Cortical network underlying speech production during delayed auditory feedback

Muge Ozker[1], Werner Doyle[2], Orrin Devinsky[1], Adeen Flinker[1]

1. Department of Neurology, New York University Langone Medical Center, NY, USA
2. Department of Neurosurgery, New York University Langone Medical Center, NY, USA

Abstract

Accurate and fluent production of speech strongly depends on hearing oneself which allows for the detection and correction of vocalization errors in real-time. When auditory feedback is disrupted with a time delay (e.g. echo on a conference call), it causes slowed and stutter-like speech in humans. Impaired speech motor control during delayed auditory feedback is implicated in various neurological disorders ranging from stuttering to aphasia, however the underlying neural mechanisms are poorly understood. Here, we investigated auditory feedback control in human speech by obtaining electrocorticographic recordings from neurosurgical subjects performing a delayed auditory feedback (DAF) task. We observed a significant increase in neural activity in auditory sites that scaled with the duration of feedback delay and correlated with response suppression during normal speech, providing direct evidence for a shared mechanism between sensitivity to altered feedback and speech-induced auditory suppression in humans. Furthermore, we find that when subjects robustly slowed down their speech rate to compensate for the delay, the dorsal division of the precentral gyrus was preferentially recruited to support articulation during an early time frame. This recruitment was accompanied by response enhancement across a large speech network commencing in temporal cortex and then engaging frontal and parietal sites. Our results highlight the critical components of the human speech network that support auditory feedback control of speech production and the temporal evolution of their recruitment.

Abbreviations: ECoG = Electrocorticography; DAF = Delayed auditory feedback; STG = Superior temporal gyrus; IFG = Inferior frontal gyrus; NMF = Non-negative matrix factorization; DTW = Dynamic time warping

Introduction

Human speech production is strongly influenced by the auditory feedback it generates. When we speak, we continuously monitor our vocal output and adjust our vocalization to maintain fluency. For example, speakers involuntarily raise their voice in the presence of background noise to be more audible (Lombard 1911, Zollinger and Brumm 2011). Similarly, when speakers hear themselves with a delay (e.g. voice delays or echoes in teleconferencing), they compensate for the delay by slowing down and resetting their speech. This compensatory adjustment of human vocalization provides evidence for a mechanism which detects and corrects vocal errors in real time. Abnormal sensorimotor network interactions have been implicated in various disorders including stuttering, aphasia, Parkinson's disease, autism spectrum disorder and schizophrenia (Goldberg, Gold et al. 1997, Civier, Tasko et al. 2010, Liu, Wang et al. 2012, Lin, Mochida et al. 2015, Hardy, Bond et al. 2018), however the neural underpinnings of this dynamic system remain poorly understood.

Sensorimotor processing shares a common mechanism known as corollary discharge across the animal kingdom including crickets, songbirds, mice and primates (Crapse and Sommer 2008, Schneider and Mooney 2018). Corollary discharge suppresses sensory responses to self-generated motor actions and renders the brain more sensitive to detecting external sensory stimuli as well as motor errors. In human speech production, this mechanism is described in models which suggest that neural response in the auditory cortex is suppressed during speech production. Furthermore, when there is a mismatch between intended speech and its perceived auditory feedback, the auditory response is enhanced to encode the mismatch. This auditory error signal is then relayed to vocal-motor regions for the real time correction of vocalization in order to produce the intended speech (Hickok, Houde et al. 2011, Houde and Nagarajan 2011,

Tourville and Guenther 2011). In support of these models, electrophysiological studies in non-human primates demonstrated increased activity in auditory neurons when the frequency of the auditory feedback is shifted during vocalization (Eliades and Wang 2008). Behavioral evidence in human studies showed that when formant frequencies of a vowel or the fundamental frequency (pitch) is shifted, speakers change their vocal output in the opposite direction of the shift to compensate for the spectral perturbation (Houde and Jordan 1998, Jones and Munhall 2000, Niziolek and Guenther 2013). In line with non-human primate studies, human neurosurgical recordings as well as neuroimaging studies demonstrated that these feedback-induced vocal adjustments are accompanied by enhanced neural responses in auditory regions (Tourville, Reilly et al. 2008, Behroozmand, Karvelis et al. 2009, Behroozmand, Shebek et al. 2015). However, all these studies manipulate auditory feedback by altering the spectral features of vocalizations. A more ecologically relevant manipulation in humans is altering the temporal features of auditory feedback by delaying the voice onset in real time, termed "delayed auditory feedback (DAF)".

First described in the 1950s, DAF strongly disrupts speech fluency leading to slower speech rate, pauses, syllable repetitions and increased voice pitch or intensity (Lee 1950, Black 1951, Fairbanks 1955). Further, higher susceptibility to DAF occurs in autism spectrum disorder, non-fluent primary progressive aphasia, schizophrenia and other neurological disorders (Goldberg, Gold et al. 1997, Lin, Mochida et al. 2015, Hardy, Bond et al. 2018). Interestingly, DAF improves speech fluency in individuals who stutter and is a therapeutic approach in speech therapy for stuttering and Parkinson's Disease (Bloodstein 1969, Kalinowski, Stuart et al. 1996, Blanchet and Hoffman 2014). Although behavioral effects of DAF have been widely studied in both normal and clinical groups, only a few neuroimaging studies investigated the neural

responses. These studies demonstrated enhanced responses in bilateral posterior superior cortices during DAF compared with normal auditory feedback conditions (Hirano, Kojima et al. 1997, Hashimoto and Sakai 2003, Takaso, Eisner et al. 2010). However, the temporal aspects of sensorimotor integration during speech production with DAF remain unknown.

To address this issue, we leveraged the excellent spatial and temporal resolution of electrocorticography (ECoG). Using ECoG, we acquired direct cortical recordings from 15 epilepsy patients while they read aloud words and sentences. As they spoke, we recorded their voice and played it back to them through earphones either simultaneously or with a delay (50, 100 and 200 milliseconds). We found that subjects slowed down their speech to compensate for the delay, and more profoundly so when producing sentences. There was a significant increase in neural activity across a large speech network encompassing temporal, parietal and frontal sites that scaled with the duration of feedback delay. In auditory sites, response enhancement during DAF correlated with speech-induced auditory suppression during normal speech. Critically, when speech was slowed down and was effortful, the dorsal division of the precentral gyrus was preferentially recruited at an early timing to support articulation. To our knowledge, we introduce the first DAF investigation with invasive human electrophysiology in which we reveal the spatiotemporal dynamics of the neural mechanism underlying compensatory adjustment of human vocalization.

Materials and methods

Subject information

All experimental procedures were approved by the New York University School of Medicine Institutional Review Board. 15 neurosurgical epilepsy patients (8 females, mean age: 34, 2 right,

9 left and 4 bilateral hemisphere coverage) implanted with subdural and depth electrodes provided informed consent to participate in the research protocol. Electrode implantation and location were guided solely by clinical requirements. 3 patients were consented separately for higher density clinical grid implantation, which provided denser sampling of underlying cortex.

