PROCEEDINGS of the
23rd International Congress on Acoustics
9 to 13 September 2019 in Aachen, Germany

Using functional near infrared spectroscopy to assess auditory responses in auditory and lateral frontal cortex

Min Zhang[(1)], Antje Ihlefeld[(2)]

[(1)]New Jersey Institute of Technology and Rutgers University, USA, mz86@njit.edu
[(2)]New Jersey Institute of Technology, USA, ihlefeld@njit.edu

Abstract

A growing literature reveals that brain mechanisms contribute to symptoms of hearing loss. To study how hearing loss affects brain function, several groups including our own have developed light-based methods via functional near-infrared spectroscopy (fNIRS). Using fNIRS, we recently confirmed that lateral frontal cortex (LFC) engages when normal-hearing listeners direct attention to words in background sound versus listen passively. We now test how sensory resolution affects LFC recruitment. As a control, we also record over auditory cortex (AC). Using simulated cochlear implant speech, we ask listeners to perform a word detection task in a situation with competing background speech, as a function of the amount of sensory detail in the cochlear implant simulation. One possibility is that impoverished sensory cues reduce AC or LFC engagement, as compared to high-fidelity cues, limiting the potential usefulness auditory attention. Alternatively, poorly resolved sensory detail may increase AC or LFC engagement, causing it to saturate and limiting overall speech intelligibility. Recruitment of AC and LFC will be discussed in context of the saturation hypothesis.

Keywords: functional near infrared spectroscopy, cochlear implant

1 INTRODUCTION

Studies increasingly show how hearing loss shifts the balance of excitation and inhibition along the auditory pathway of the brain (e.g. Chambers et al. 2016). However, when cochlear implants are fitted to an individual user, device programming is mainly based on an individual's audiometric thresholds, ignoring auditory brain health. This may be a reason why these devices generally provide good audibility in quiet environments but provide poor performance in situations with background sound for the majority of users. The goal of the current work is to develop a tool by which we can assess auditory brain function in people with cochlear implants, especially in situations with background sound.

When background sound is present, auditory attention should act as a gatekeeper, such that only a small fraction of the incoming auditory information enters higher-order processing (Ihlefeld & Shinn-Cunningham, 2008). Indeed, cochlear implant users can direct auditory attention (Paredes-Gallardo et al. 2018). In normal-hearing listeners auditory attention engages auditory cortex (AC) and auditory-specific regions in lateral frontal cortex (LFC; Michalka et al. 2015). Using fNIRS, we widened this approach and confirmed that LFC engages more strongly when a listener actively attends to sound in the presence of a noisy background, versus listens passively (Zhang et al. 2018). To understand how AC and LFC of cochlear implant users may respond in these kinds of acoustic situations, we now simulate coarse cochlear-implant-like speech and present these simulations to normal-hearing listeners.

2 METHODS

A total of 11 normal-hearing listeners (aged 21-40, 3 female) participated in the study. All listeners gave written informed consent to participate in the study. All testing was administered according to the guidelines of the institutional review board of the New Jersey Institute of Technology.

Overall methods were similar to previous work (for detailed description see: Zhang et al. 2018), except that 1)

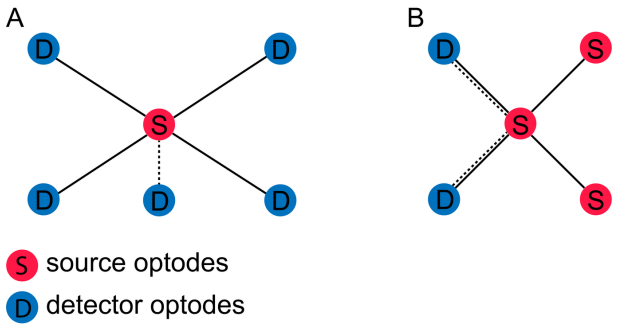

Figure 1. Optode configurations A) AC B) LFC. Solid lines denote deep, dashed lines shallow channels.

we presented acoustically simulated cochlear-implant speech and 2) we refined the placement of optode source-detector pairs to improve recording robustness.

