

4-15-2022

Neural Correlates of Impaired Vocal Feedback Control in Post-stroke Aphasia

Roozbeh Behroozmand

Leonardo Bonilha

Chris Rorden

University of South Carolina, rorden@mailbox.sc.edu

Gregory Hickok

Julius Fridriksson

Follow this and additional works at: https://scholarcommons.sc.edu/psyc_facpub



Part of the [Psychology Commons](#)

Publication Info

Published in *NeuroImage*, Volume 250, 2022, pages 118938-.

This Article is brought to you by the Psychology, Department of at Scholar Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Scholar Commons. For more information, please contact digres@mailbox.sc.edu.



Neural correlates of impaired vocal feedback control in post-stroke aphasia

Roozbeh Behroozmand^{a,*}, Leonardo Bonilha^b, Chris Rorden^c, Gregory Hickok^d, Julius Fridriksson^e

^a Speech Neuroscience Lab, Department of Communication Sciences and Disorders, Arnold School of Public Health, University of South Carolina, 915 Greene Street, Columbia, SC 29208, United States

^b Department of Neurology, Medical University of South Carolina, Charleston, SC 29425, United States

^c Department of Psychology, University of South Carolina, Columbia, SC 29208, United States

^d Department of Cognitive Sciences, University of California, Irvine, Irvine CA 92697, United States

^e The Aphasia Lab, Department of Communication Sciences and Disorders, Arnold School of Public Health, University of South Carolina, 915 Greene St., Columbia, SC 29208, United States

ARTICLE INFO

Keywords:

Vocal motor control
Sensorimotor integration
Effort copy
Auditory feedback
Aphasia

ABSTRACT

We used left-hemisphere stroke as a model to examine how damage to sensorimotor brain networks impairs vocal auditory feedback processing and control. Individuals with post-stroke aphasia and matched neurotypical control subjects vocalized speech vowel sounds and listened to the playback of their self-produced vocalizations under normal (NAF) and pitch-shifted altered auditory feedback (AAF) while their brain activity was recorded using electroencephalography (EEG) signals. Event-related potentials (ERPs) were utilized as a neural index to probe the effect of vocal production on auditory feedback processing with high temporal resolution, while lesion data in the stroke group was used to determine how brain abnormality accounted for the impairment of such mechanisms. Results revealed that ERP activity was aberrantly modulated during vocalization vs. listening in aphasia, and this effect was accompanied by the reduced magnitude of compensatory vocal responses to pitch-shift alterations in the auditory feedback compared with control subjects. Lesion-mapping revealed that the aberrant pattern of ERP modulation in response to NAF was accounted for by damage to sensorimotor networks within the left-hemisphere inferior frontal, precentral, inferior parietal, and superior temporal cortices. For responses to AAF, neural deficits were predicted by damage to a distinguishable network within the inferior frontal and parietal cortices. These findings define the left-hemisphere sensorimotor networks implicated in auditory feedback processing, error detection, and vocal motor control. Our results provide translational synergy to inform the theoretical models of sensorimotor integration while having clinical applications for diagnosis and treatment of communication disabilities in individuals with stroke and other neurological conditions.

1. Introduction

Voluntary control of the larynx is a key innovation to the evolution of human speech and involves complex neuro-computational mechanisms that incorporate sensory feedback for vocal production (Fischer, 2017; Fitch, 2017, 2010; Hickok, 2017; Kuypers, 1958a, 1958b). These mechanisms mediate segmental (voicing) and supra-segmental (e.g., prosody, rhythm, stress) processes that are critical for speech communication and rely on brain networks that support sensorimotor integration for online monitoring of auditory feedback information for regulating vocal motor output (Pichon and Kell, 2013; Tang et al., 2017).

Research on neurologically intact populations has provided evidence for the role of auditory feedback in motor control of vocalization (Behroozmand et al., 2016; Burnett et al., 1998; Chang et al.,

2013; Chen et al., 2007; Larson, 1998) and speech (Cai et al., 2011; Niziolek and Guenther, 2013; Tourville et al., 2008). These studies have demonstrated that speakers detect and correct for alterations (i.e. errors) in their online auditory feedback by generating compensatory motor responses that change their vocal output in the opposite direction to external stimuli. According to the dual-stream model (Hickok, 2012; Hickok et al., 2011; Hickok and Poeppel, 2004; Rauschecker, 2011), this function is mediated by predominantly left-lateralized sensorimotor networks that use an internal forward model to translate efference copies of motor commands to predict auditory consequences of intended vocal outputs. This internally established forward prediction provides the system with the advantage to execute rapid vocal corrections in case of erroneous productions even before the actual feedback has become available. In addition, when the online feedback is altered during pro-

* Corresponding author at: 915 Greene Street.

E-mail address: r-behroozmand@sc.edu (R. Behroozmand).

<https://doi.org/10.1016/j.neuroimage.2022.118938>.

Received 8 September 2021; Received in revised form 31 December 2021; Accepted 25 January 2022

Available online 29 January 2022.

1053-8119/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

duction, the comparison between internally predicted and actual auditory feedback gives rise to an error signal that triggers corrective motor commands to drive compensatory behavior.

Electrophysiological recordings have provided the temporal resolution to understand the complex dynamics of neural mechanisms, and their subcomponents, that regulate rapid interactions within sensorimotor networks for vocal feedback control. In humans and non-human primates, neural recordings during vocalization and listening tasks have shown that vocal production under normal auditory feedback (NAF) results in central cancellation and, therefore, suppression of auditory responses that match the internal representation of predicted feedback provided by efference copies (Behroozmand and Larson, 2011; Eliades and Wang, 2003; Houde et al., 2002). In contrast, vocal production under pitch-shifted altered auditory feedback (AAF) results in the enhancement (i.e. increase) of auditory neural responses to vocalization feedback compared with listening, which is assumed to reflect the mismatch between predicted and perceived signals (Behroozmand et al., 2009; Chang et al., 2013; Eliades and Wang, 2008; Greenlee et al., 2013). This latter effect has been argued to be accounted for by the top-down influence of efference copies on modulating auditory neural sensitivity for vocal feedback error detection and motor control. Findings of these studies have emphasized the involvement of efference copies in vocal feedback control; however, limitations of neurophysiological data, especially associated with their lack of spatial resolution, have precluded us from understanding the role of underlying brain networks in different aspects of such neural processes.

In a previous study from our lab (Behroozmand et al., 2018), we addressed this limitation via examining the lesion correlates of behavioral vocal impairments in individuals with brain damage due to left-hemisphere stroke. Our data revealed that, compared to neurotypical speakers, the stroke group showed deficits that corresponded to reduced magnitude of compensatory vocal responses to pitch-shifted auditory feedback alterations, and this effect was predicted by damage to distributed sensorimotor networks within the frontal, temporal, and parietal cortices. This finding raises a key question as to how such deficits are accounted for by the impairment of underlying neural mechanisms due to structural and functional abnormalities within the audio-vocal integration networks. To address this question, it is crucial to develop methods that rule out behavioral variability arising from peripheral vocal impairments (i.e., changes in laryngeal biomechanics resulting from damage to anatomical structures and/or muscle innervations) and establish a direct link between deficits in cortical neural responses and impaired sensorimotor brain regions.

In this study, we aimed to address this gap by using data from individuals with aphasia as a model to examine anatomical lesion and neural activity correlates of efference copies for audio-vocal integration and their impaired function due to left-hemisphere stroke. Pitch-shift stimuli (PSS) were utilized to alter the fundamental frequency (F0) in auditory feedback to investigate the underlying mechanisms of vocal sensorimotor function in speakers with aphasia and a matched control group. Electroencephalography (EEG) signals were concurrently recorded during vocalization and listening tasks to determine the neural correlates of efference copies in individuals with aphasia compared with controls. In addition, we used lesion-mapping analysis to study the relationship between structural brain abnormalities and functional EEG activity to determine how pathological changes in neural activity is predicted by damage to left-hemisphere sensorimotor networks implicated in audio-vocal integration. This highly novel approach allowed us to overcome limitations in identifying the lesion correlates of impaired efference copies as indexed by pathological modulation of neural activity during active vocalization compared with listening to the playback of self-vocalizations in individuals with aphasia compared with controls.

We examined the suppression of temporally specific event-related potential (ERP) components during vocal production compared with listening under NAF to probe deficits in vocal efference copy mechanisms of natural speech vowel sound vocalizations. The hypothesis was that

the impairment of efference copies would result in diminished suppression of auditory neural activity during vocal production in speakers with aphasia due to left-hemisphere stroke (Behroozmand and Larson, 2011; Houde et al., 2002). In addition, we hypothesized that the impairment of efference copies during vocalization error detection and motor correction would result in diminished enhancement of auditory neural responses (i.e. lowered sensory sensitivity) to pitch-shift AAF stimuli during vocal production compared with listening, and weaker compensatory responses (i.e. reduced motor correction) to alterations in the auditory feedback in individuals with aphasia (Behroozmand et al., 2009; Chang et al., 2013; Greenlee et al., 2013). Furthermore, we anticipated that pathological changes in neural activity correlates of efference copies would be accounted for by distinctive patterns of damage within the frontal, temporal, and parietal cortical areas that provide neural scaffolding and the interface for sensorimotor integration in the audio-vocal system (Hickok and Poeppel, 2004; Rauschecker, 2011).

2. Materials and methods

2.1. Subjects

A total of 34 subjects with post-stroke aphasia (22 males; age range: 42–80 yrs; mean age: 61.2 yrs), and 46 neurologically intact control subjects (23 males; age range: 44–82 yrs; mean age: 63.6 yrs) were recruited. All subjects with aphasia were recruited from the Center for the Study of Aphasia Recovery (C-STAR) at the University of South Carolina. All aphasia subjects had undergone testing with the Western Aphasia Battery (WAB) (Kertesz, 2007, 1982) as well as high-resolution T1-MRI scanning. At the time of testing, all subjects in the aphasia group were at least 6 months post stroke, with a mean age of 58.64 years old at the time of stroke (SD = 12.17) and a mean time post stroke of 39.83 months (SD = 53.66). The mean Aphasia Quotient, a measure of aphasia severity on the WAB was 64.97 (SD = 19.86). Based on the WAB aphasia classification system, the distribution of aphasia types across the 34 subjects was as following: Anomic = 7; Broca's = 18; Conduction = 8; and Global = 1. In addition, 19 subjects in the stroke group exhibited co-existing symptoms associated with apraxia of speech (AOS) with 18 subjects having mild-to-moderate and only 1 subject showing severe impairments as determined by the AOS Rating Scale (Strand et al., 2014). Neurotypical subjects in the control group had no history of speech, language, or neurological disorders, and were recruited from the greater Columbia, SC area through word-of-mouth and flyers. Subjects in both the post-stroke aphasia and control groups passed a binaural hearing screening and had thresholds of 40 dB or less at 500, 1000, 2000, and 4000 Hz. Informed consent was obtained from all subjects and the research was approved by the University of South Carolina Institutional Review Board. All subjects were monetarily compensated for their participation time.