Experiment setup

Subjects were tested while resting in their hospital bed in the epilepsy-monitoring unit. Visual stimuli were presented on a laptop screen positioned at a comfortable distance from the subject. Auditory stimuli were presented through earphones (Bed Phones On-Ear Sleep Headphones Generation 3) in the Delayed Auditory Feedback experiment and through speakers in the Auditory Repetition experiment. Subjects' voice was recorded using an external microphone (Zoom H1 Handy Recorder).

Delayed auditory feedback experiment

The experiment consisted of a word-reading session and a sentence-reading session. 10 different 3-syllable words (e.g. document) were used in the word reading session, and 6 different 8-word sentences (e.g. The cereal was fortified with vitamins and nutrients) were used in the sentence reading session. Text stimuli were visually presented on the screen and subjects were instructed to read them out loud. As subjects spoke, their voices were recorded using the laptop's internal microphone, delayed at 4 different amounts (no delay, 50, 100, 200ms) using custom script (MATLAB, Psychtoolbox-3) and played back to them through earphones. A TTL pulse marking the onset of a stimulus, the delayed feedback voice signal (what the patient heard) and the actual microphone signal (what the patient spoke) were fed in to the EEG amplifier as an auxiliary

input in order to acquire them in sync with the EEG samples. Trials, which consisted of different stimulus-delay combinations, were presented randomly (3 to 8 repetitions) with a 1 second inter-trial-interval.

Auditory repetition experiment

Stimuli consisted of 50 different auditory words presented randomly (2 repetitions) through speakers. Subjects were instructed to listen to the presented words and repeat them out loud at each trial. Both the microphone signal and a TTL pulse marking the onset of a stimulus were fed in to the EEG amplifier as an auxiliary input in order to acquire them in sync with EEG samples.

Electrocorticography (ECoG) recording

ECoG was recorded from implanted subdural platinum-iridium electrodes embedded in flexible silicon sheets (2.3 mm diameter exposed surface, 8 x 8 grid arrays and 4 to 12 contact linear strips, 10 mm center-to-center spacing, Ad-Tech Medical Instrument, Racine, WI) and penetrating depth electrodes (1.1 mm diameter, 5-10 mm center-to-center spacing 1 x 8 or 1 x 12 contacts, Ad-Tech Medical Instrument, Racine, WI). Three subjects consented to a research hybrid grid implanted which included 64 additional electrodes between the standard clinical contacts (16 × 8 grid with sixty-four 2 mm macro contacts at 8 x 8 orientation and sixty-four 1 mm micro contacts in between, providing 10 mm center-to-center spacing between macro contacts and 5 mm center-to-center spacing between micro/macro contacts, PMT corporation, Chanassen, MN). Recordings were made using one of two amplifier types: NicoletOne amplifier (Natus Neurologics, Middleton, WI), bandpass filtered from 0.16-250 Hz and digitized at 512 Hz. Neuroworks Quantum Amplifier (Natus Biomedical, Appleton, WI) recorded at 2048

Hz, bandpass filtered at 0.01682.67 Hz and then downsampled to 512 Hz. A two-contact subdural strip facing toward the skull near the craniotomy site was used as a reference for recording and a similar two-contact strip screwed to the skull was used for the instrument ground. Electrocorticography and experimental signals (trigger pulses that mark the appearance of visual stimuli on the screen, microphone signal from speech recordings and auditory playback signal that was heard by the patients through earphones) were acquired simultaneously by the EEG amplifier in order to provide a fully synchronized dataset. For the DAF task, recorded microphone and feedback signals were analyzed to ensure that the temporal delay manipulation by our MATLAB code produced the intended delay.

Electrode localization

Electrode localization in subject space as well as MNI space was based on co-registering a preoperative (no electrodes) and postoperative (with electrodes) structural MRI (in some cases a postoperative CT was employed depending on clinical requirements) using a rigid-body transformation. Electrodes were then projected to the surface of cortex (preoperative segmented surface) to correct for edema induced shifts following previous procedures (Yang, Wang et al. 2012) (registration to MNI space was based on a non-linear DARTEL algorithm (Ashburner 2007). Within subject anatomical locations of electrodes was based on the automated FreeSurfer segmentation of the subject's pre-operative MRI. All middle and caudal superior temporal gyrus electrodes were grouped as superior temporal gyrus (STG), all parsopercularis and pars triangularis electrodes were grouped as inferior frontal gyrus (IFG) electrodes. Precentral electrodes with an x coordinate larger than ±40 were grouped as dorsal precentral, and those with smaller than ±40 were grouped as ventral precentral.

Neural data analysis

A common average reference was calculated by subtracting the average signal across all electrodes from each individual electrode's signal (after rejection of electrodes with artifacts caused by line noise, poor contact with cortex and high amplitude shifts). Continuous data streams from each channel were epoched into trials (from -1.5 s to 3.5 s for word stimuli and from -1.5 s to 5.5 s for sentence stimuli with respect to speech onset). Line noise at 60, 120 and 180 Hz were filtered out and the data was transformed to time-frequency space using the multitaper method (MATLAB, FieldTrip toolbox) with 3 Slepian tapers; frequency windows from 10 to 200 Hz; frequency steps of 5 Hz; time steps of 10 ms; temporal smoothing of 200 ms; frequency smoothing of ±10 Hz. The high gamma broadband response (70-150 Hz) at each time point following stimulus onset was measured as the percent signal change from baseline, with the baseline calculated over all trials in a time window from -500 to -100 ms before stimulus onset. High gamma response duration for each electrode was measured by calculating the time difference at full width quarter maximum of the response curve.

Electrode selection

We recorded from a total of 1693 subdural and 608 depth electrode contacts in 15 subjects. Electrodes were examined for speech related activity defined as significant high gamma broadband responses. For auditory repetition and DAF word-reading tasks, electrodes that showed significant response increase ($p < 10^{-4}$, unpaired t-test) either before (-0.5 to 0 s) or after speech onset (0 to 0.5 s) with respect to a baseline period (-1 to -0.6 s) and at the same time had a large signal-to-noise ratio ($\mu/\sigma > 0.7$) during either of these time windows were selected. For DAF sentence-reading task, the same criteria were applied, except the time window after speech

onset was longer (0 to 3 s). Electrode selection was first performed for each task separately, then electrodes that were commonly selected for both tasks were further analyzed.

