Defining the Human Neural Network for Vocal Production and Control

Farhad Tabasi[1], Roozbeh Behroozmand[2], Hyunjin Cho[3], Christopher K. Kovach[4], Christopher I. Petkov[1], Matthew A. Howard III[1], and Jeremy D. Greenlee[1]*

1 Department of Neurosurgery, The University of Iowa, Iowa City, IA, USA
2 Department of Speech, Language, and Hearing, The University of Texas at Dallas, Richardson, TX, USA
3 Department of Computer Science, The University of Iowa, Iowa City, IA, USA
4 Department of Neurosurgery, University of Nebraska Medical Center, Omaha, NE, USA
Corresponding author: jeremy-greenlee@uiowa.edu

IOWA HEALTH CARE Human Brain Research Lab

Introduction

The ability to control vocalization is an integral part of human speech and verbal communication, relying heavily on a constellation of cortical regions. Changing vocal pitch is a highly complex process in spoken communication and requires a coordinated sensory-motor network to monitor the output in real-time and post-production [1].
Different brain regions are suggested as the main coordinators of vocalization and pitch change, including two distinct cortical regions that control laryngeal movements: the ventral laryngeal motor cortex (vLMC) and the dorsal premotor region (dLMC) [2].
These primarily motor regions, engaged in planning and executing motor commands, are also active during auditory information processing [3, 4].
Although known as motor regions responsible for motor planning and execution, neural activity patterns in non-speech vocalization and monitoring auditory feedback remain understudied.

Methods

Design and Recording
We investigated alterations in intracranially recorded local field potentials (LFPs) in three right-handed neurosurgery patients (**Table 1**) requiring stereo-electroencephalography (SEEG) and electrocorticography (ECoG) recordings for clinical monitoring.
The recordings were conducted during a task in which the participants learned to control their voice to create sustained vowel vocalizations with cued dynamic pitch changes compared to passively listening to the playback of the same self-produced sounds (**Figure 1**). All poorly performed trials (i.e., those where the subjects couldn't increase their pitch to half of the goal pitch or vocalization onset was off by more than 0.5 seconds from the cue) were excluded.

Electrophysiology analysis
LFPs recorded from surface electrodes sampled at 2 kHz online were included in the analysis. First, signals were downsampled to 1 kHz and denoised using a demodulated band transform (DBT) with a bandwidth of 0.25 [5]. Singular value decomposition was implemented to discard the first principal component based on the covariance matrix computed from the high-pass (300 Hz) filtered data [6]. Broadband (1-150 Hz) event-related band power (ERBP) was calculated using the DBT method from LFPs recorded in the dorsolateral prefrontal cortex, inferior frontal gyrus, superior temporal gyrus, and precentral gyrus using surface grids during both vocalization and passive listening trials.
For both vocalization and passive listening responses, ERBP was calculated by normalizing post-vocalization spectral power relative to the baseline power from 1 second to 100 milliseconds before vocalization onset (i.e., during silence). Spatiotemporal dynamics of cortical activity at each frequency range (defined as 1-4 Hz 'Delta,' 4-8 Hz 'Theta,' 8-14 Hz 'Alpha,' 14-30 Hz 'Beta,' 30-70 Hz 'Low-Gamma,' 70-150 Hz 'High-Gamma') were overlaid on the native mesh surface of individual participants. Electrode response density was calculated for each electrode displaying a given response profile in 1-cm-diameter regions of the canonical cortical surface (**Figure 4**).

Table 1. Participants demographics

Subject	Age (y)	Sex	Handedness	Language Dominant hemisphere (Wada)	Recorded side	Note
L315	31	M	Right-handed (+80)	Unknown	Right	Right hemisphere lesion limited to posterior premotor cortex and superior frontal gyrus
R322	28	F	Right-handed (+100)	Unknown	Right	-
L362	58	M	Right-handed (+100)	Both	Right and Left	-