2.1 BEHAVIORAL DATA ACQUISITION

Briefly, using a simulation technique called vocoding, we processed single words from a closed-set corpus that were uttered in isolation by 1) bandpass-filtering the speech into consecutive bands, 2) multiplying the Hilbert envelope of each band with a sine-wave carrier and 3) adding the component bands to form the simulated speech signal. Stimuli could either be unprocessed speech ("normal"), or vocoded with 32, 16 or 8 bands. Each utterance was time scaled to be of 300 ms duration. Testing a 2-talker mixture of 15 seconds of concatenated vocoded speech, we then asked young normal-hearing listeners to report when target key words are uttered by clicking a button on a handheld response interface. Each 15-second task epoch was followed by a 30 to 40 second silent rest period (duration randomly jittered). A listener's response was scored a "hit" if a button press occurred within 1200 ms after the onset of a target utterance, and a "miss" if no button press occurred within 1200 ms after target onset. If the listener pushed a button and no target occurred within 1200 ms prior to button push, this was scored a "false alarm." All other non-responses were scored "correct reject." A listener's behavioral sensitivity was then calculated by simulating chance performance via random guessing and subtracting the z-score of this simulated chance performance from that of the observed percent correct, resulting in d' scores.

2.2 PHYSIOLOGICAL DATA ACQUISITION

Using a 3D-location tracking system (Brainsight 2.0 software and hardware by Rogue Research Inc., Canada), we recorded four landmark coordinates on each listener's head: nasion, inion, and bilateral preauricular points. Relative to these four landmarks and using the Montreal Neurological Institute ICBM-152 brain atlas (Talairach 1988), we referenced four regions of interest (ROI), specifically, the right AC, left AC, right LFC, and left LFC. Infrared optodes, embedded in a custom-made headcap were placed directly above each of the four ROIs, and held in place with adjustable straps.

Using two different wavelengths, 690 and 830 nm, 50 Hz sampling frequency and a total of 8 source and 14 detector optodes, a continuous-wave diffuse-optical NIRS system simultaneously recorded light absorption (CW6; TechEn Inc., Milford, MA). We used two recording depths. To estimate the neurovascular response of cortical tissue between 0.5 to 1 cm below the surface of the skull, deep channels had a 3 cm source-detector distance (solid lines in Fig. 1 A and B). In contrast, shallow channels, used to partial out physiological noise, had a source-detector distance of 1.5 cm (dotted lines in Fig. 1 A and B). A subset of all potential combinations of optode-detector pairs was interpreted as response channels and further analyzed. To improve test-retest reliability, based on pilot testing, the optode design matrices differed across the AC and LFC recording sites. Specifically, at each of the two LFC recording sites we combined one optical source and five surrounding de-

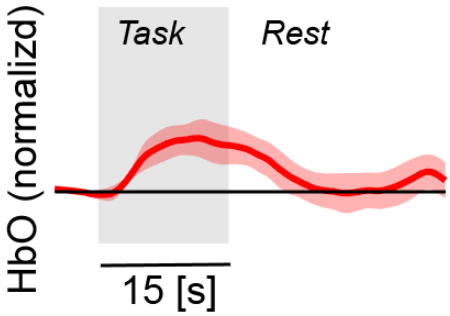

Figure 2. Representative across-listener average HbO trace during auditory task engagement with normal speech in the right AC versus rest. Ribbon shows one standard error of the mean across listeners.

tectors into one probe set, resulting in four deep and one shallow channel (Fig. 1A). At each of the AC sites, we used a more compact channel design, combining source and detector optodes at the four corners to form two deep channels diagonally, and combining the center optode with two corner detector optodes to form to shallow channels (dashed lines in Fig. 1B).