Clustering analysis

Non-negative matrix factorization (NMF) was used to identify major response patterns across different brain regions during speech production. NMF is an unsupervised dimensionality reduction (or clustering) technique that reveals the major patterns in the data without specifying any features (Ding, He et al. 2005). We performed the clustering analysis using the data from the word-reading with DAF task, since this data set contained a large number of trials. We combined responses from all subjects by concatenating trials and electrodes forming a large data matrix A (electrodes-by-timepoints). Matrix A was factorized into two matrices W and H by minimizing the root mean square residual between A and W*H (nnmf function in MATLAB). Factorization was performed based on two clusters to represent the two major predicted speech related components in the brain; auditory and motor.

Suppression index calculation

Average response across trials was calculated over two time periods in the Auditory Repetition task. First time period was during listening the stimulus (0-0.5 s) and the second time period was during speaking (0-0.5 s). Suppression index was obtained by comparing the average responses in these two time periods: Listen-Speak/Listen+Speak. For the neural response, raw high gamma broadband signal was used instead of the percent signal change to ensure that the suppression index values varied between -1 to 1, indicating a range from complete enhancement to complete suppression respectively.

Dynamic time warping analysis

For each trial of the DAF sentence-reading task, the speech spectrogram was averaged across frequencies. Then, the mean spectrograms were averaged across trials of the same sentence stimuli (e.g. averaged over Sentence #1 trials). Dynamic time warping (DTW) was performed separately for different sentence stimuli. The first 900 time points of the resulting DTW paths were applied to the neural response signal at each trial (representing approximately 9 seconds in the common warped time). Finally, the transformed neural responses were averaged across trials for each sentence stimuli. This procedure was performed to compare two conditions that resulted in the largest neural response difference (no delay versus 200 ms delay).

Statistical analysis

The effect of DAF on speech behavior was determined by performing one-way ANOVA across subjects using articulation duration of words and sentences as the dependent variable and delay condition as the independent variable. To determine a significant difference in the amplitude of neural response between conditions, the average high gamma activity in a specified time window was compared by performing one-way ANOVA across all trials in all electrodes using delay condition as the independent variable. Similarly, a significant difference in the duration of neural response was determined by performing one-way ANOVA across subjects using response duration as the dependent variable and delay condition as the independent variable. Significance levels were computed at a p-value of 0.01. To compare sensitivity to DAF for word and sentence-reading, sensitivity indices were compared across electrodes using an unpaired t-test. To reveal how response enhancement to DAF changed across time during the sentence-reading task, we performed a one-way ANOVA at each time point using the neural response in each

electrode as a dependent variable and delay as an independent variable. We performed a permutation test at each timepoint to assess a significance threshold for the F-value. We shuffled the delay condition labels 1000 times and performed an ANOVA for each timepoint at each iteration, then we set the threshold the 0.999 quantile of the F-value distribution.

Results

Subjects (N = 15) performed a word-reading task (single 3-syllable words) while their voice onset was delayed (no delay, 50, 100 and 200 ms) and played back to them through earphones in real time, a paradigm known as delayed auditory feedback (DAF). We first analyzed the voice recordings of subjects and measured the articulation duration at different amount of delays to establish the behavioral effect of DAF (**Fig. 1A**). Articulation duration increased slightly with delay: average articulation duration across subjects was 0.698, 0.726, 0.737 and 0.749 milliseconds for no delay, 50, 100 and 200 milliseconds delay conditions respectively. While this increase was not significant (**Fig. 1B, ANOVA:** $F = 1.985$ $p = 0.165$) we observed robust neural changes.

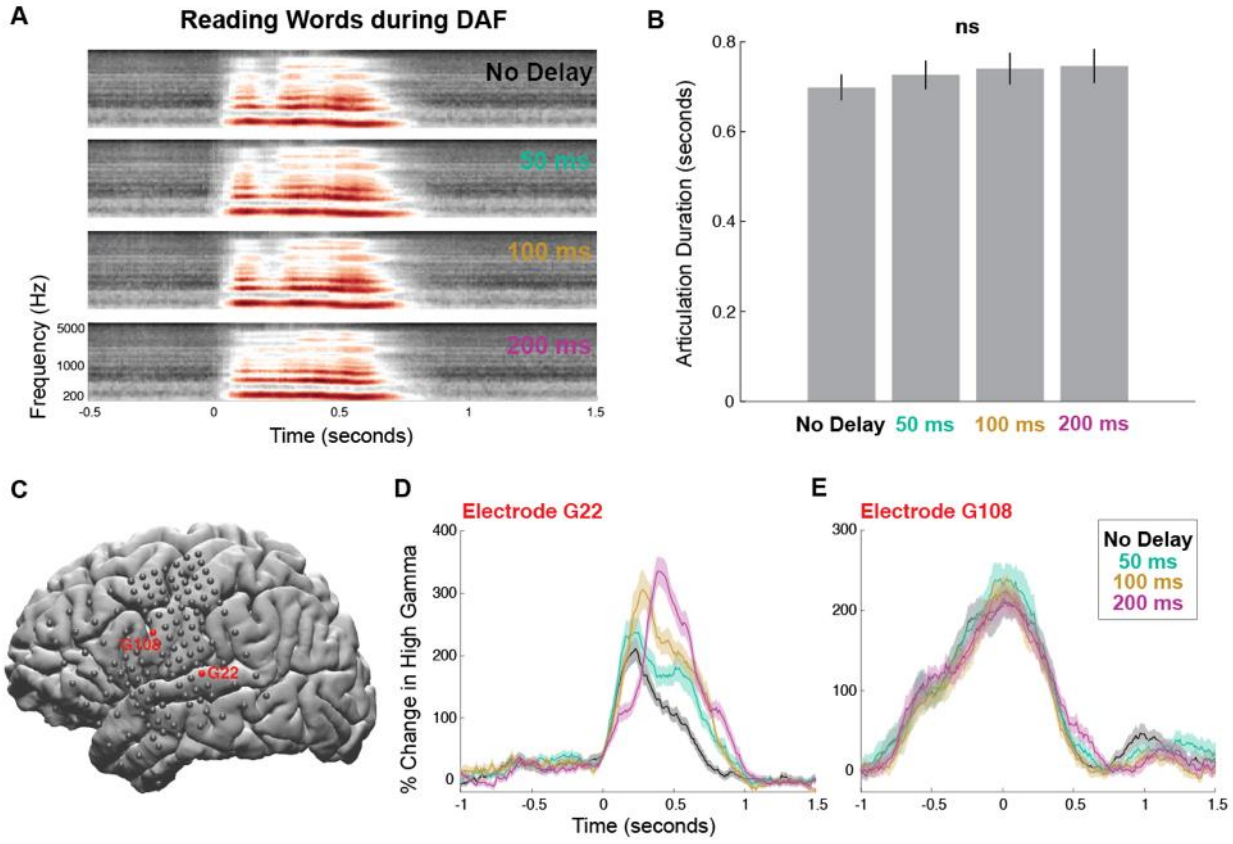

Figure 1 Behavioral and neural responses during word reading with DAF. (A) Speech spectrogram of a single subject articulating words during DAF conditions. **(B)** Mean articulation duration of words during DAF conditions averaged across subjects. Error bars show SEM over subjects. **(C)** Cortical surface model of the left hemisphere brain of a single subject. Gray circles indicate the implanted electrodes. Red highlighted electrodes are located on the STG (G22) and on the precentral gyrus (G108). **(D)** High gamma responses in an auditory electrode (G22) to articulation of words during DAF conditions (color coded). Shaded regions indicate SEM over trials. **(E)** High gamma responses in a motor electrode (G108) to articulation of words during DAF conditions (color coded). Shaded regions indicate SEM over trials. **SEM** = Standard error of the mean