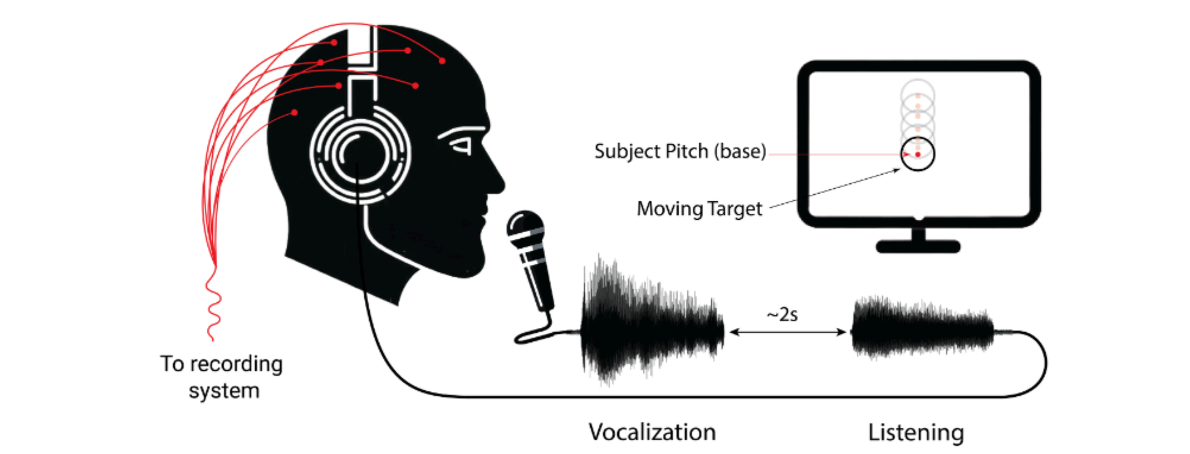

***Figure 1.* Schematic of task design.** The target's starting point was calibrated to each subject's baseline pitch, and a subsequent increase of 300 cents per second above this baseline was introduced after approximately 1 second. Participants' real-time vocal pitch (F0) was extracted and displayed on the screen. They were instructed to modulate their vowel vocalizations by following a visual target to match their vocal pitch to the visual target on the screen. Each vocalization trial was recorded and immediately played back to the participant through insert earmolds for the listening trials.

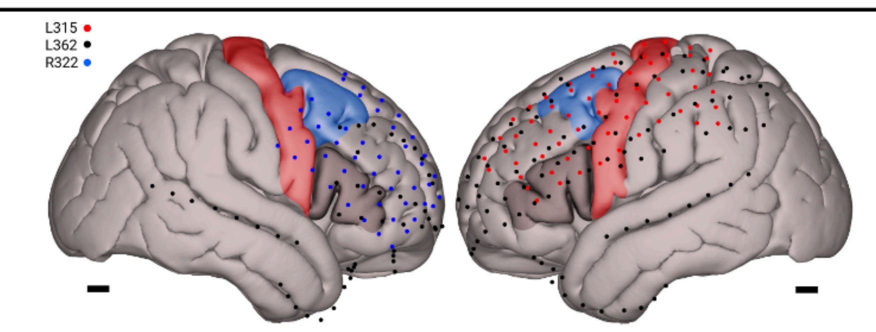

***Figure 2.* Surface electrode coverage.** ECoG coverage of all three participants was non-linearly co-registered and transformed into Montreal Neurological Institute (MNI) space. The precentral gyrus (red), caudal middle frontal gyrus (blue), and inferior frontal gyrus (black) are highlighted in color.