2.2. Experimental design

The experiment was conducted in a sound-attenuated booth in which subjects' voice and EEG signals were recorded. All subjects in the aphasia and control groups completed a vocalization task under AAF in which they were instructed to produce steady phonations of the vowel sound /a/ at their conversational pitch and loudness after a human face visual cue was presented on the screen. During each vocalization trial, subjects maintained their vocalizations for 2–3 s while a brief pitch-shift stimulus with 200 ms duration was applied to alter their auditory feedback at randomized ± 100 cents (1 semitone) magnitudes. For each trial, the onset time of the pitch-shift stimulus was randomized between 750 and 1250 ms relative to the onset of vocalization. In addition, all subjects in the aphasia group and 25 out of 46 subjects in the control group completed the AAF paradigm during a listening task in which they received the same pitch-shift stimuli while they remained silent

and listened to the playback of their own pre-recorded vowel sound vocalizations following the presentation of a human ear visual cue on the screen. For subjects who completed the listening task, the order of vocalization and listening trials was interleaved so that each vocalization trial was immediately followed by the playback of its pre-recorded version during the succeeding listening trial. The inter-trial interval (ITI) between vocalization and listening trials was approximately 2–3 s. During vocalization trials, the gain of the auditory feedback signal was adjusted 10 dB higher than subjects' voice level to partially mask bone or air-borne conduction effects. In addition, the gain of the auditory feedback was equalized between both vocalization and listening tasks. Data were collected for 200 vocalization and 200 listening trials with 100 trials per stimulus direction during each task, separately. At the beginning of each session, subjects were provided with a brief practice to ensure they understood the experimental tasks and were producing vowel sounds steadily and with adequate length. A major advantage of our AAF paradigm is that it involves tasks that are both motorically and perceptually simple (i.e. steady vowel productions and listening to their playback), and therefore, could be successfully performed even by subjects in the stroke group who often exhibit limited production and comprehension abilities due to co-existing aphasia and AOS symptoms. For all subjects, the experimenters verified correct task performance during practice session before data collection started. Subjects were monitored throughout the data recording session to ensure that they continued vocal production and listening tasks as directed and were offered breaks if they appeared to be experiencing vocal fatigue. All experimental parameters including the timing, order, and the type of visual cues and pitch-shift stimuli were controlled by a custom-made program in Max 5.0 (Cycling '74, Inc). Transistor-transistor logic (TTL) pulses were also generated to synchronize the timing of visual cues and pitch-shift stimuli with subjects' behavioral voice and neurophysiological EEG signals during the experiment.

2.3. Voice data acquisition and analysis

Subjects' voice signal was picked up using a head-mount AKG condenser microphone (model C520), amplified by a Motu Ultralite-MK3, and recorded at 44.1 kHz on a laboratory computer. Data were analyzed to extract the behavioral measure of vocal compensation responses relative to the onset of pitch-shift stimuli. First, the pitch frequency of the recorded voice signals was extracted in Praat (Boersma and Weenik, 1996) using an autocorrelation method and then exported to a custom-made MATLAB code for further processing. The extracted pitch frequencies were segmented into epochs ranging from –100 ms before to 500 ms after the onset of pitch-shift stimuli. Pitch frequencies were then converted from Hertz to Cents scale to calculate vocal compensation magnitude in response to the pitch-shift stimulus using the following formula:

$$\text{Vocal Compensation Magnitude} = 1200 \times \log_2(F/F_{\text{Baseline}})$$

Here, F is the post-stimulus pitch frequency and F_{Baseline} is the baseline pitch frequency from –100 to 0 ms pre-stimulus. Artefactual responses to pitch shifts in the auditory feedback due to large-magnitude voluntary vocal pitch modulations were rejected by removing trials in which vocal responses exceeded ± 200 cents in magnitude. The extracted pitch contours were then averaged on remaining trials for each subject in response to upward and downward pitch shifts across aphasia and control groups, separately.

2.4. EEG data acquisition and analysis

Electrophysiological responses were measured during the experiment by recording EEG signals from 64 BrainVision actiCAP active electrodes (Brain Products GmbH, Germany) following the standard 10–10 montage and a common average reference. A BrainVision actiCHamp amplifier (Brain Products GmbH, Germany) integrated with

the Pycorder software was used to record EEG signals at 1 kHz sampling rate after applying a low-pass anti-aliasing filter with 200 Hz cut-off frequency. Electrode impedances were kept below 5 k Ω for all channels. The EEGLAB toolbox (Delorme and Makeig, 2004) was used for pre-processing of data by first band-pass filtering the EEG signals at 1–30 Hz (–24 dB/oct), correcting for muscle artefacts (e.g., eye movement, saccades, blinks etc.) using independent component analysis (ICA), and then segmenting them into epochs from –200 to 500 ms relative to the onset of voice and pitch-shift stimuli. The extracted epochs were baseline corrected at –200 to –100 ms and then averaged across trials to calculate event-related potentials (ERPs) for each subject across groups (aphasia vs. control) and tasks (vocalization vs. listening), separately. The ERP components at different latencies reflect positive or negative voltage deflections recorded on the surface of the scalp as a result of phase-synchronized neuronal activities time-locked to the onset of different events. In our analysis, ERP responses for each subject were calculated for vocalization and listening tasks during NAF and AAF conditions by averaging neural responses for a minimum number of 150 epochs time-locked to the onset of voice and pitch-shift stimuli, respectively.

2.5. Statistical analysis

Statistical analysis of data was performed in SPSS v.27 using general linear models (GLMs) to analyze the effects of group (aphasia vs. control) on the magnitude of behavioral vocal compensation responses to AAF. ERP components were analyzed using topographical analysis of variance (TANOVA) in CURRY 8.0 (Compumedics Neuroscan, Inc) to examine the effects of group (aphasia vs. control), task (vocalization vs. listening), and their interactions on responses to NAF at the onset of vocalization and AAF at the onset of pitch-shift stimuli. The choice of these factors was prioritized based on our research questions and hypotheses to determine how neural responses to NAF and AAF are modulated during vocalization and listening tasks for subjects in the aphasia and control groups, irrespective of the difference in the direction of pitch-shift stimuli during AAF. Therefore, to keep our analysis consistent and comparable across NAF and AAF conditions, stimulus direction was not included as a factor of interest. In addition, this approach helped reduce the number of factors in our analysis to maintain statistical power for the sample size in the present study. TANOVA was used as a non-parametric permutation model to determine statistical significance by assessing global dissimilarity of neural activities in spatially organized topographical maps while correcting data for multiple comparisons on a temporal basis. The main advantage of TANOVA is that it allows to test specific hypotheses via an independent choice of electrodes and time points yielding significant results without requiring a-priori assumptions about the spatiotemporal characteristics of the underlying data (Wagner et al., 2017). In our analysis, we used TANOVA to identify regions of interest (ROIs) for ERP components with significant effects, and then submitted those data to GLM analysis to further examine the main effects of group, task, and their interactions. Data normality and homogeneity of variance assumptions were examined using the Shapiro-Wilk and Mauchly's sphericity tests, respectively. For data violating the normality assumption, a rank-based inverse normal transformation was applied (Templeton, 2011) and p-values were reported using Greenhouse-Geisser's correction for data violating homogeneity of variances assumption. Partial Eta squared (η_p^2) was reported as an index of the effect size for significant main effects and post-hoc tests for significant interactions were performed using t-tests with Bonferroni's correction with Cohen's d reported as a measure of effect size.

2.6. MRI data acquisition

MRI data were acquired with a 3T Siemens Trio system fitted with a 12-channel head-coil. All subjects with aphasia were scanned with two MRI sequences: (i) T1-weighted imaging sequence using a 3D MP-RAGE (magnetization-prepared rapid-gradient echo) [TFE (turbo field echo)]

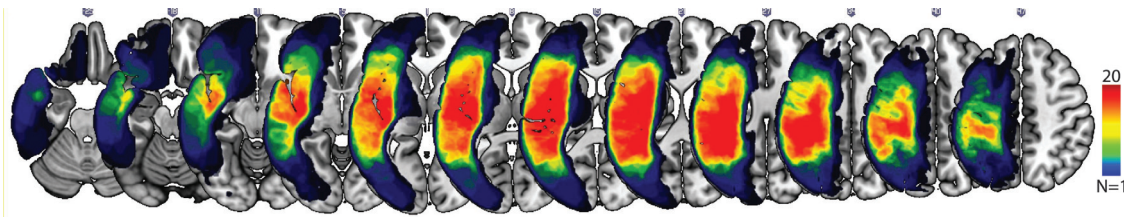


Fig. 1. Lesion overlap maps in individuals with post-stroke aphasia ($n = 34$). The maps show lesion distribution on coronal (top) slices in MNI space for the sample, with warmer colors representing more lesion overlap across aphasic speakers (dark red areas represent lesion overlap across at least $N = 20$ stroke subjects). Maximum overlap areas include the left superior and middle temporal gyrus, Heschl's gyrus, precentral and postcentral gyrus, inferior and middle frontal gyrus, Rolandic operculum, insula, supramarginal gyrus, angular gyrus, and inferior and superior parietal gyrus where nearly 60% (20 out of 34) of subjects had damage.

sequence with voxel size = 1 mm^3 , FOV (field of view) = $256 \times 256 \text{ mm}$, 192 sagittal slices, 9° flip angle, TR (repetition time) = 2250 ms, TI (inversion time) = 925 ms, TE (echo time) = 4.15 ms, GRAPPA (generalized autocalibrating partial parallel acquisition) = 2, and 80 reference lines; and (ii) T2-weighted MRI for the purpose of lesion demarcation with a 3D sampling perfection with application optimized contrasts by using different flip angle evolutions protocol with the following parameters: voxel size = 1 mm^3 , FOV = $256 \times 256 \text{ mm}$, 160 sagittal slices, variable flip angle, TR = 3200 ms, TE = 352 ms, and no slice acceleration. The same slice center and angulation were used as in the T1 sequence.