To quantify the neural response, we used the high gamma broadband signal (70-150 Hz, *see Methods*), a widely used index of cortical activity which correlates with underlying neuronal spike rates (Mukamel, Gelbard et al. 2005, Nir, Fisch et al. 2007, Ray and Maunsell 2011). Two response patterns emerged among the electrodes that showed significant activity during speech production (*see Electrode Selection in Methods*). In the first pattern, shown on a representative auditory electrode located in the STG (**Fig. 1C**), neural response started after speech onset and its amplitude increased significantly with delay (**Fig. 1D, ANOVA:** $F = 37$, $p = 1.55\times10^{-8}$). In the second pattern, shown on a representative motor electrode located in ventral precentral gyrus

(**Fig. 1C**), neural response started before speech onset and its amplitude was not affected by delay (**Fig. 1E, ANOVA:** F = 0.084, p = 0.772). This result demonstrated that although DAF did not significantly affect speech behavior (i.e. articulation duration), it affected the neural response in auditory sites that are involved in speech processing.

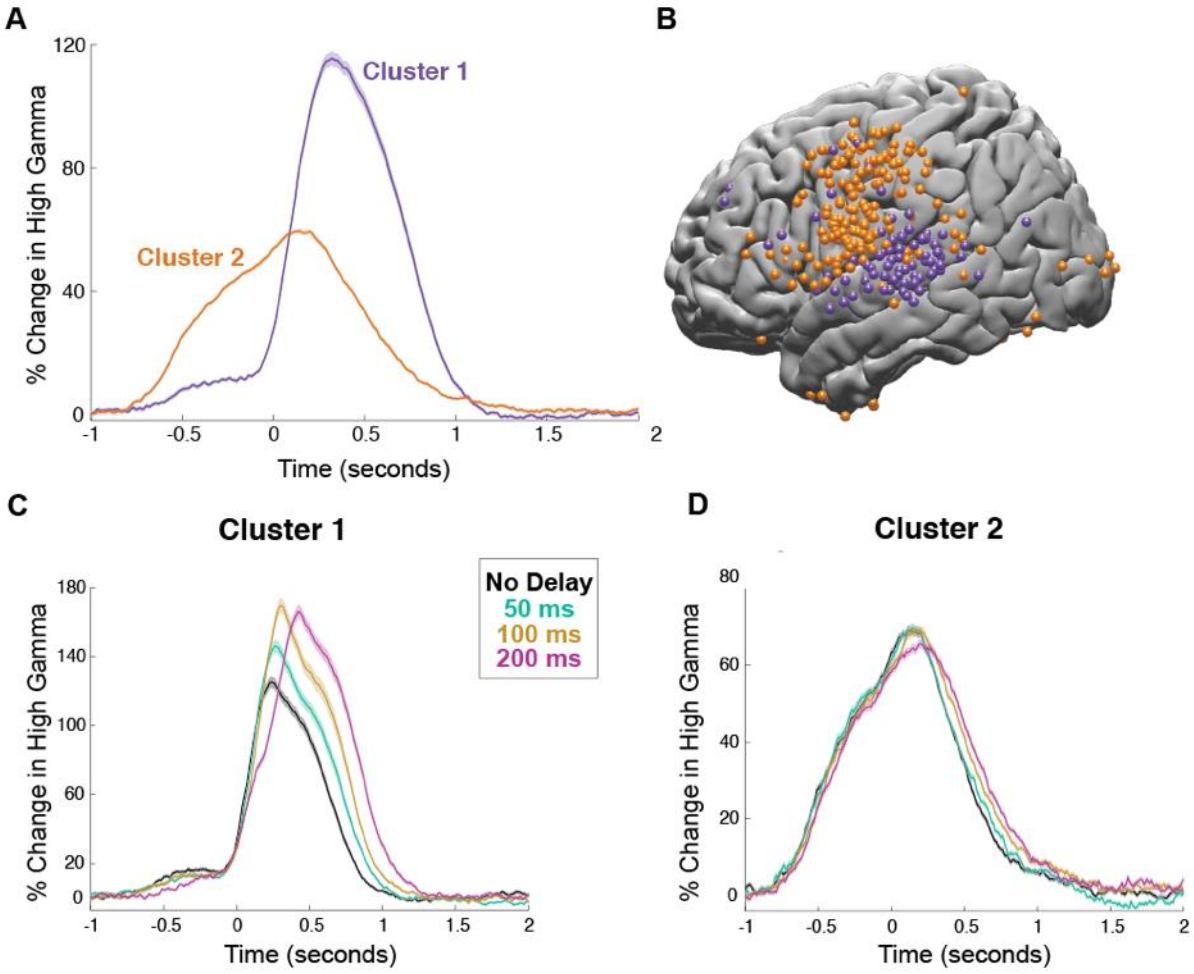

Figure 2 Clustering with non-negative matrix factorization. (A) High gamma responses averaged across electrodes in the two clusters provided by the unsupervised NMF. Shaded regions indicate SEM over trials. **(B)** Spatial distribution on cortex of electrodes in the two clusters displayed on the left hemisphere of a template brain. **(C)** High gamma responses to articulation of words during DAF conditions averaged across electrodes in Cluster 1. Shaded regions indicate SEM over trials. **(D)** High gamma response to articulation of words during DAF averaged across electrodes in Cluster 2. Shaded regions indicate SEM over trials. **SEM** = Standard error of the mean

To characterize the two major response patterns in the brain we chose to use an unbiased, data-driven approach which does not impose any assumptions or restrictions on the selection of

responses. We performed an unsupervised clustering analysis using the NMF algorithm on neural responses across all delay conditions, brain sites and subjects (Ding, He et al. 2005, Hamilton, Edwards et al. 2018). The clustering analysis identified the major response patterns represented by two distinct clusters, which corroborated our representative results shown in a single subject (**Fig. 1C-E**) as well as visual inspection of the data across subjects. The first response pattern (Cluster 1, N = 125 electrodes) started after speech onset and peaked at 320 ms reaching 115 percent change in amplitude. The second response pattern (Cluster 2, N = 253 electrodes) started much earlier approximately 750 ms prior to speech onset and peaked 140 ms after speech onset reaching 60 percent change in amplitude (**Fig. 2A**). These two clusters had a distinct anatomical distribution (**Fig. 2B**): Cluster 1 was mainly localized to STG suggesting an auditory function while Cluster 2 was localized to frontal cortices suggesting a pre-motor and motor function. Next, we examined the effect of DAF on these two clusters. The amplitude of the neural response increased significantly with delay in Cluster 1 (**Fig. 2C, ANOVA:** $F = 5.35$, $p = 0.02$), but not in Cluster 2 (**Fig. 2D, ANOVA:** $F = 1.65$, $p = 0.2$). The duration of the neural response did not show a significant increase in either of the clusters (**ANOVA:** $F = 1$, $p = 0.32$ for Cluster 1 and $F = 0.01$, $p = 0.92$ for Cluster 2).

