detailed in *C* and *D* examined how distortion type interacted with intelligibility (*C*) and trial accuracy (*D*), respectively. Subscripts *c* and *i* indicate correct or incorrect trials. HbO, oxy-hemoglobin; HbR, deoxy-hemoglobin.

A. Hemoglobin (2) X Noise Status (3)

Hemoglobin		Noise Status	Condition
1 HbO	1	No Noise	SQ
	2	+10 dB SNR	HN
	3	−4 dB SNR	LN_c
2 HbR	1	No Noise	SQ
	2	+10 dBSNR	HN
	3	−4 dB SNR	LN_c

B. Hemoglobin (2) X Vocoding Status (3)

Hemoglobin		Vocoded Status	Condition
1 HbO	1	No Vocoding	SQ
	2	8-channel Vocoding	HV
	3	Noise at +7 dB SNR, then 8-ch. vocoding	LV_c
2 HbR	1	No Vocoding	SQ
	2	8-channel Vocoding	HV
	3	Noise at +7 dB SNR, then 8-ch. vocoding	LV_c

C. Hemoglobin (2) X Distortion (2) X Intelligibility (2)

Hemoglobin		Distortion		Intelligibility	Condition
1 HbO	1	Background Noise	1	High	HN
			2	Low	LN_c
	2	Vocoded	1	High	HV
			2	Low	LV_c
2 HbR	1	Background Noise	1	High	HN
			2	Low	LN_c
	2	Vocoded	1	High	HV
			2	Low	LV_c

D. Hemoglobin (2) X Distortion (2) X Accuracy (2)

Hemoglobin		Distortion		Accuracy	Condition
1 HbO	1	Background Noise	1	Correct	LN_c
			2	Incorrect	LN_i
	2	Vocoded	1	Correct	LV_c
			2	Incorrect	LV_i
2 HbR	1	Background Noise	1	Correct	LN_c
			2	Incorrect	LN_i
	2	Vocoded	1	Correct	LV_c
			2	Incorrect	LV_i

Table 2

Results of behavioral performance per condition and paired samples t-tests between each condition (mean/ standard deviation as percentage).

A. Performance		B. Pairwise Comparisons		
Condition	Mean % (Std. Dev. +/−)	Comparison	t	Sig. (2-tailed)
HN	99.7 (.01)	HN – HV	7.97	<.001
HV	92.5 (.06)	HN – SQ	−2.09	.044
LN	47.7 (.12)	HV – SQ	−8.35	<.001
LV	50.3 (.11)	LN – LV	−1.599	.118

Table 3

Significant effects/interactions and their respective cluster regions[b] and coordinates are listed by ANOVA for the image-based analyses. MNI coordinates (x, y, z) report center of mass for each cluster effect. IFG, inferior frontal gyrus; MFG, middle frontal gyrus; MTG, middle temporal gyrus; Hb, Hemoglobin.

ANOVA Factors (no. levels)	Effect/ Interaction	Cluster(s)[a]	Cluster location[b]	X	Y	Z	Spatial Extent (2 mm²)	F	Effect Size (η2)
A) Hb (2) X Noise Level (3)	Hb	1**	IFG	43	−45	15	86	11.62	.068
		2*	MTG	64	9	−13	81	6.81	.097
		3***	MTG	67	32	−5	65	10.34	.114
	Hb X Noise Level	1**	MFG	42	−50	8	135	5.87	.077
		2**	IFG[c]	52	−32	−2	109	6.41	.125
B) Hb (2) X Vocoding Level (3)	Hb	1**	MFG	43	−49	9	273	10.61	.164
		2**	MTG	64	11	−10	101	7.77	.122
	Hb X Vocoding Level	1*	MTG	66	29	−3	148	6.59	.086
C) Hb (2) X Distortion (2) X Intelligibility (2)	Hb	1***	MFG	43	−46	15	114	12.32	.967
		2**	MTG	64	10	−12	73	7.06	.087
	Hb x Distortion	1*	IFG	52	−33	0	93	6.08	.064
		2**	MTG	67	32	−6	52	9.06	.077
	Hb x Intelligibility	1***	MTG	66	27	−4	150	18.59	.095
		2*	MFG	43	−49	7	97	5.99	.123
D) Hb (2) X Distortion (2) X Accuracy (2)	Hb	1**	IFG	46	−46	12	96	10.35	.055
		2**	IFG	50	−36	−6	92	9.74	.105
		3**	MTG	66	30	−5	86	8.26	.147
	Hb x Distortion	1*	IFG	53	−32	4	88	5.89	.050
	Hb x Accuracy	1**	MFG	42	−50	4	184	9.72	.104
		2**	MTG	64	10	13	86	7.32	.096

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

[a] Significance is marked as follows:

*
p 0.05

**
p 0.01

***
p 0.001.

[b] All regions are from the left hemisphere.

[c] Greenhouse-Geisser correction applied when necessary.

Table 4

Results of follow-up paired samples t-tests for the interactions between Hb X Noise Level (ANOVA A) and Hb X Vocoding Level (ANOVA B).

Interaction	Comparison[a]	t	Sig. (2-tailed)[b]
Hb X Noise Level (cluster 1)	SQ – HN	.031	1.00
	SQ – LN$_c$	−2.97	.015
	HN – LN$_c$	−2.97	.015
Hb X Noise Level (cluster 2)	SQ – HN	1.88	.201
	SQ – LN$_c$	2.95	.018
	HN – LN$_c$	2.15	.114
Hb X Vocoding Level (cluster 1)	SQ – HV	2.99	.015
	SQ – LV$_c$	−391	1.00
	HV – LV$_c$	−3.23	.009

[a] The values that are being compared are the mean differences between HbO and HbR for each condition.

[b] Bonferoni correction applied for multiple comparisons.