2.7. Preprocessing of structural MRI

Images were converted to NIfTI format using dcm2nii (Li et al., 2016). Stroke lesions were demarcated by a neurologist (L.B.) in MRIcron (Rorden et al., 2012) on individual T2 MRIs (in native space). Note that the lesions demarcated on the T2-MRI images were used for the purpose of normalization and to estimate lesion size, which was included as a covariate factor in the lesion-mapping analyses. The greatest gray-matter lesion overlap among the aphasic speakers was in the left superior and middle temporal gyrus, Heschl's gyrus, precentral and postcentral gyrus, inferior and middle frontal gyrus, Rolandic operculum, insula, supramarginal gyrus, angular gyrus, and inferior and superior parietal gyrus where nearly 60% (20 out of 34) of subjects had damage. The overlaid maps of lesion distribution across all aphasic subjects are shown in Fig. 1. Preprocessing began with the coregistration of the T2 MRI to match the T1 MRIs, aligning lesions to native T1 space. Images were warped into standard space using a custom MATLAB script according to the enantiomorphic segmentation-normalization method (Nachev et al., 2008) to warp images into an age-appropriate template included with the SPM Clinical Toolbox (Rorden et al., 2012). The normalization parameters were used to reslice lesions into standard space using linear interpolation, with the resulting lesion maps stored at $1 \times 1 \times 1 \text{ mm}$ resolution and binarized using a 50% threshold. This latter procedure was undertaken because interpolation can lead to fractional probabilities, and therefore, this step ensures that each voxel is categorically either lesioned or unlesioned without biasing overall lesion volume. All normalized images were visually inspected to verify the quality of preprocessing.

2.8. Regions of interest

The primary analyses of this study related z-score-transformed mean image intensities corrected for family-wise error due to multiple comparisons in 12 *a priori* selected regions of interest (ROIs) in the left hemisphere (Table 1) to the neurophysiological measures of ERP modulation in response to normal or altered vocal auditory feedback in speakers with aphasia compared with normal control subjects. These ROIs were selected based on a review of the relevant literature to encompass cortical regions within the dorsal stream networks implicated in vocal sensorimotor processing (Fridriksson et al., 2016; Hickok and Poeppel, 2007, 2004, 2000; Poeppel and Hickok, 2004). The ROIs were se-

Table 1

Left-hemisphere regions of interest (ROIs) used in lesion-mapping analysis.

Inferior frontal gyrus (pars opercularis)
Inferior frontal gyrus (pars orbitalis)
Inferior frontal gyrus (pars triangularis)
Precentral gyrus
Postcentral gyrus
Rolandic operculum
Angular gyrus
Supramarginal gyrus
Inferior parietal gyrus
Heschl's gyrus
Superior temporal gyrus
Middle temporal gyrus

lected from the “Automated Anatomical Labeling (AAL)” atlas (Tzourio-Mazoyer et al., 2002) in which the gray matter tissue was segmented after a 50% threshold was applied during the normalization process to obtain a smoothed and interpolated lesion mask that minimizes jagged edges at the boundary between lesion and gray matter tissue.

2.9. Lesion-mapping analysis

The NiiStat toolbox (www.nitrc.org/projects/niiestat) was used to conduct lesion-mapping analyses to identify localized brain lesions within the selected ROIs that predict impaired efferece copy and vocal sensorimotor integration mechanisms, as indexed by modulation of neurophysiological responses in aphasic speakers compared with controls. In order to obtain a normalized distribution of neurophysiological response modulation within the aphasic group, a measure of ERP Modulation Index (EMI) was calculated for each aphasic subject based on the log-transformed ratio of ERP response modulation relative to the mean modulation of the same ERP response across all control subjects according to the following formula:

$$ERP \text{ Modulation Index} = 10 \times \log_{10} \left(\Delta ERP_{Aphasia} / \overline{\Delta ERP}_{Control} \right)$$

In this formula, $\Delta ERP_{Aphasia}$ is the magnitude of ERP modulation for a given component during vocalization vs. listening task in individuals with aphasia, and $\overline{\Delta ERP}_{Control}$ is the mean magnitude of ERP modulation for the same component during vocalization vs. listening across the control group. The log-transformation function was used to ensure that the data were normally distributed for statistical analysis. Lesion-mapping analysis of neurophysiological responses was performed using a computational model in which the neuroanatomical maps of individual lesion volumes within each ROI were regressed against the measures of ERP modulation index to determine lesion correlates of impaired vocal sensorimotor processing in response to normal and altered auditory feedback. For each ERP component, the corresponding time window was divided into the first vs. second half to determine lesion predictors of early vs. late phases of neural responses with higher temporal resolution, and results were reported separately if different regions were identified. ROIs for which at least ten subjects had damage were included

and statistical significance was determined by ROI-based thresholding at 3000 permutations to control for multiple comparisons at $\alpha = 0.05$. This procedure yielded standardized brain maps showing the statistical likelihood that lesions in localized brain regions predict impaired vocal efference copy and sensorimotor integration function based on modulation of ERP components in aphasia relative to control subjects. The statistical brain maps were first calculated using t-scores with degrees of freedom $df = n - 2$ (n : total number of samples), and then transformed into z-scores for standardization. Since the overall lesion size could potentially be correlated with diminished neurophysiological responses, this parameter was regarded as a nuisance variable of no interest, and therefore, was entered as a covariate in order to factor out its effect in the lesion-mapping analysis of ERP data. Type-II error (β) was reported as an estimate of statistical power for lesion-mapping analysis.

3. Results

3.1. Neural responses to NAF

TANOVA analysis of ERP responses to NAF at vowel vocalization onset indicated significant main effects of group, task, and group \times task interaction in multiple time windows. The main effect of group was associated with significantly stronger ERP activity in control compared with aphasia in three time windows, one before and two after the onset of voice: 1) the P0 component (i.e. the positive potential emerging before the onset of vocalization) at -50 to 0 ms ($p < 0.01$) with the largest contribution from the left fronto-central electrodes, 2) the N1 component (i.e. the first negative potential after vocalization onset) at 25 – 100 ms ($p < 0.01$), and 3) the P1 components (i.e. the first positive potential after vocalization onset) at 150 – 250 ms ($p < 0.01$) with the largest contributions from the left fronto-central and temporo-parietal electrodes. The overlaid profiles of grand-average ERP responses to NAF across left fronto-central electrodes for vocalization and listening tasks are shown in Fig. 2A–B. The topographical distribution maps are shown separately for the aphasia and control groups in Fig. 2C. For the main effect of task, ERP activity was significantly stronger for vocalization vs. listening before voice onset for the P0 component ($p < 0.01$). However, after the onset of vocal production, an opposite response modulation pattern was observed and ERP activity was significantly suppressed during vocalization vs. listening for the N1 ($p < 0.01$) and P1 ($p < 0.01$) components. In order to further examine the significant group \times task interaction, post-hoc tests with Bonferroni's correction were conducted for each group to analyze ERP responses to NAF during vocalization and listening tasks. Results revealed significant enhancement of the P0 component during vocalization vs. listening ($p < 0.05$), and significant suppression of the N1 and P1 components in control subjects ($p < 0.01$); however, for subjects in the aphasia group, the only significant effect was observed as vocalization-induced suppression of the N1 component ($p < 0.01$). In addition, we found that the magnitude of vocalization-induced enhancement of the P0 component was significantly smaller in aphasia vs. controls ($p < 0.01$), and the magnitude of vocalization-induced suppression of the N1 ($p < 0.05$) and P1 ($p < 0.01$) components was significantly smaller in aphasia vs. controls (Fig. 2D).

3.2. Neural responses to AAF

TANOVA analysis of ERP responses to AAF stimuli indicated significant main effects of group, task, and group \times task interaction in multiple time windows. The main effect of group was associated with significantly stronger ERP activity in control compared with aphasia in three time windows after the onset of pitch-shift stimuli: 1) the P1 component (i.e. the first positive potential after AAF) at 50 to 100 ms ($p < 0.01$) with the largest contribution from the left fronto-central, 2) the N1 component (i.e. the first negative potential after AAF) at 100 – 150 ms ($p < 0.01$), and 3) the P2 components (i.e. the second positive potential after AAF) at 200 – 300 ms ($p < 0.01$) with the largest contributions from the

left fronto-central and temporo-parietal electrodes. For the main effect of task, ERP activity was significantly stronger for vocalization vs. listening only for the P1 ($p < 0.05$) and N1 ($p < 0.01$) component. The overlaid profiles of grand-average ERP responses to AAF across left fronto-central electrodes for vocalization and listening tasks are shown in Fig. 3A–B. The topographical distribution maps are shown separately for the aphasia and control groups in Fig. 3C. The significant group \times task interaction was further examined using post-hoc tests with Bonferroni's correction for each group separately. Results revealed significantly stronger vocalization-induced response enhancement of the P1 ($p < 0.01$) and N1 ($p < 0.01$) components in control subjects; however, for subjects in the aphasia group, the only significant effect was observed as vocalization-induced response enhancement of the N1 component ($p < 0.01$). In addition, we found that the magnitude of vocalization-induced response enhancement was significantly smaller in the aphasia vs. control group for the P1 ($p < 0.05$) and N1 ($p < 0.01$) components (Fig. 3D).

3.3. Vocal responses to AAF

The magnitude of compensatory vocal responses to pitch-shift stimuli were analyzed within three time windows: 1) 150 – 250 ms for response rising time when vocal compensation ascended toward the peak, 2) 250 – 350 ms for response peak time when vocal compensation reached the maximum magnitude, and 3) 350 – 450 ms for response rebound time when vocal compensation descended toward the pre-stimulus baseline (Fig. 4). The time windows were selected to capture the temporal dynamics of vocal compensatory response profiles in the aphasia and control groups. Results of our analysis for the vocal responses to pitch-shift stimuli revealed a main effect of group with significantly diminished magnitude of compensation in aphasia vs. control only within the rise ($F(1,78) = 11.04$, $p < 0.01$, $\eta_p^2 = 0.124$) and the peak ($F(1,78) = 5.64$, $p < 0.05$, $\eta_p^2 = 0.087$) time windows at 150 – 250 ms and 250 – 350 ms, respectively. As can be seen in Fig. 4A, both groups compensated by generating responses that deviated from the baseline at approximately 100 ms following the onset of pitch-shift stimuli in the auditory feedback. The group and individual subject data for vocal compensation responses averaged within the rise and peak time windows (150 – 350 ms) are shown in Fig. 4B ($t(78) = 2.61$, $p < 0.05$, $d = 0.591$). In addition, the histogram plot and normal distribution curves for vocal compensation magnitudes in aphasia vs. control groups averaged within the rise and peak time windows is presented in Fig. 4C.