Our clustering analysis identified two response components that were mostly anatomically distinct reflecting an auditory response to self-generated speech and a motor response to articulation. The auditory component was unique in exhibiting an enhanced response that varied as a function of feedback delay, likely representing an auditory error signal encoding the mismatch between the expected and the actual feedback. We quantified this error signal by calculating a sensitivity index for each electrode in the auditory component (Cluster 1). Sensitivity to DAF was defined as the trial-by-trial correlation between the delay condition and

the neural response averaged over the 0 to 1 s time window. A large sensitivity value indicated a strong response enhancement with increasing delays. DAF sensitivity across auditory electrodes exhibited a large variability ranging from -0.2 to 0.7 (r values).

We hypothesized that sensitivity to DAF should be tightly linked to the degree of speech-induced auditory suppression, as predicted by current models (Houde and Nagarajan 2011,

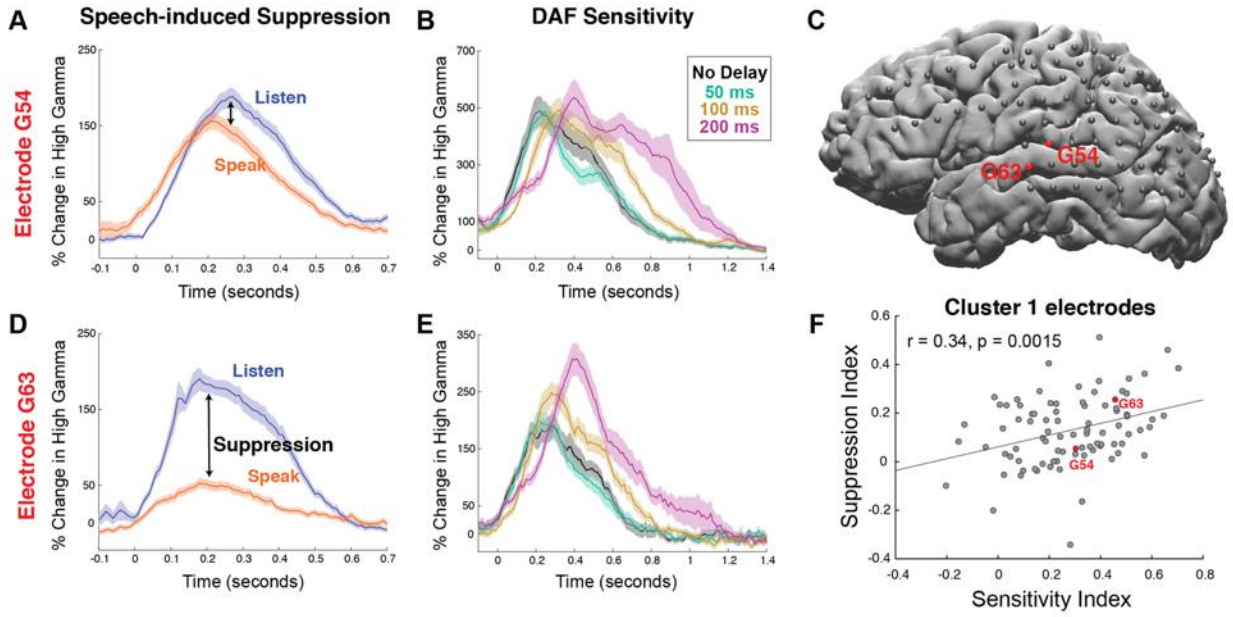

Figure 3 Correlation between auditory suppression and DAF sensitivity. (A) High gamma response in electrode G54 showing a small degree of auditory suppression during speaking words compared to listening. Shaded regions indicate SEM over trials. **(B)** High gamma response in electrode G54 locked to articulation of words during DAF. Shaded regions indicate SEM over trials. **(C)** Cortical surface model of the left hemisphere brain of a single subject. Gray circles indicate the implanted electrodes. Red highlighted electrodes are located on the caudal (G54) and middle STG (G63). **(D)** High gamma responses in electrode G63 showing a large amount of auditory suppression during speaking words compared to listening to the same words. Shaded regions indicate SEM over trials. **(E)** High gamma responses in electrode G63 to articulation of words with DAF. Shaded regions indicate SEM over trials. **(F)** Scatter plot and fitted regression showing a significant correlation between sensitivity to DAF and speech-induced auditory suppression across electrodes in Cluster 1. Each circle represents one electrode's sensitivity and suppression index. **SEM** = Standard error of the mean

Tourville and Guenther 2011). In order to test our hypothesis, we performed an additional auditory repetition experiment. We observed that individual electrodes in a single subject (**Fig. 3C**) that showed weak auditory suppression during speaking (**Fig. 3A**) did not exhibit a strong sensitivity to DAF (**Fig. 3B**), while nearby electrodes that showed strong auditory suppression during speaking (**Fig. 3D**) exhibited a strong sensitivity to DAF (**Fig. 3E**). We quantified a

suppression index for each electrode by comparing the responses during listening versus speaking (SI = Listen-Speak/Listen+Speak; *see Methods*). The degree of suppression across electrodes exhibited a large variability ranging from SI = -0.34 to SI = 0.51. Negative suppression values indicating neural response enhancement to self-generated speech were observed only for a few electrodes. To test the relationship between speech-induced auditory suppression and DAF sensitivity, we calculated the correlation between suppression and sensitivity indices of electrodes in Cluster 1 (**Fig. 3F**). We found a significant correlation between the two measures ($r = 0.34$, $p = 0.0015$) providing evidence for a common neural mechanism. It is noteworthy that while non-human primate studies have demonstrated this neural mechanism in auditory neurons with suppression indices that vary as a function of spectral feedback alteration (Eliades and Wang 2008), human studies have yet to replicate this to date.