Listeners initially performed a controlled breathing task, followed by auditory testing (for details see Zhang et al. 2018). Using HOMER2 (Huppert et al. 2009), a series of signal processing steps and the modified Beer–Lambert law (Cope & Delpy, 1988; Kocsis et al. 2006), we extracted oxy-hemoglobin (HbO) traces. We then normalized all HbO traces from the auditory task conditions by dividing each trace by the maximal HbO concentration change in that source-detector pair during controlled breathing. We then fitted the normalized HbO traces by four general linear models (GLM), one GLM for each ROI of the form: $y(t) = x_{normal}(t) * \beta_{normal} + x_{32bands}(t) * \beta_{32} + x_{16bands}(t) * \beta_{16} + x_{8bands}(t) * \beta_{8} + x_{nuisance}(t) * \beta_{nuisance} + \varepsilon(t)$ where y is the normalized HbO trace, t is time, and the β values indicate the activation levels of each of the task regressors. Specifically, β_i was change attributed to vocoding condition i, $\beta_{nuisance}$ was the HbO change attributed to non-neurally driven events, and $\varepsilon(t)$ was the residual error of the GLM. Note that the task regressors x_i in the GLM design matrix contained reference functions for the appropriate vocoding condition, each convolved with a canonical hemodynamic response function (for further details on data processing, see Zhang et al. 2018).

3 RESULTS

Figure 2 shows the across-listener average HbO trace for a representative task condition and ROI. As expected, HbO in AC increases during auditory task engagement and decreases during the silent rest period. Similar HbO traces were obtained for all combinations of four ROIs and four vocoding conditions. For each ROI, Figure 3A-D shows β_i HbO activation levels due to task engagement, as derived from the GLM models versus the behavioral d′ scores. Task-driven responses in all four ROIs were significantly different from zero. To estimate the within-listener association between behavior and fNIRS response, for each ROI, we calculated the repeated measures correlations between the β_i at that ROI and d′ (Bakdash et al. 2018). These repeated measures correlations were robust in right AC but did not reach statistical significance in the other three ROIs. Specifically, for right AC, r(32) = 0.45, p <0.01. For left AC, r(32) =- 0.05, p = 0.79. For right LFC, r(32) = -0.14, p = 0.41. For left LFC, r(32) =0.09, p = 0.59. However, a repeated measures analysis of variance on the HbO activation levels β in right AC did not find a main effect of sensory detail [F(3,30) = 0.3, p > 0.01].

4 DISCUSSION

The current study intended to assess viability of fNIRS for measuring auditory brain function. Here, listeners detected target keywords among two competing streams of rapidly changing words. Results show that right AC

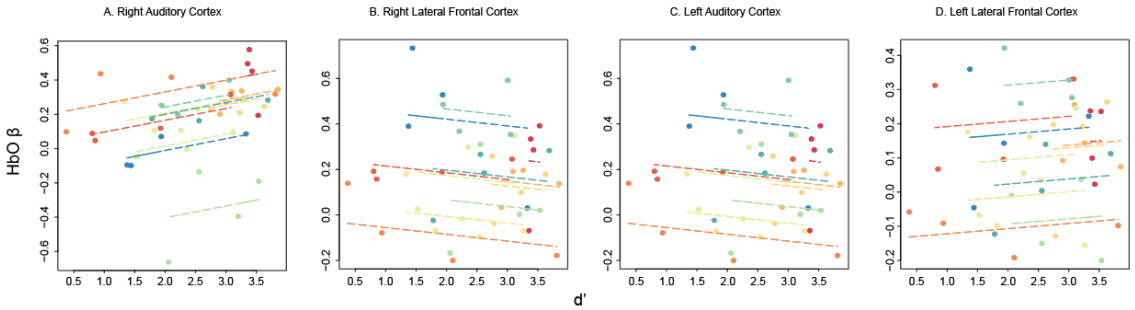

Figure 3. HbO activation strength (β_i), as attributed to hemodynamic response function, versus behavioral sensitivity (d′) at each of the four ROIs. Each of the seven colors denotes one listener.