3.4. Lesion-mapping analysis

The relationship between stroke-induced cortical damage and impaired efference copies for vocal production under NAF was investigated using lesion-mapping analysis of the relative degree of diminished vocalization-induced modulation of the P0, N1, and P1 components in response to NAF as measured by the EMI in speakers with aphasia compared with controls. As described earlier, the EMI is a neural indicator that measures the degree of aberrant vocalization-induced modulation of ERP activity in individuals with aphasia normalized to the mean of the control group (see Section 2.9). In our analysis, EMIs were calculated for the ERP components and electrodes showing a significant group \times task interaction based on the TANOVA results. For ERP components, amplitudes were averaged within their specified time window and were submitted to ROI-based lesion-mapping analysis for each individual aphasia subject. Results yielded statistically significant effects for damage to the left inferior frontal gyrus pars orbitalis predicting diminished vocalization-induced enhancement of the P0 component in responses to NAF in speakers with aphasia ($z = -2.57$, $p < 0.05$, $\beta = 0.26$). In addition, significant reduction in vocalization-induced suppression of the early phase of N1 response (N1a: 0 – 50 ms) in aphasia was predicted by damage to a network comprising the left inferior frontal gyrus pars orbitalis ($z = 3.28$, $p < 0.05$, $\beta = 0.08$) and precentral gyrus areas ($z = 2.81$,

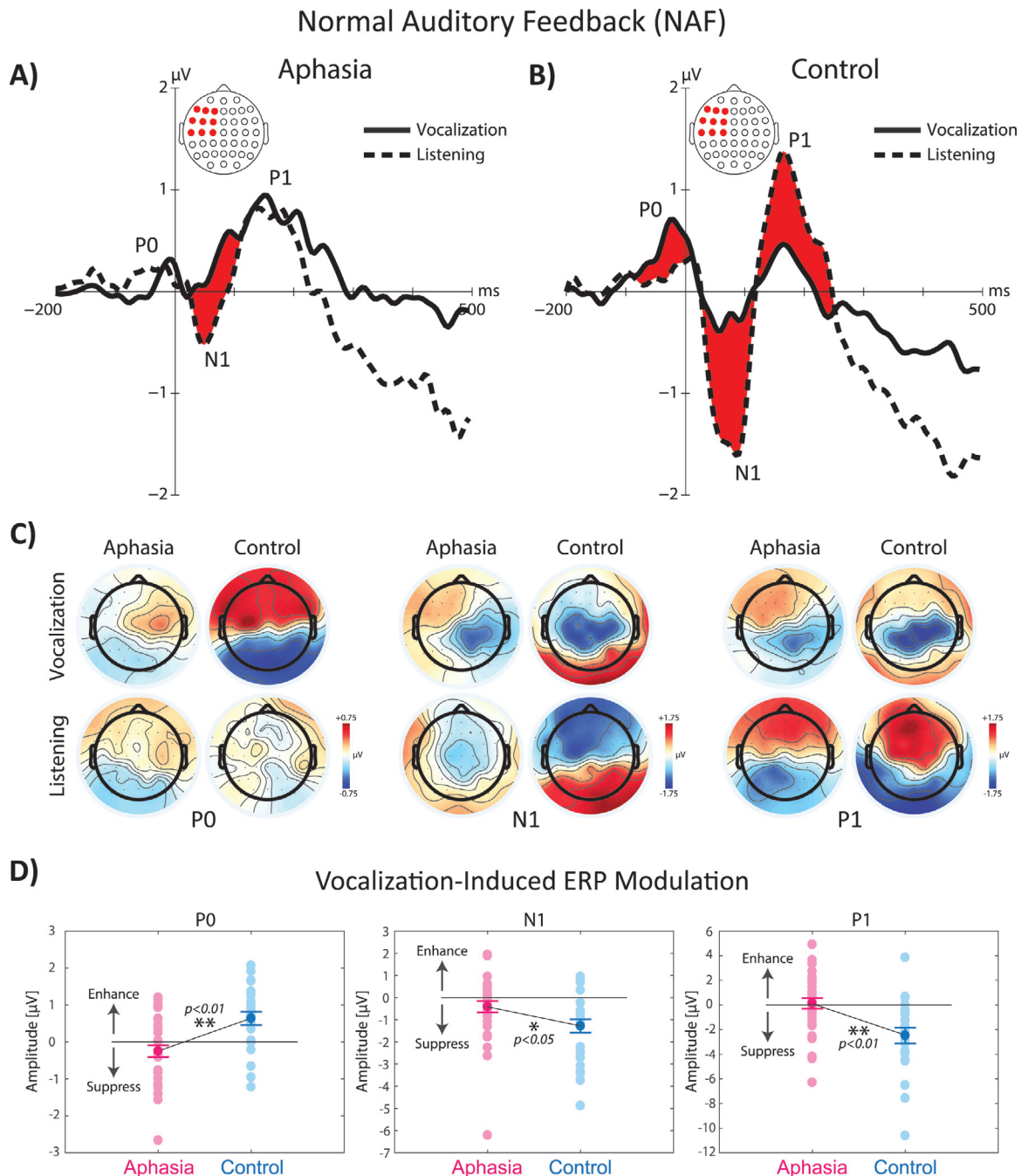


Fig. 2. The overlaid profiles of ERP responses to NAF at the onset of vowel sound production during vocalization and listening tasks in fronto-central electrode in A) aphasia (left panel) vs. B) control subjects (right panel). Highlighted areas in red color indicate time ranges of significant differences between vocalization vs. listening. C) Shows the topographical distribution maps of ERP activity in 64 electrodes for the P0, N1, and P1 ERP responses for vocalization and listening tasks across aphasia and control subjects. D) Shows the raster plot of individual subjects and group differences in the amplitude of vocalization-induced modulation of ERP activity in aphasia vs. control group.

$p < 0.05$, $\beta = 0.19$). Furthermore, significant reduction in vocalization-induced suppression of the late phase of N1 response (N1b: 50–100 ms) in aphasia was predicted by damage to a network comprising the left inferior frontal gyrus pars orbitalis ($z = 2.79$, $p < 0.05$, $\beta = 0.19$) and supramarginal gyrus areas ($z = 2.54$, $p < 0.05$, $\beta = 0.27$). Lastly, significant reduction in vocalization-induced suppression of the P1 response in aphasia was predicted by damage to a network comprising the left precentral gyrus ($z = 2.90$, $p < 0.05$, $\beta = 0.16$) and superior temporal gyrus areas ($z = 2.72$, $p < 0.05$, $\beta = 0.21$). Our results showed that the extent of damage involving these regions was positively correlated with

N1 and P1 modulation during vocalization vs. listening, indicating that greater damage to these areas was associated with reduced vocalization-induced suppression of these ERP components in speakers with aphasia due to left-hemisphere stroke. The overlaid maps of statistically significant lesion predictors of diminished vocalization-induced suppression of ERPs in response to NAF are shown in Fig. 5A.

For neural responses to AAF, ROI-based lesion-mapping analysis of EMI was conducted for ERP components with significant group \times task interaction to identify damage to left-hemisphere brain networks associated with the impairment of efference copies during vocal error detec-