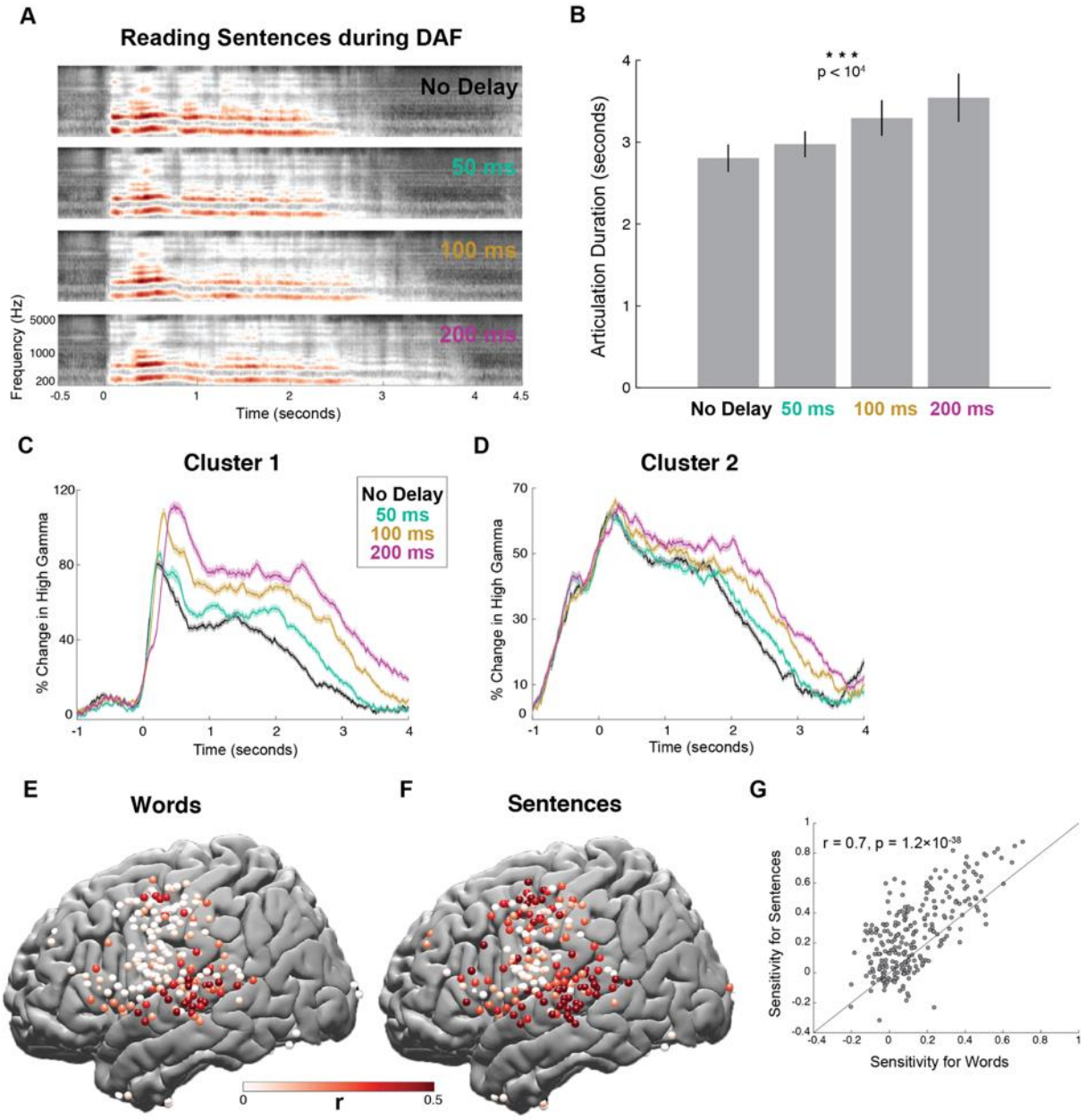

Figure 4 Behavioral and neural responses during sentence reading with DAF. (A) Speech spectrogram of a single subject articulating sentences during DAF conditions showing a marked increase in articulation duration. **(B)** Mean articulation duration of sentences during DAF conditions averaged across subjects showing a significant effect of duration. Error bars show SEM over subjects. **(C)** High gamma responses to articulation of sentences during DAF conditions averaged across electrodes in Cluster 1. Shaded regions indicate SEM over trials. **(D)** High gamma responses to articulation of sentences during DAF conditions averaged across electrodes in Cluster 2. Shaded regions indicate SEM over trials. **(E)** Anatomical map of electrodes across all subjects displayed on the left hemisphere of a template brain showing the neural sensitivity to DAF during word-reading. **(F)** Anatomical map of electrodes across all subjects displayed on the left hemisphere of a template brain showing the neural sensitivity to DAF during sentences-reading. **(G)** Scatter plot and fitted regression showing significant correlation between sensitivity to DAF for the word-reading and sentence-reading tasks. Each circle represents one electrode. **SEM** = Standard error of the mean

We have shown that reading words with DAF did not prolong articulation duration (**Fig. 1B**) and while it increased neural responses in auditory regions, it did not affect responses in motor regions (**Fig. 2B-D**). We hypothesized that a longer and more complex stimulus may elicit a stronger behavioral response and predicted that motor regions will show an effect of DAF when articulation is strongly affected. To test this prediction, we performed another experiment in which subjects read aloud sentences during DAF. Indeed, articulating longer speech segments (8-word sentences) during DAF resulted in a significantly stronger behavioral effect (**Fig. 4A**). Articulation duration increased significantly with delay: average articulation duration across subjects was 2.761, 2.942, 3.214 and 3.418 seconds for no delay, 50, 100 and 200 milliseconds delay conditions respectively (**Fig. 4B, ANOVA:** $F = 17.11$, $p = 0.0001$).

Next, we examined the neural response to DAF in the two electrode clusters we identified previously (**Fig. 2B**). When reading words during DAF, amplitude of the neural response increased with delay in Cluster 1 but not in Cluster 2 (**Fig. 2C and 2D**). However, when reading sentences during DAF, neural response in both clusters showed a sustained effect (**Fig. 4C and 4C, ANOVA:** $F = 18$, $p = 2.95\times10^{-5}$ for Cluster 1 and $F = 4.8$, $p = 0.03$ for Cluster 2). Also, when reading words during DAF, duration of the neural response in neither of the clusters showed a significant effect of delay. However, when reading sentences during DAF, neural response duration in both clusters increased significantly with delay paralleling the significant behavioral effect of DAF on articulation duration (**ANOVA:** $F = 21.6$, $p = 10^{-5}$ for Cluster 1 and $F = 35.5$, $p = 10^{-8}$ for Cluster 2).

To identify brain regions that showed enhanced response to reading sentences during DAF, we recalculated the sensitivity index for each electrode by measuring the trial-by-trial correlation between the delay and the neural response averaged over a 0 to 3 s time window.