activation is significantly correlated with behavioral sensitivity to the task. This is consistent with data from functional magnectic resonance imaging which show that the blood-oxygen-level-dependent response is a viable indicator of perceived sensory detail, at least in the AC (Sohoglu et al. 2012). Our current results suggest that right AC is a promising site for assessing auditory brain function via fNIRS. However, using fNIRS, the current data show that the differential recruitment due to changes in sensory detail is smaller than what we can currently statistically resolve at the population level with 11 listeners. One possibility is that right AC reflects behavioral accuracy. Alternatively, the effect size in HbO activation levels may be too small considering the overall variance in the fNIRS traces. Indeed, to date, no standardized method exists for estimating brain activity from fNIRS recordings (e.g., Knauth et al. 2017). Future work will continue to test signal processing algorithms to statistically partial out nuisance signals such as cardiac rhythm, respiratory induced change, and blood pressure variations from the desired hemodynamic response driven by neural events in cortex.

5 CONCLUSIONS

Using a target detection paradigm with rapid concurrent two-talker speech, we find that when fNIRS recordings are fitted with hemodynamic response functions, responses in right AC correlate robustly with behaviorally observed sensitivity.

ACKNOWLEDGEMENTS

This work was supported by the New Jersey Health Foundation (PC 24-18 to A. I.) and the National Science Foundation (MRI CBET 1428425).

REFERENCES

[1] Bakdash, J. Z., & Marusich, L. R. (2017). Repeated Measures Correlation. Frontiers in Psychology, 8. doi:10.3389/fpsyg.2017.00456

[2] Cope, M., & Delpy, D. T. (1988). System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. Medical & Biological Engineering & Computing, 26(3), 289-294. doi:10.1007/bf02447083

[3] Chambers, A.R., Resnik, J., Yuan, Y., Whitton, J.P., Edge, A.S., Liberman, M.C. & Polley, D.B., 2016. Central gain restores auditory processing following near-complete cochlear denervation. Neuron, 89(4), pp.867-879.

[4] Huppert, T. J., Diamond, S. G., Franceschini, M. A., & Boas, D. A. (2009). HomER: A review of time-series analysis methods for near-infrared spectroscopy of the brain. Applied Optics, 48(10). doi:10.1364/ao.48.00d280

[5] Ihlefeld, A., & Shinn-Cunningham, B. (2008). Spatial release from energetic and informational masking in a selective speech identification task. The Journal of the Acoustical Society of America, 123(6), 4369-4379. doi:10.1121/1.2904826

[6] Knauth, M., Heldmann, M., Münte, T. F., Royl, G. (2017). Valsalva-induced elevation of intracranial pressure selectively decouples deoxygenated hemoglobin concentration from neuronal activation and functional brain imaging capability. NeuroImage 162: 151–161. doi:10.1016/j.neuroimage.2017.08.062.

[7] Kocsis, L., Herman, P., & Eke, A. (2006). The modified Beer–Lambert law revisited. Physics in Medicine and Biology, 51(5). doi:10.1088/0031-9155/51/5/n02

[8] Michalka, S., Kong, L., Rosen, M., Shinn-Cunningham, B., & Somers, D. (2015). Short-Term Memory for Space and Time Flexibly Recruit Complementary Sensory-Biased Frontal Lobe Attention Networks. Neuron, 87(4), 882-892. doi:10.1016/j.neuron.2015.07.028

[9] Paredes-Gallardo, A., Innes-Brown, H., Madsen, S. M., Dau, T., & Marozeau, J. (2018). Auditory Stream Segregation and Selective Attention for Cochlear Implant Listeners: Evidence From Behavioral Measures and Event-Related Potentials. Frontiers in Neuroscience, 12. doi:10.3389/fnins.2018.00581

[10] Sohoglu, E., Peelle, J.E., Carlyon, R.P. and Davis, M.H., (2012). Predictive top-down integration of prior knowledge during speech perception. Journal of Neuroscience, 32(25), pp.8443-8453.

[11] Talairach, J., Rayport, M., Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: An approach to cerebral imaging, Stuttgart, Germany: Thieme.

[12] Zhang, M., Ying, Y.-L.M., & Ihlefeld, A. (2018). Spatial Release From Informational Masking: Evidence From Functional Near Infrared Spectroscopy. Trends in Hearing. doi:10.1177/2331216518817464