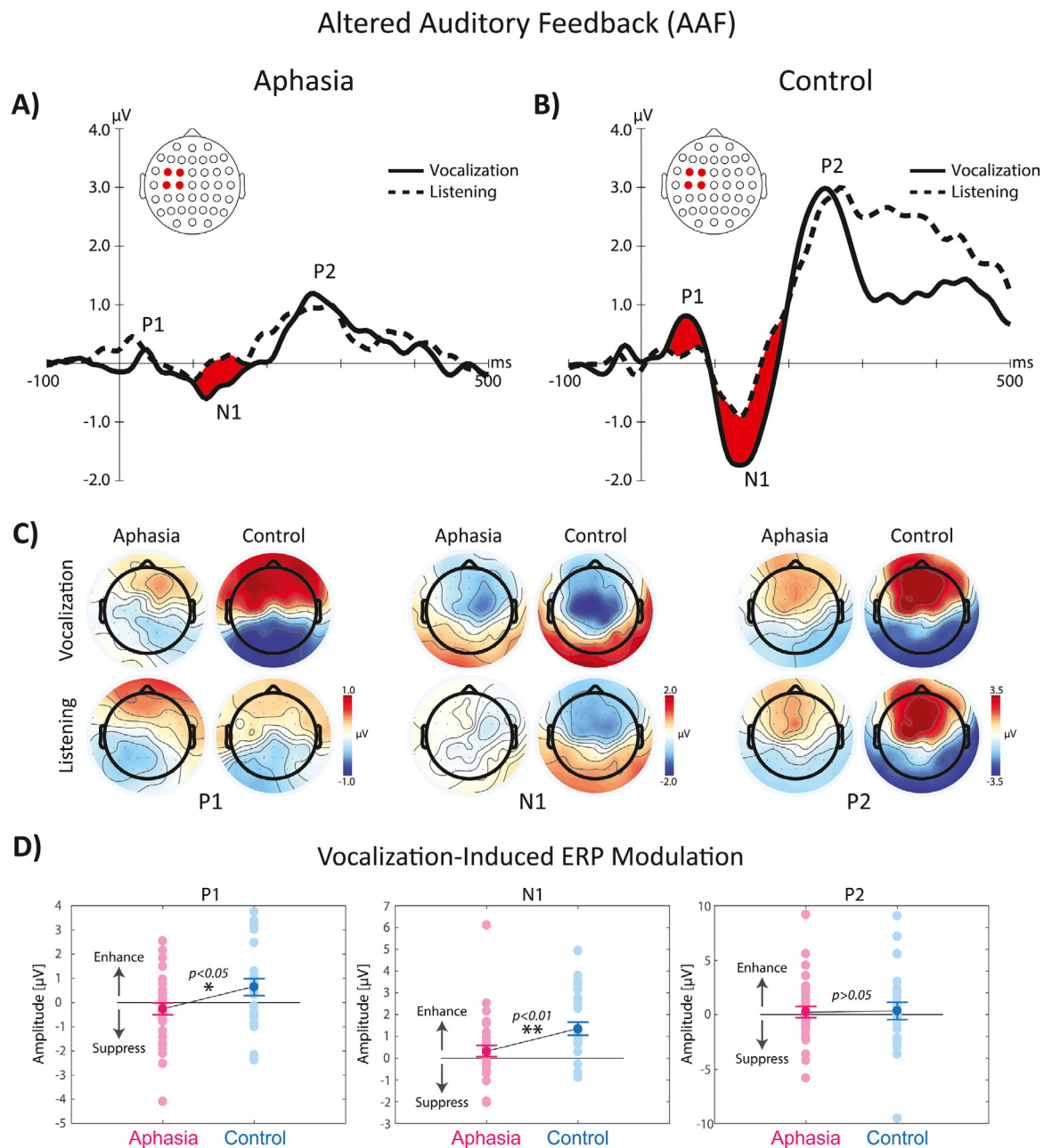


Fig. 3. The overlaid profiles of ERP responses to AAF at the onset of pitch-shift stimuli during vowel sound vocalization and listening tasks in fronto-central electrode in A) aphasia (left panel) vs. B) control subjects (right panel). Highlighted areas in red color indicate time ranges of significant differences between vocalization vs. listening. C) Shows the topographical distribution maps of ERP activity in 64 electrodes for the P1, N1, and P2 ERP responses for vocalization and listening tasks across aphasia and control subjects. D) Shows the raster plot of individual subjects and group differences in the amplitude of vocalization-induced modulation of ERP activity in aphasia vs. control group.

tion and motor correction in post-stroke aphasia. Results yielded statistically significant effects for damage to a network comprising the inferior frontal gyrus pars orbitalis ($z = -2.78, p < 0.05, \beta = 0.20$), inferior frontal gyrus pars opercularis ($z = -2.74, p < 0.05, \beta = 0.21$), and angular gyrus ($z = -2.85, p < 0.05, \beta = 0.17$) in the left hemisphere predicting the diminished vocalization-induced enhancement of the P1 component in response to AAF in aphasia. For the N1 component, diminished response enhancement in aphasia was found to be predicted by damage to a left-hemisphere network comprising the angular gyrus ($z = -2.69, p < 0.05, \beta = 0.22$) and supramarginal gyrus ($z = -2.63, p < 0.05, \beta = 0.24$) areas. Our results showed that the extent of damage involving these regions was negatively correlated with P1 and N1 modulation during vocalization vs. listening, indicating that greater damage to these ar-

eas was associated with reduced vocalization-induced enhancement of these ERP components in speakers with aphasia due to left-hemisphere stroke. The overlaid maps of statistically significant lesion predictors of diminished vocalization-induced enhancement of ERPs in response to AAF are shown in Fig. 5B.

3.5. Correlation analysis

The relationship between efference copies and behavioral responses to AAF was investigated by examining the correlations between neural activity in electrodes that showed significant vocalization-induced modulation of ERPs and the measures of vocal compensation for pitch-shift stimuli in speakers with aphasia and controls. Results of the analysis re-

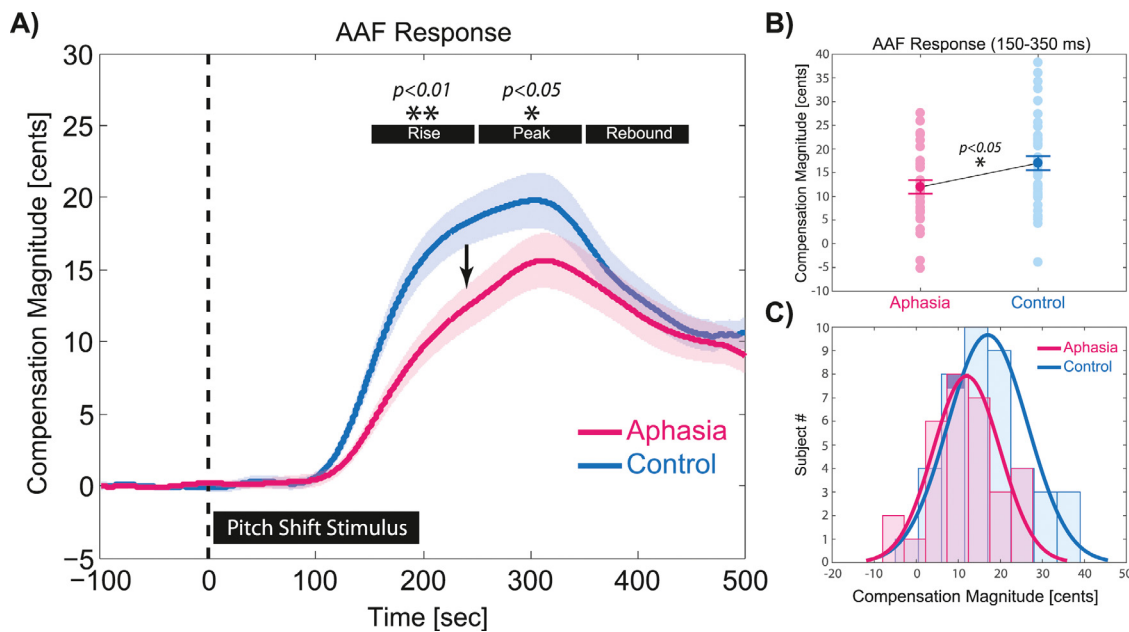


Fig. 4. A) The overlaid profiles of vocal compensation responses averaged for upward (+100 cents) and downward (-100 cents) pitch-shift stimuli in aphasia ($n = 34$) and control ($n = 46$) subjects. Highlighted time windows represent the rise (150 – 250 ms), peak (250 – 350 ms), and rebound (350 – 450 ms) phases of responses in each group. B) Shows the raster plot of individual subjects and group differences in the magnitude of vocal compensations averaged within the rise and peak time windows (150 – 350 ms) in aphasia vs. control group. C) Shows the histogram plot and normal distribution curves for vocal compensation magnitudes in aphasia vs. control groups averaged within the rise and peak time windows.

vealed that vocalization-induced enhancement of the P0 ERP component in response to NAF was positively correlated with stronger vocal compensation responses to pitch-shift stimuli in the control group ($r = 0.49$, $p < 0.05$); however, no such effect was found for individuals with aphasia ($r = -0.02$, $p > 0.05$). The topographical distribution maps of correlation between the P0 component and vocal compensation responses along with the scatter plots of a representative electrode over the left fronto-central electrode (F5) are shown separately for the aphasia and control groups in Fig. 6. In addition, for responses to NAF, vocalization-induced suppression of the P1 component was negatively correlated with stronger vocal compensation responses for controls ($r = -0.48$, $p < 0.05$), but not the aphasia group ($r = -0.25$, $p > 0.05$). For responses to AAF, we found a positive correlation between vocalization-induced enhancement of the N1 ERP component and stronger vocal compensation responses to pitch-shift stimuli only in the control group ($r = 0.69$, $p < 0.01$) but not aphasia ($r = 0.33$, $p > 0.05$).

4. Discussion

Previous studies on behavioral lesion-mapping analysis of impaired vocal production and speech have been limited by lack of high-temporal resolution neurophysiological data to investigate the rapid neural dynamics of these mechanisms and their deficits in post-stroke aphasia (Behroozmand et al., 2018; Fridriksson et al., 2013). On the other hand, studies on neurotypical adults that used high-temporal resolution recordings of electrophysiological data did not provide a unified account to determine the neuroanatomical correlates of sensorimotor integration mechanisms implicated in vocal feedback control (Behroozmand et al., 2009; Chang et al., 2013; Greenlee et al., 2013). In the present study, we aimed to overcome these limitation by combining data from high-temporal resolution ERP components with lesion profiles to determine how deficits in vocal feedback control mechanisms at different time scales are accounted for by the functional and structural brain abnormalities in individuals with post-stroke aphasia. This novel approach allowed us to examine the relationship between aberrant neural activity directly recorded from the surface of the scalp and their underlying

stroke-induced brain damage in aphasic speakers. Our data provided evidence for impaired vocal feedback control under normal and altered auditory feedback as indexed by the diminished pattern of vocalization-induced modulation of temporally specific ERP components in aphasia compared with controls. We found that when speakers with aphasia produced a speech vowel sound under normal feedback, vocalization-induced suppression of their ERP responses were significantly diminished over the left fronto-central and temporo-parietal electrodes compared with control subjects. Lesion-mapping analysis revealed that this effect was predicted by damage to a distributed sensorimotor network comprising regions within the frontal, temporal, and parietal cortices. In addition, our data revealed deficits in vocal feedback error detection and motor correction under altered feedback by showing significant reduction in vocalization-induced enhancement of ERP responses in the left fronto-central and temporo-parietal electrodes in aphasia compared with controls. Lesion-mapping analysis showed that this latter effect was predicted by damage to a distinguishable but overlapping sensorimotor network within the frontal and parietal cortices. Moreover, our analysis revealed that vocalization-induced modulation of ERP responses to normal and altered auditory feedback were correlated with the magnitude of compensatory vocal responses to pitch-shift stimuli in control speakers; however, no such effect was observed for individuals in the aphasia group. These findings provide multi-modal evidence from functional neurophysiology, neuroanatomical lesion data, and behavioral responses confirming the role of feedback control mechanisms in vocal production and their impairment in individuals with post-stroke aphasia.