Comparing sensitivity indices for word-reading and sentence-reading tasks revealed that electrodes that were sensitive to DAF for the word-reading task were also sensitive for the sentence-reading task ($r = 0.7$, $p = 1.2\times10^{-38}$, **Fig. 4G**), however the majority of electrodes showed larger sensitivity to DAF for the sentence-reading task. Moreover, several sites such as IFG and dorsal precentral gyrus showed increased sensitivity in the sentence-reading task (**Fig. 4E-F**). This result suggests that articulating longer speech segments during DAF results in stronger overall sensitivity across auditory and motor regions and engages a larger brain network uniquely recruiting additional frontal regions.

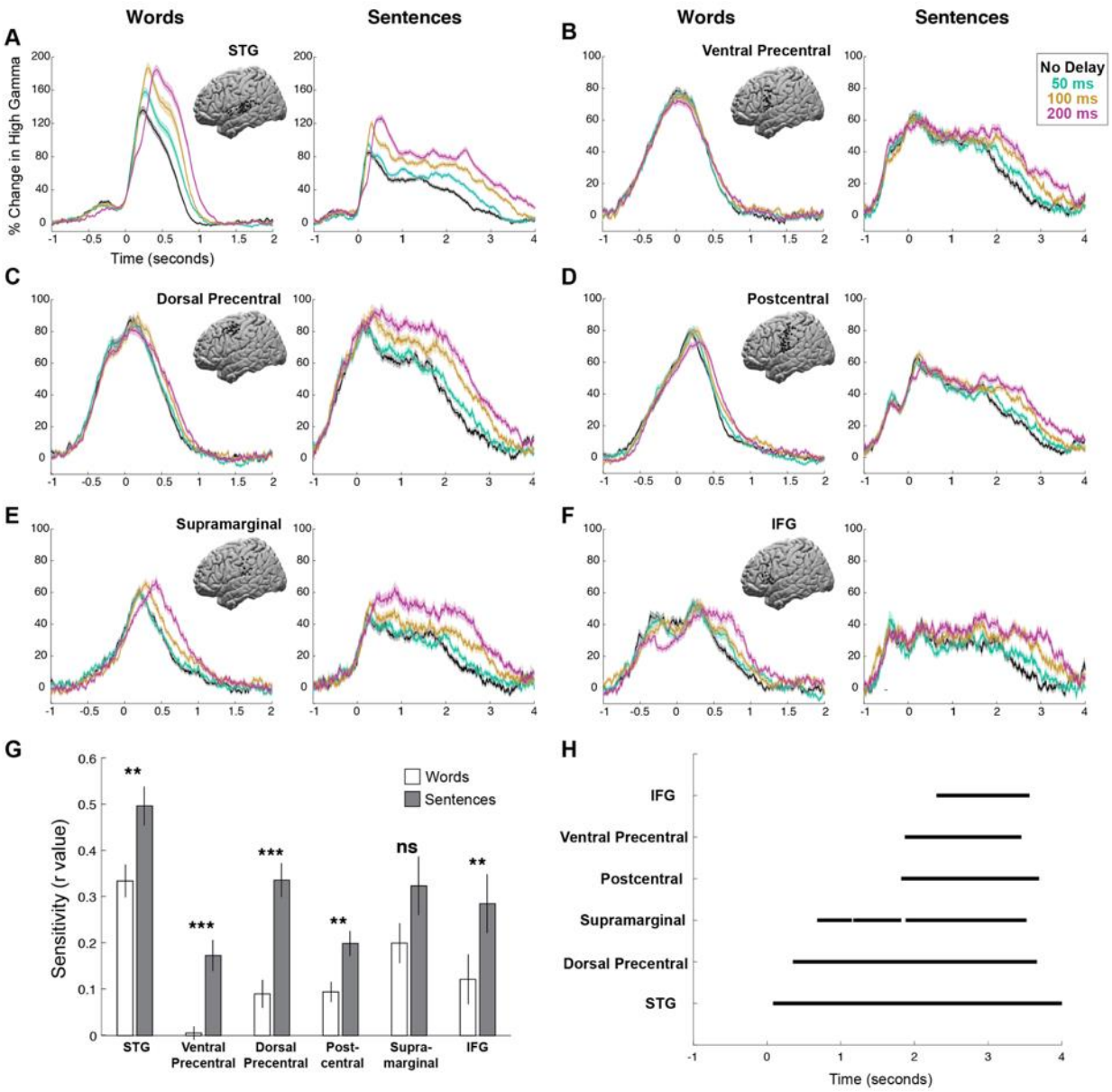

Figure 5 Neural responses to DAF by regions. (A-F) High gamma responses to articulation of words and sentences during DAF in six different regions: STG (a), ventral precentral (b), dorsal precentral (c), postcentral (d), supramarginal (e) and IFG (f). Inset brain figures shows the location of electrodes across all subjects on the left hemisphere of a template brain. Colors represent the various DAF conditions and shaded regions indicate SEM over trials. **(G)** Sensitivity to DAF during word-reading and sentence-reading tasks averaged across electrodes (error bars indicate SEM over electrodes) in six different regions. **(H)** Time intervals when the neural response to reading sentences during DAF diverged significantly across conditions within each of the six different regions. Significance is assessed using a shuffled permutation approach. **SEM** = Standard error of the mean

We further examined the neural response to DAF in six different regions of interest based on within subject anatomy: STG, ventral precentral gyrus, dorsal precentral gyrus, postcentral

gyrus, supramarginal gyrus and IFG (**Fig. 5A-F**). Comparing sensitivity indices for word-reading and sentence-reading tasks in these regions revealed that all six regions showed larger sensitivity to DAF during sentence-reading (unpaired ttest; STG: $t = 3.13$, $p = 0.0025$, ventral precentral: $t = 4.55$, $p = 2.7\times10^{-5}$, dorsal precentral: $t = 5.15$, $p = 3\times10^{-6}$, postcentral: $t = 3.1$, $p = 0.0024$, supramarginal: $t = 1.87$, $p = 0.07$, IFG: $t = 2.89$, $p = 0.0069$; **Fig. 5G**). To reveal how response enhancement to DAF changed across time during the sentence-reading task, we performed a one-way ANOVA at each time point (*see Methods*) and marked the timepoints when the neural response to the four delay conditions were significantly different for at least 200 consecutive milliseconds. Significant divergence onset during sentence-reading started the earliest in STG at 80 ms after speech onset, followed by dorsal precentral gyrus at 350 ms and supramarginal gyrus at 680 ms, and lasted throughout the stimulus. In postcentral, ventral precentral gyrus and IFG, responses diverged much later, at 1.82, 1.87 and 2.30 s respectively (**Fig. 5H**). These timings reflect when cortical regions become sensitive to DAF and provide evidence for two distinct timeframes of early (STG, dorsal precentral, supramarginal) and late (postcentral, ventral precentral, IFG) recruitment.