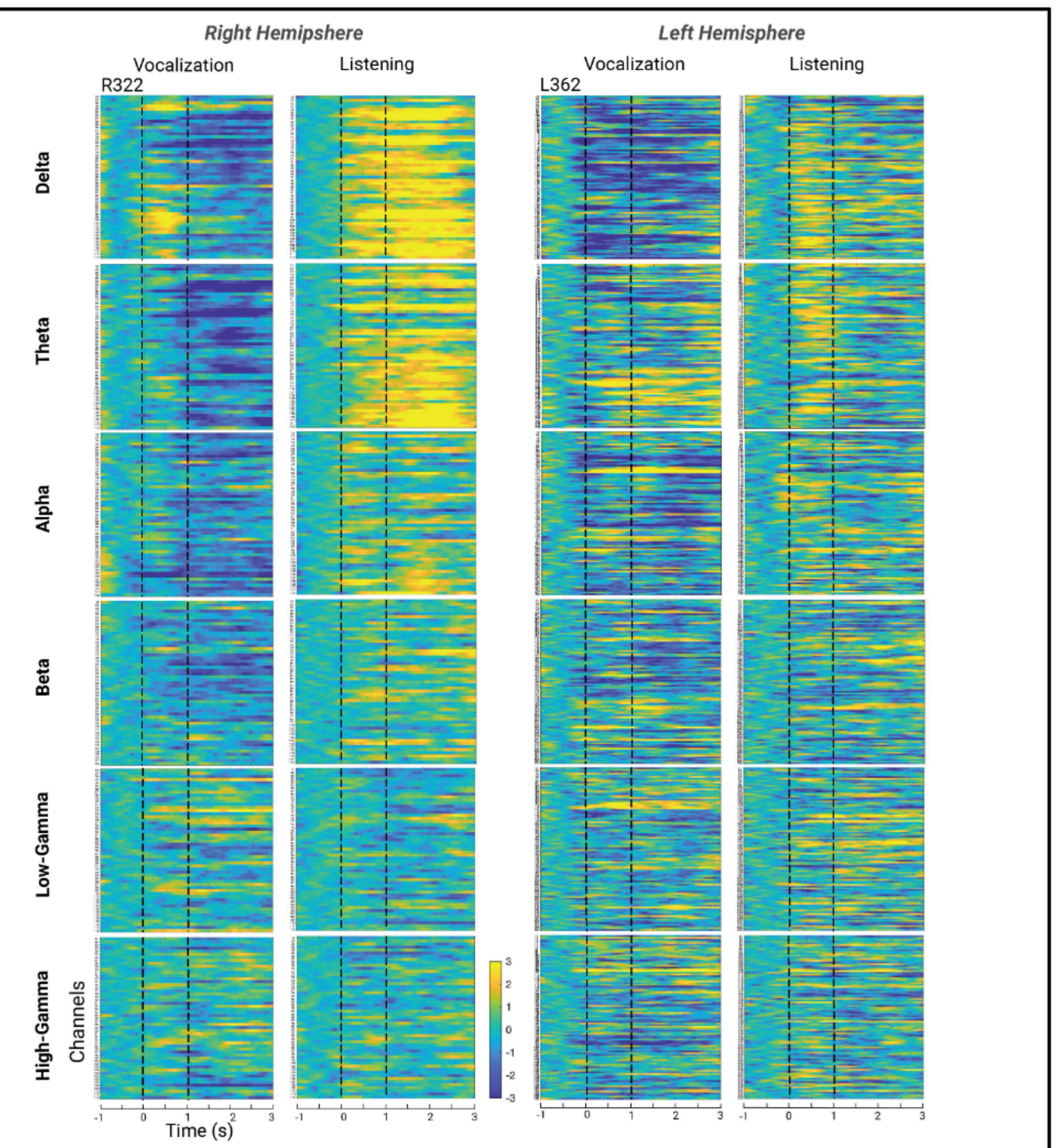

***Figure 3.* Network activity in vocalization and listening**. Exemplar of left vs. right hemisphere (from subjects R362 and R322, respectively) during vocalization and passive listening to playback for each frequency range. Vertical lines are aligned to the main time points: t = 0 (the beginning of the vocalization) and t = 1 (the time the subject starts to increase voice pitch).