The patterns of brain damage associated with impaired vocal feedback control in aphasia have highlighted two distinct, and yet overlapping, networks within the left-lateralized dorsal stream pathways that regulate vocal production under normal and altered auditory feedback. According to the dual-stream model of speech (Hickok and Poeppel, 2007, 2004), the dorsal stream provides a neural interface to integrate sensorimotor networks for online monitoring of auditory feedback, and to compare predicted and actual inputs for vocal error detection and motor correction. In the context of this model, the sensorimotor interface

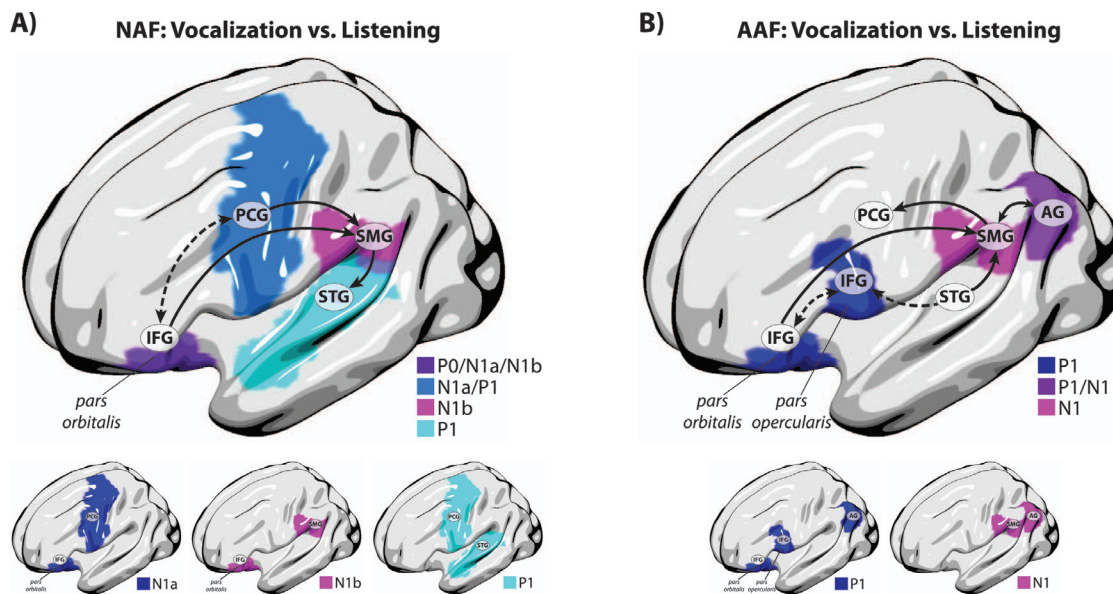


Fig. 5. Anatomical representation of significant lesion correlates of aberrant vocalization-induced modulation of ERP activity in aphasia for A) the P0, N1, and P1 components in responses to NAF, and B) the P1, and N1 components in responses to AAF. In panel A, the dashed arrow highlights connectivity between the left frontal cortical areas involved in internal monitoring and regulation of vocal output before and shortly after the onset of vocalization when auditory feedback is not available, and the solid arrows highlight connectivity between the left frontal, parietal, and temporal cortical areas implicated in auditory feedback-based vocal monitoring and regulation during vocalization under NAF. In panel B, the solid arrows highlight connectivity between the left frontal, parietal, and temporal cortical areas implicated in auditory feedback-based vocal error detection and motor control during vocalization under AAF (dorsal stream), and the dashed arrows highlight connectivity between the left frontal and temporal cortical areas that provide an interface to process AAF errors for updating the neural representations of the internal forward model in the inferior frontal cortex (ventral stream).

provides two potential mechanisms for vocal control: First, an internal mechanism in which forward predictions are checked against their sensory targets allowing for error correction prior to and shortly after the onset of utterance when the actual auditory feedback has not yet become available (Hickok, 2012); and second, an external mechanism in which sensory representations of vocal output are compared against the actual auditory feedback for error detection and motor correction. Evidence from previous studies has supported the notion of an internal control mechanism by showing that neurotypical speakers rapidly (i.e., within ~50 ms) move their peripheral productions toward the center of their vowel space in the absence of any external feedback alterations, and this effect is associated with reduced motor-induced suppression of neural activities compared with production trials that closely match sensory targets near the center of the intended vowel sounds (Cheng et al., 2021; Niziolek et al., 2013). In the present study, we found that when speakers with aphasia produced vowel sounds under normal feedback, damage to left-hemispheric frontal networks involving the inferior frontal *pars orbitalis* and precentral gyri predicted the impairment of vocal feedback control in the earlier phases of neural processing as indexed by the aberrant modulation of short-latency N1a ERP responses at 0–50 ms following vocalization onset. In addition, our data showed that the *pars orbitalis* was also a robust predictor of diminished neural response modulations prior to the onset of vocalization as indexed by the P0 ERP component elicited at latencies around –50 ms before vowel sound production. These findings suggest that areas within the left frontal cortex are involved in regulating internal control mechanisms before and shortly after the onset of vocalization when external auditory feedback is not available (see the dashed arrow in Fig. 5A). Based on previous evidence (Conner et al., 2019), we suggest that *pars orbitalis* is likely to play a high-level regulatory role enabling efficient engagement of the vocal sensorimotor networks for adaptive responses to changing feedback conditions while the precentral gyrus may be involved in maintaining efference copies for internal vocal motor correction. Results of our analysis also revealed that vocalization-induced modulation of ERP responses to normal auditory feedback were correlated with compen-

satory behavior only in the control, but not aphasia, group. This finding suggests that the internal mechanisms can mediate online vocal control via generating rapid motor commands to correct erroneous productions even in the absence of incoming auditory feedback information.

Moreover, our data revealed that when speakers with aphasia produced vowel sounds under normal auditory feedback, the aberrant patterns of neural activity modulation in the long-latency N1b (50–100 ms) and P1 (150–250 ms) ERPs were predicted by damage to supramarginal and superior temporal gyri networks of the left hemisphere, respectively. Based on the latency of their underlying neural response components, we argue that these latter networks are likely implicated in regulating external vocal control mechanisms via matching sensory predictions against actual auditory feedback for error detection and motor control. In this context, the supramarginal gyrus is a candidate for the sensorimotor interface that transforms motor efference copies (i.e. forward predictions) to sensory representations for feedback monitoring and/or error detection and converts sensory errors into motor representations that drive corrective vocal behavior in response to external alternations. This notion is supported by previous studies emphasizing the role of a localized region in the Sylvian fissure at the boundary between the parietal and temporal lobes (i.e., area Spt) in speech sensorimotor integration (Hickok et al., 2003, 2009). We argue that transformations between motor efference copies and their sensorimotor representations within supramarginal gyrus is mediated via white matter tracts of the anterior segment of the arcuate fasciculus connecting frontal areas of *pars orbitalis* within the inferior frontal gyrus and ventral precentral gyrus with the posterior, inferior parietal lobe (Cocquyt et al., 2020; Fridriksson et al., 2013). This notion converges with evidence from previous neuroimaging studies indicating that damage to the posterior region of the parietal-temporal boundary, which includes Spt, accounts for speech repetition impairments as a hallmark of sensorimotor deficit in individuals with conduction aphasia (Buchsbaum et al., 2011; Hickok, 2012; Hickok et al., 2011; Rogalsky et al., 2015). In addition, indirect transformations between motor efference copies and their sensory representations via supramarginal gyrus may underlie white matter

ERP vs. Vocal Response Correlation

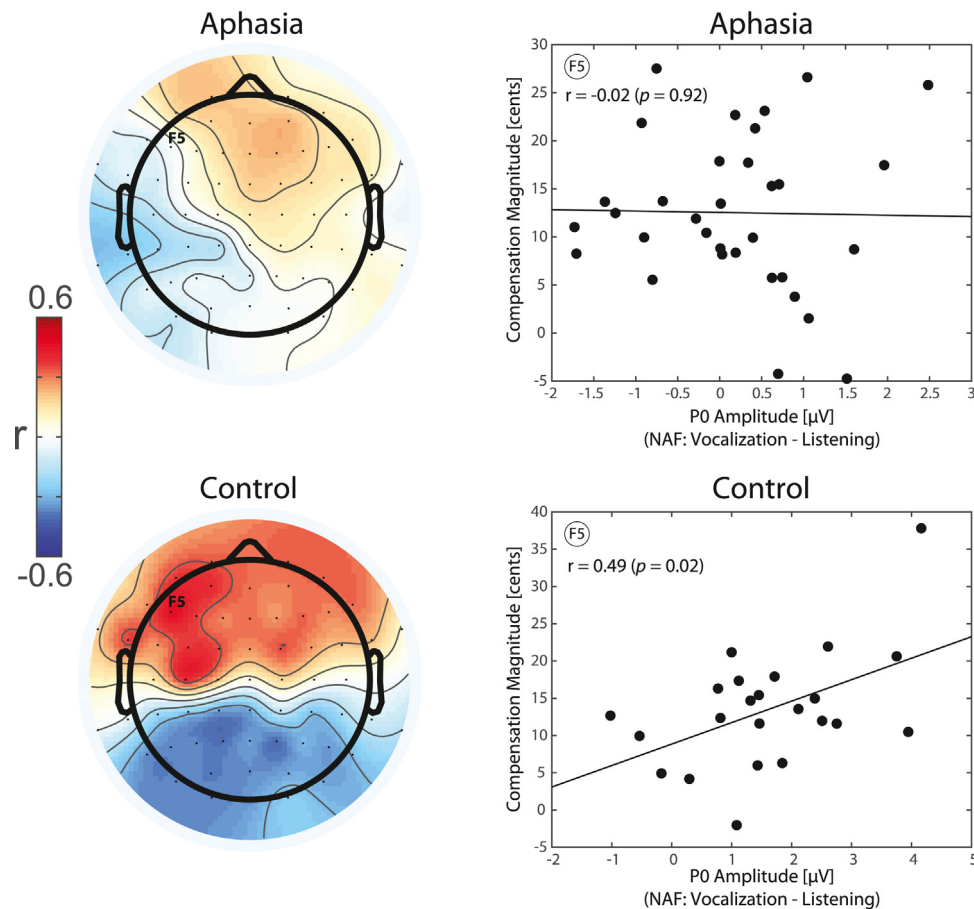


Fig. 6. Topographical distribution maps and raster plots of correlation between vocalization-induced modulations of ERP activity for the representative P0 component elicited before the onset of vowel production and the magnitude of compensatory vocal responses to pitch-shift AAF stimuli in aphasia and control groups.

tracts in the posterior segment of the arcuate fasciculus that connect inferior parietal regions to posterior superior temporal cortex (Dick et al., 2014; Hula et al., 2020). Alternatively, a pathway within the long segment of the arcuate fasciculus that connects the inferior frontal cortex to posterior temporal lobe may also involve direct transformations between motor efference copies and their sensory representations in the auditory cortex (see the solid arrows in Fig. 5A). Despite the existing evidence in support of the involvement of the arcuate fasciculus tracts in regulating sensorimotor integration within the dorsal stream network, it should be noted that its frontal connectivity pattern has remained largely controversial. In an alternative view, one study (Martino et al., 2013) has shown that the frontal termination points of the arcuate fasciculus are predominantly toward the posterior inferior frontal gyrus areas that do not include *pars orbitalis*. In this context, it is reasonable to suggest that *pars orbitalis* may also be connected to the sensorimotor networks via white matter tracts including the uncinate and the inferior fronto-occipital fasciculus.

For responses to altered auditory feedback, lesions within a distinguishable, and yet overlapping, left-hemispheric network predicted temporal-specific patterns of neural deficits in aphasia as indexed by the diminished vocalization-induced enhancement of ERP responses to pitch-shift stimuli. We found that the aberrant modulation of short-latency P1 component at 50 – 100 ms was predicted by damage to the *pars orbitalis* and *pars opercularis* within the inferior frontal gyrus as well as the angular gyrus within the inferior parietal cortex, whereas neural deficits for the long-latency N1 component at 100 – 150 ms were pre-

dicted by lesions to the angular and supramarginal gyri within the inferior parietal cortical areas. These findings identified the inferior frontal *pars opercularis* and angular gyri as lesion predictors of neural deficits in aphasia only in response to altered, but not normal, auditory feedback, suggesting that they are specifically involved in regulating vocal control under conditions where auditory feedback is altered by an external stimulus. Considering the temporal profile of compensatory vocal responses (see Fig. 4), we argue that these anatomical areas support different functional mechanisms for vocal feedback control. The *pars opercularis* predicted neural deficits in the P1 ERP responses that were elicited in the time window before the onset of vocal compensation (i.e. < 100 ms), suggesting that this area is primarily implicated in neural mechanisms of vocal feedback error detection rather than motor correction. However, the angular gyrus was identified as a lesion predictor of neural deficits in ERP components before (i.e. P1) and after (i.e. N1) the onset of vocal compensation, suggesting that this area supports neural processes that are activated during vocal error detection and motor correction. This latter notion is supported by our data showing that the magnitude of vocal compensations were only correlated with the long-latency N1 ERPs in control subjects and this effect was degraded in individuals with post-stroke aphasia. Although the functional role of the angular gyrus is not clearly understood, evidence from a previous study has suggested that this area interacts with networks within the intraparietal sulcus areas (e.g., supramarginal gyrus) that are activated during sensorimotor tasks (Buchsbbaum et al., 2011), and may help regulate attentional resources for disengaging from a current state and switching to a new plan.

Taken together, our findings suggest that the computational neural processes of vocal feedback control are mediated via two anatomically segregated and functionally distinct networks in the left hemisphere: First, a dorsal stream network mediated via the arcuate fasciculus that compares incoming feedback from posterior superior temporal gyrus with forward predictions maintained within the inferior frontal and supramarginal gyri, giving rise to error signals that drive compensatory vocal behavior in response to external alterations in the auditory feedback (see the solid arrows in Fig. 5B); and second, a ventral stream network mediated via the extreme capsule and/or uncinata fasciculus that compares incoming feedback from anterior superior temporal gyrus with forward predictions maintained in the frontal cortex, giving rise to error signals in the *pars opercularis* gyrus as a result of mismatching signals (see the dashed arrows in Fig. 5B). In this context, we argue that while the dorsal networks are involved in error detection to drive compensatory vocal behavior, the ventral network may provide an interface to process errors for updating the neural representations of the internal forward model in the inferior frontal cortex. Although the notion of a ventral stream network for sensorimotor integration is not elaborately discussed in the dual-stream model (Hickok and Poeppel, 2007, 2004), evidence from other studies has provided support for the existence of such pathways for vocal production in humans and non-human primates (Friederici et al., 2006; Rauschecker, 2011; Remedios et al., 2009).

Despite the critical role of left-hemispheric brain networks in regulating vocal production and motor control, previous findings have indicated that areas within the ventral cortical motor and pre-motor regions of the right hemisphere show increased activation in response to altered speech feedback, suggesting that these brain networks may also contribute to auditory control mechanisms during speech production (Tourville et al., 2008). This effect raises the question whether the engagement of right-hemispheric brain networks can compensate for impaired vocal feedback control in left-hemisphere stroke survivors with aphasia. In the present study, we did not find evidence in support of this notion as the results of our ERP analysis did not reveal a significant group effect on neural responses elicited in scalp electrodes over the right hemisphere. This finding was further confirmed by the results of our correlation analysis indicating no significant relationship between the behavioral measures of vocal compensation and ERP responses in electrodes over the right hemisphere. One possible explanation for this difference is that subjects in Tourville et al.'s study (Tourville et al., 2008) were recruited from younger adult populations (23 – 36 yrs.) whereas subjects of both groups in the present study were older adults who may have diminished functional neural capacities to engage compensatory right-hemispheric mechanisms for vocal feedback control. In addition, since the anatomical location of scalp-recorded ERP neural generators cannot be determined without conducting source estimation analysis, further studies are warranted to more reliably investigate the contribution of right-hemispheric mechanisms to vocal auditory feedback control.

A potential limitation of the present study was that the distribution of lesion maps across the aphasia group did not include certain anatomical areas involved in vocal sensorimotor integration (e.g., supplementary motor and anterior cingulate cortex). This limitation was due to the fact that stroke-induced damage to left-hemisphere brain networks did not affect those regions in some subjects or the overlap in those areas did not reach the threshold for ROI-based lesion mapping analysis. Therefore, statistical discrimination of all anatomical structures involved in vocal sensorimotor control was not possible in this study. This issue is a common effect to lesion studies in stroke survivors, and its constraints are dictated by cerebrovascular anatomy and co-occurrence of frequent lesions within similar vascular perfusion bed across individual subjects. This issue is further compounded by the choice of stringent statistical thresholds to control for family-wise error, limiting our statistical power in regions with low injury incidence. In addition, these challenges are aggravated by the lack of advanced diffusion-based imaging data and the

lack of unified accounts to define speech and language deficits in individuals with post-stroke aphasia. Therefore, recording of functional neuroimaging data such as magnetic resonance imaging (fMRI) and near-infrared spectroscopy (fNIRS) can provide solutions to tackle these limitations in future studies. Moreover, due to the cross-sectional design, relatively small sample size, and heterogeneous distribution of aphasia sub-types in the stroke group recruited for this study, characterizing the differences in the behavioral and neural correlates of sensorimotor deficits was not feasible within each aphasia category. This limitation should be addressed by larger sample size studies in the future. Lastly, subjects in the present study were not screened for conditions associated with impaired pitch processing due to the development of amusia and/or aprosodia following stroke, and therefore, future investigations are warranted to examine their effects on the vocal feedback control mechanisms.

5. Conclusions

In conclusion, it seems evident that damage to left-hemisphere brain networks can have negative impacts on the neural and behavioral correlates of vocal production and sensorimotor control in aphasia. Data in the present study provided supporting evidence for the notion that such deficits are predominantly accounted for by damage to the dorsal stream networks that are implicated in generating and/or maintaining feedback control mechanisms for vocal error detection and motor correction. Future studies are warranted to determine how such deficits may vary across different aphasia sub-types and production modalities (e.g., vocal production vs. continuous speech). Studies on sensorimotor control mechanisms of vocalization (Behroozmand et al., 2016; Burnett et al., 1998; Chang et al., 2013; Chen et al., 2007; Larson, 1998) and speech (Cai et al., 2011; Niziolek and Guenther, 2013; Tourville et al., 2008) using altered auditory feedback have shown strong parallels in the automaticity and time course of compensatory behavior, suggesting that the underlying mechanisms of laryngeal and supra-laryngeal (i.e. vocal tract) control may use similar neuro-computations and homologous if not shared neural circuitries. This in turn suggests that studying the neural mechanisms of vocal sensorimotor control may provide an ideal model to understand how speech sensorimotor networks are organized in the brain, how they may break down in neurological conditions leading to speech disorders, and how they can be treated. These findings motivate interventions targeted toward normalizing functional neural activities via techniques such as transcranial electric stimulation (tES), transcranial magnetic stimulation (TMS), and neurofeedback training for vocal and speech rehabilitation in stroke-induced aphasia and other neurological populations.

CRedit author statements

R.B. and J.F. designed the research, and R.B. supervised data collection for the experiments. R.B. analyzed the collected data with assistance from J.F., L.B., and C.R. R.B. wrote the paper and L.B., C.R., G.H. and J.F. provided feedback. All authors reviewed and approved the final draft.

Data availability

The data and code used in this study are available from the corresponding author upon reasonable request and filing for a formal data sharing agreement.

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgement

This research was supported by funding from NIH/NIDCD Grants R01-DC018523 and K01-DC015831 (PI: Behroozmand) and R21-

DC014170 and P50-DC014664 (PI: Fridriksson). The authors wish to thank Lorelei Phillip Johnson and Stacey Sangtian for their assistance with data collection.