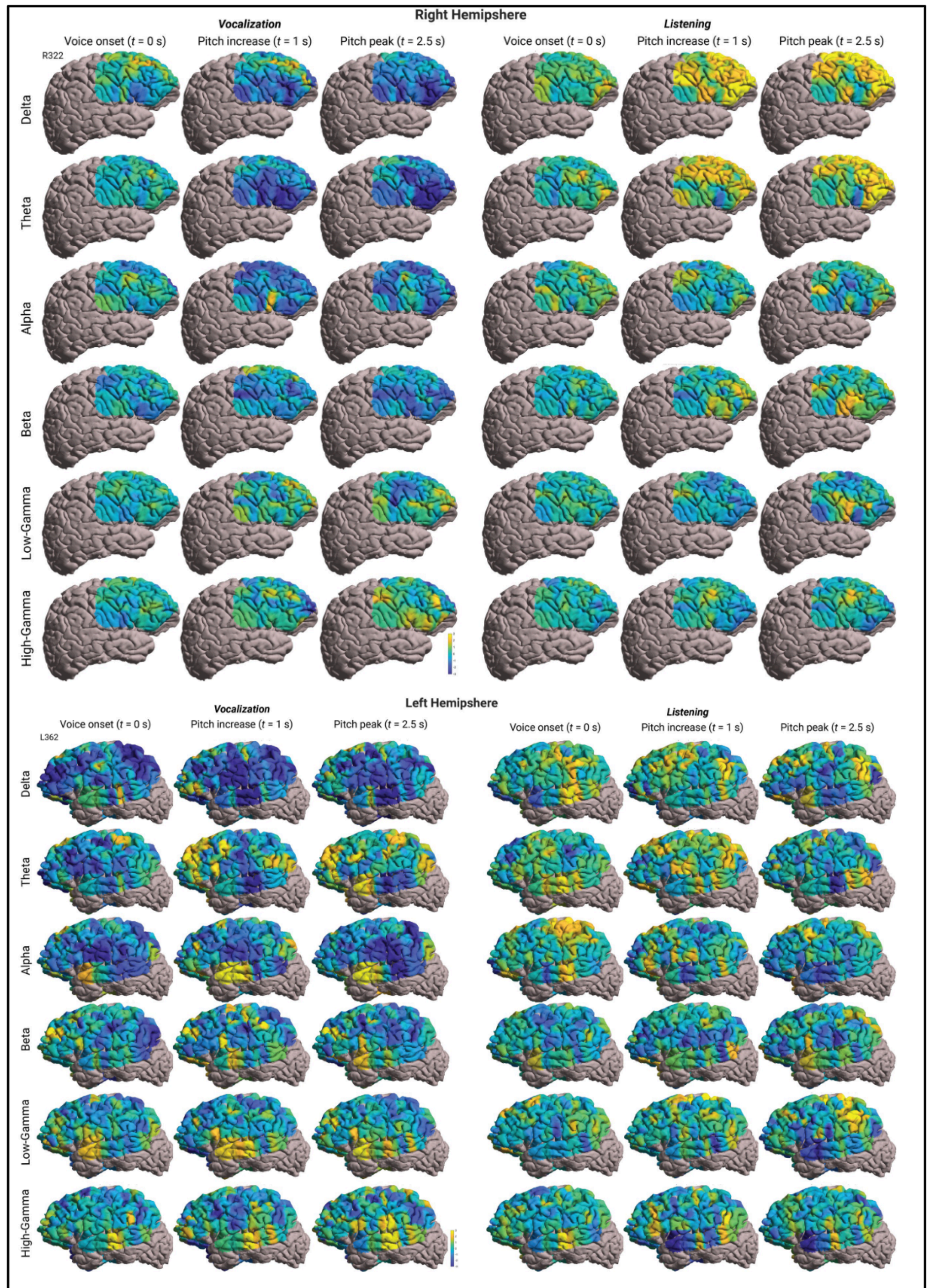

***Figure 4.* Spatiotemporal pattern of cortical activation during vocalization and listening to playback for exemplar left and right hemispheres at each frequency range.** We defined three main time points to examine the cortical activity pattern: 1) Voice onset (t = 0): When the subject sees the cue on the screen and starts vocalization. 2) Pitch increase (t = 1s): When the target starts moving upward, the subject should increase their pitch. 3) Pitch peak (t = 2.5): When the subject reaches the maximum level of increased pitch.
The right hemisphere is less active at lower frequencies during vocalization, as opposed to higher frequencies during the same task, and also at lower frequencies during listening, especially in the inferior frontal gyrus (IFG). The left hemisphere shows high activity at lower frequencies (theta) and in the gamma range over the IFG and posterior premotor area. However, this activity is more evident in these regions at lower frequencies during listening.

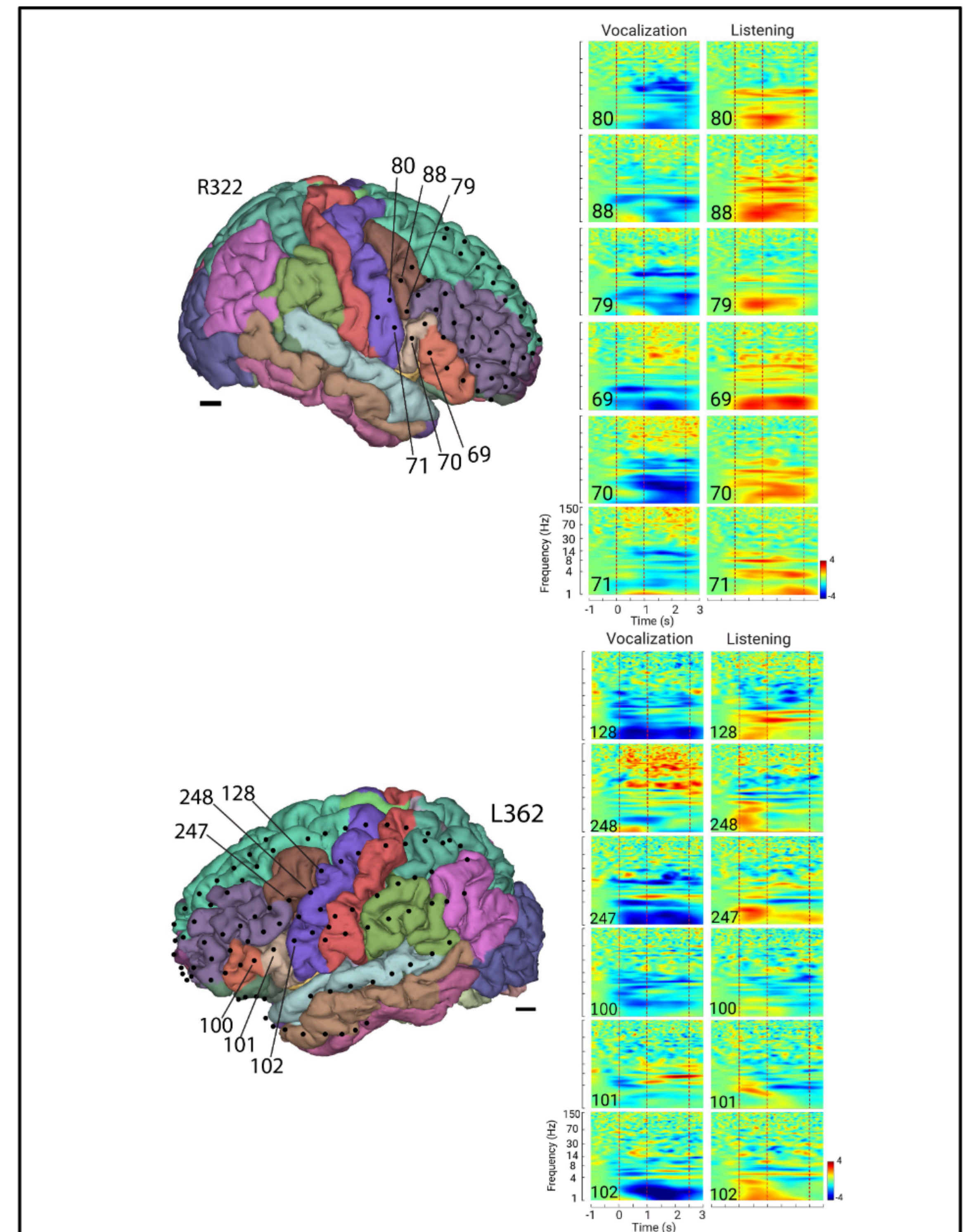

***Figure 5.* The ERBP of right and left hemisphere exemplars** show the pattern of cortical activity in the posterior premotor area and caudal inferior frontal gyrus during vocalization and listening. In the left hemisphere, increased activity at higher frequencies during vocalization and at lower frequencies during listening is evident. In the right hemisphere, the activity pattern (increased or suppressed) is noticeable at lower frequencies either during vocalization or listening.

Summary

These results show that regions primarily known as motor areas are also engaged in the perception of sustain vowel vocalization during listening to the playback. Although this study is limited to three subjects, hemisphere lateralization may be considered in pitch production. However, both hemispheres seem active and engaged in the perception of auditory stimuli, not only in the auditory cortex but also in motor regions responsible for vocal motor planning and execution.

Acknowledgment

We thank Haiming Chen for his help in data collection and Juan Vivanco-Suarez for his comments.
This work is supported by grant R01DC04290.

Contact information: Farhad Tabasi, MD. farhad-tabasi@uiowa.edu

References

1. Behroozmand, R., et al., Sensory–motor networks involved in speech production and motor control: An fMRI study. NeuroImage, 2015. 109: p. 418-428.
2. Hickok, G., J. Venezia, and A. Teghipco, Beyond Broca: neural architecture and evolution of a dual motor speech coordination system. Brain, 2023. 146(5): p. 1775-1790.
3. Du, Y., et al., Noise differentially impacts phoneme representations in the auditory and speech motor systems. Proc Natl Acad Sci U S A, 2014. 111(19): p. 7126-31.
4. Greenlee, J., et al., Sensorimotor integration during human self-vocalization: Insights from invasive electrophysiology. Proceedings of Meetings on Acoustics, 2013. 19(1).
5. Kovach, C.K. and P.E. Gander, The demodulated band transform. Journal of Neuroscience Methods, 2016. 261: p. 135-154.
6. Kocsis, Z., et al., Immediate neural impact and incomplete compensation after semantic hub disconnection. Nature Communications, 2023. 14(1): p. 6264.