References

- Behroozmand, R., Karvelis, L., Liu, H., Larson, C.R., 2009. Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clin. Neurophysiol.* 120, 1303–1312. S1388-2457(09)00340-X [pii]. doi:10.1016/j.clinph.2009.04.022.
- Behroozmand, R., Larson, C.R., 2011. Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neurosci.* 12, 54. doi:10.1186/1471-2202-12-54.
- Behroozmand, R., Oya, H., Nourski, K., Kawasaki, H., Larson, C.R., Brugge, J.F., Howard III, M.A., Greenlee, J.D.W., 2016. Neural correlates of vocal production and motor control in human Heschl's gyrus. *J. Neurosci.* 36, 2302–2315.
- Behroozmand, R., Phillip, L., Johari, K., Bonilha, L., Rorden, C., Hickok, G., Fridriksson, J., 2018. Sensorimotor impairment of speech auditory feedback processing in aphasia. *Neuroimage* 165, 102–111. doi:10.1016/j.neuroimage.2017.10.014.
- Boersma, P., Weenik, D., 1996. PRAAT: a system for doing phonetics by computer. *Rep. Inst. Phonetic Sci. Univ. Amsterdam*.
- Buchsbaum, B.R., Baldo, J., Okada, K., Berman, K.F., Dronkers, N., D'Esposito, M., Hickok, G., 2011. Conduction aphasia, sensory-motor integration, and phonological short-term memory - An aggregate analysis of lesion and fMRI data. *Brain Lang* 119, 119–128. doi:10.1016/j.bandl.2010.12.001.
- Burnett, T.a, Freedland, M.B., Larson, C.R., Hain, T.C., 1998. Voice F0 responses to manipulations in pitch feedback. *J. Acoust. Soc. Am.* 103, 3153–3161. doi:10.1121/1.423073.
- Cai, S., Ghosh, S.S., Guenther, F.H., Perkell, J.S., 2011. Focal manipulations of formant trajectories reveal a role of auditory feedback in the online control of both within-syllable and between-syllable speech timing. *J. Neurosci.* 31, 16483–16490. doi:10.1523/JNEUROSCI.3653-11.2011.
- Chang, E.F., Niziolek, C.a, Knight, R.T., Nagarajan, S.S., Houde, J.F., 2013. Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc. Natl. Acad. Sci. U. S. A.* 110, 2653–2658. doi:10.1073/pnas.1216827110.
- Chen, S.H., Liu, H., Xu, Y., Larson, C.R., 2007. Voice F 0 responses to pitch-shifted voice feedback during English speech. *J. Acoust. Soc. Am.* 121, 1157–1163. doi:10.1121/1.2404624.
- Cheng, H.S., Niziolek, C.A., Buchwald, A., McAllister, T., 2021. Examining the relationship between speech perception, production distinctness, and production variability. *Front. Hum. Neurosci.* 15. doi:10.3389/fnhum.2021.660948.
- Cocquyt, E.M., Lanckmans, E., van Mierlo, P., Duyck, W., Szmalec, A., Santens, P., De Letter, M., 2020. The white matter architecture underlying semantic processing: a systematic review. *Neuropsychologia* 136, 107182. doi:10.1016/j.neuropsychologia.2019.107182.
- Conner, C.R., Kadipasaoglu, C.M., Shouval, H.Z., Hickok, G., Tandon, N., 2019. Network dynamics of Broca's area during word selection. *PLoS ONE* 14, 1–30. doi:10.1371/journal.pone.0225756.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi:10.1016/j.jneumeth.2003.10.009.
- Dick, A.S., Bernal, B., Tremblay, P., 2014. The language connectome: new pathways, new concepts. *Neuroscientist* 20, 453–467. doi:10.1177/1073858413513502.
- Eliades, S.J., Wang, X., 2008. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* 453, 1102–1106. doi:10.1038/nature06910.
- Eliades, S.J., Wang, X., 2003. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J. Neurophysiol.* 89, 2194–2207. doi:10.1152/jn.00627.2002.
- Fischer, J., 2017. Primate vocal production and the riddle of language evolution. *Psychon. Bull. Rev.* 24, 72–78. doi:10.3758/s13423-016-1076-8.
- Fitch, W.T., 2017. Empirical approaches to the study of language evolution. *Psychon. Bull. Rev.* 24, 3–33. doi:10.3758/s13423-017-1236-5.
- Fitch, W.T., 2010. *The Evolution of Language*. Cambridge University Press., Cambridge, UK.
- Fridriksson, J., Guo, D., Fillmore, P., Holland, A., Rorden, C., 2013. Damage to the anterior arcuate fasciculus predicts non-fluent speech production in aphasia. *Brain* 136, 3451–3460. doi:10.1093/brain/awt267.
- Fridriksson, J., Yourganov, G., Bonilha, L., Basilakos, A., Ouden, D.D., 2016. Revealing the dual streams of speech processing. *PNAS* 113, 15108–15113. doi:10.1073/pnas.1614038114.
- Friederici, A.D., Bahlmann, J., Heim, S., Schubotz, R.I., Anwander, A., 2006. The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proc. Natl. Acad. Sci. U. S. A.* 103, 2458–2463. doi:10.1073/pnas.0509389103.
- Greenlee, J.D.W., Behroozmand, R., Larson, C.R., Jackson, A.W., Chen, F., Hansen, D.R., Oya, H., Kawasaki, H., Howard, M.a, 2013. Sensory-motor interactions for vocal pitch monitoring in non-primary human auditory cortex. *PLoS ONE* 8. doi:10.1371/journal.pone.0060783.
- Hickok, G., 2017. A cortical circuit for voluntary laryngeal control: implications for the evolution language. *Psychon. Bull. Rev.* 24, 56–63. doi:10.3758/s13423-016-1100-z.
- Hickok, G., 2012. The cortical organization of speech processing : feedback control and predictive coding the context of a dual-stream model. *J. Commun. Disord.* 45, 393–402. doi:10.1016/j.jcomdis.2012.06.004.
- Hickok, G., Buchsbaum, B., Humphries, C., Muftuler, T., 2003. Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area spt. *J. Cogn. Neurosci.* 15, 673–682. doi:10.1162/089892903322307393.
- Hickok, G., Houde, J., Rong, F., 2011. Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69, 407–422. doi:10.1016/j.neuron.2011.01.019.
- Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–403.
- Hickok, G., Poeppel, D., 2004. Dorsal and ventral streams : a framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99. doi:10.1016/j.cognition.2003.10.011.
- Hickok, G., Okada, K., Serences, J.T., 2009. Area spt in the human planum temporale supports sensory-motor integration for speech processing. *J. Neurophysiol.* 101, 2725–2732. doi:10.1152/jn.91099.2008.
- Hickok, G., Poeppel, D., 2000. Towards a functional neuroanatomy of speech perception. *Trends Cogn. Sci.* 4, 131–138. doi:10.1016/S1364-6613(00)01463-7.
- Houde, J.F., Nagarajan, S.S., Sekihara, K., Merzenich, M.M., 2002. Modulation of the auditory cortex during speech: an MEG study. *J. Cogn. Neurosci.* 14, 1125–1138. doi:10.1162/089892902760807140.
- Hula, W.D., Panesar, S., Gravier, M.L., Yeh, F.C., Dresang, H.C., Dickey, M.W., Fernandez-Miranda, J.C., 2020. Structural white matter connectometry of word production in aphasia: an observational study. *Brain* 143, 2532–2544. doi:10.1093/brain/awaa193.
- Kertesz, A., 2007. *The Western Aphasia Battery - Revised*. Grune & Stratton, New York.
- Kertesz, A., 1982. *Western Aphasia Battery*. Grune & Stratton, New York.
- Kuypers, H.G., 1958a. Corticobular connections to the pons and lower brain-stem in man: an anatomical study. *Brain* 81, 364–388.
- Kuypers, H.G., 1958b. Some projections from the peri-central cortex to the pons and lower brain stem in monkey and chimpanzee. *J. Comp. Neur.* 110, 221–255.
- Larson, C.R., 1998. Cross-modality influences in speech motor control: the use of pitch shifting for the study of F0 control. *J. Commun. Disord.* 31, 489–502 quiz 502–3; 553.
- Li, X., Morgan, P.S., Ashburner, J., Smith, J., Rorden, C., 2016. The first step for neuroimaging data analysis: DICOM to NIFTI conversion. *J. Neurosci. Methods* 264, 47–56. doi:10.1016/j.jneumeth.2016.03.001.
- Martino, J., De Witt Hamer, P.C., Berger, M.S., Lawton, M.T., Arnold, C.M., De Lucas, E.M., Duffau, H., 2013. Analysis of the subcomponents and cortical terminations of the perisylvian superior longitudinal fasciculus: a fiber dissection and DTI tractography study. *Brain Struct. Funct.* 218, 105–121. doi:10.1007/s00429-012-0386-5.
- Nachev, P., Coulthard, E., Jäger, H.R., Kennard, C., Husain, M., 2008. Enantiomorphic normalization of focally lesioned brains. *Neuroimage* 39, 1215–1226. doi:10.1016/j.neuroimage.2007.10.002.
- Niziolek, C.a, Guenther, F.H., 2013. Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations. *J. Neurosci.* 33, 12090–12098. doi:10.1523/JNEUROSCI.1008-13.2013.
- Niziolek, C.a, Nagarajan, S.S., Houde, J.F., 2013. What does motor efference copy represent? Evidence from speech production. *J. Neurosci.* 33, 16110–16116. doi:10.1523/JNEUROSCI.2137-13.2013.
- Pichon, S., Kell, C.A., 2013. Affective and sensorimotor components of emotional prosody generation. *J. Neurosci.* 33, 1640–1650. doi:10.1523/JNEUROSCI.3530-12.2013.
- Poeppel, D., Hickok, G., 2004. Towards a new functional anatomy of language. *Cognition* 92, 1–12. doi:10.1016/j.cognition.2003.11.001.
- Rauschecker, J.P., 2011. An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hear. Res.* 271, 16–25. doi:10.1016/j.heares.2010.09.001.
- Remedios, R., Logothetis, N.K., Kayser, C., 2009. An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. *J. Neurosci.* 29, 1034–1045. doi:10.1523/JNEUROSCI.4089-08.2009.
- Rogalsky, C., Poppa, T., Chen, K., Anderson, S.W., Damasio, H., Love, T., Hickok, G., 2015. Speech repetition as a window on the neurobiology of auditory – motor integration for speech : a voxel-based lesion symptom mapping study. *Neuropsychologia* 71, 18–27. doi:10.1016/j.neuropsychologia.2015.03.012.
- Rorden, C., Bonilha, L., Fridriksson, J., Bender, B., Karnath, H., 2012. Age-specific CT and MRI templates for spatial normalization. *Neuroimage* 61, 957–965. doi:10.1016/j.neuroimage.2012.03.020.
- Strand, E.A., Duffy, J.R., Clark, H.M., Josephs, K., 2014. The apraxia of speech rating scale: a tool for diagnosis and description of apraxia of speech. *J. Commun. Disord.* 51, 43–50. doi:10.1016/j.jcomdis.2014.06.008.
- Tang, C., Hamilton, L.S., Chang, E.F., 2017. Intonational speech prosody encoding in the human auditory cortex. *Science* 357 (80), 797–801. doi:10.1126/science.aam8577.
- Templeton, G.F., 2011. A two-step approach for transforming continuous variables to normal: implications and recommendations for IS research. *Commun. Assoc. Inf. Syst.* 28, 41–58. doi:10.17705/1cais.02804.
- Tourville, J.a, Reilly, K.J., Guenther, F.H., 2008. Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 39, 1429–1443. doi:10.1016/j.neuroimage.2007.09.054.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated Anatomical Labeling of Activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi:10.1006/nimg.2001.0978.
- Wagner, M., Tech, R., Fuchs, M., Kastner, J., Gasca, F., 2017. Statistical non-parametric mapping in sensor space. *Biomed. Eng. Lett.* 7, 193–203. doi:10.1007/s13534-017-0015-6.