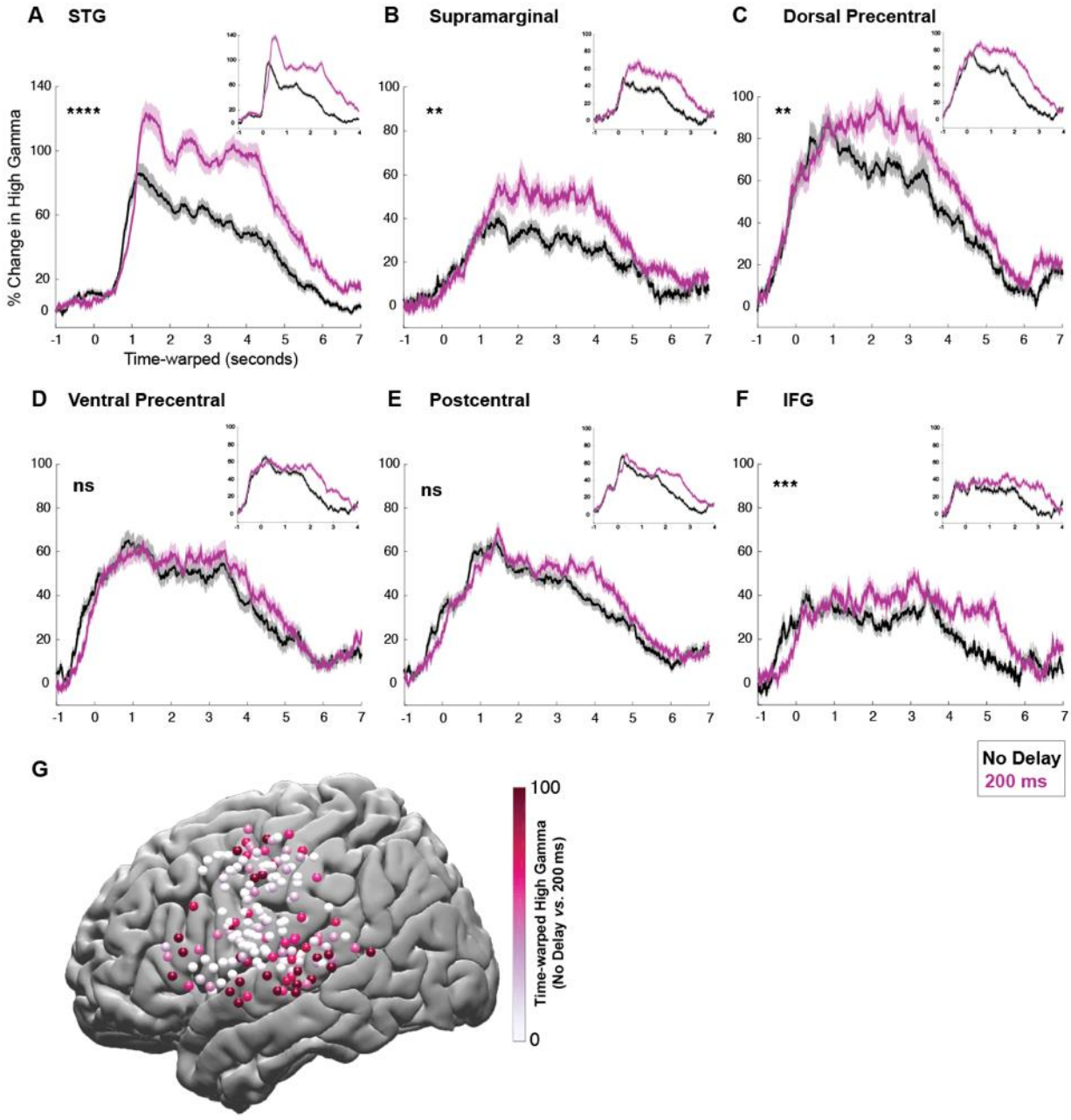

Figure 6 Time warped neural responses during sentence reading with DAF. (A-F) High gamma responses after correction for articulation duration using dynamic time warping. Activity locked to articulation of sentences is shown for no delay (black) and 200 ms delay (magenta) conditions in six different regions: STG, supramarginal, dorsal precentral, ventral precentral, postcentral and IFG. Inset figures show the uncorrected high gamma responses which include the normal articulation timing. **(G)** Spatial distribution of the increase in neural responses to sentence-reading during DAF (200 ms) compared to no delay condition. Electrodes across all subjects are displayed on the left hemisphere of a template brain.

Examining different regions of the speech network revealed variable degree of neural response enhancement to DAF. The increase in response amplitude was usually accompanied by

an increase in response duration, which was a result of longer articulation duration. In order to disentangle the enhanced neural response representing an error signal with longer articulation duration due to the exerted behavior, we applied a temporal normalization technique. We transformed the neural response time series using dynamic time warping (DTW) so that they would match in time span. DTW measures the similarity between two temporal sequences with different lengths by estimating a distance metric (a warping path) that would transform and align them in time. Matching the neural responses in time allowed us to directly compare their amplitudes and identify which brain regions produce an error signal in response to DAF rather than just sustained activity in time due to longer articulation.

We compared two conditions, which show the largest neural response difference in terms of amplitude and duration; 0 and 200 ms delay conditions (see *Dynamic Time Warping Analysis in Methods*). After aligning the responses in time, we averaged the amplitudes of the time-warped signals over time (0-6 s) and compared the two conditions by running an unpaired t-test. Amplitudes of the time-warped responses to DAF were significantly larger in STG ($t = 5.5$, $p = 7.7\times10^{-8}$), supramarginal gyrus ($t = 3.06$, $p = 0.0025$), dorsal precentral gyrus ($t = 2.4$, $p = 0.016$) and IFG ($t = 3.4$, $p = 10^{-3}$) but not in ventral precentral gyrus ($t = 0.82$, $p = 0.41$) and postcentral ($t = 1.66$, $p = 0.0971$) regions (**Fig. 6A-F**). Lastly, to examine the spatial distribution of the (time-warped) error response in more detail, we calculated the percent increase in response amplitude in single electrodes (HGB_{200} - $HGB_{no\ delay}$ / $HGB_{no\ delay}$*100). This analysis revealed that the magnitude of the error response was variable both across and within the regions of the speech network. Overall the error signal centered around four major cortical networks: STG, IFG, supramarginal gyrus and dorsal precentral gyrus (**Fig. 6G**).

Discussion

Our study is one of the few electrophysiological investigations (Swink and Stuart 2012, Toyomura, Miyashiro et al. 2020), and to our knowledge, the only ECoG investigation of neural responses to speech with delayed auditory feedback. We compared the effects of DAF on producing isolated versus connected speech by using word and sentence stimuli and showed that producing sentences during DAF had a stronger disruptive effect on speech. We used an unsupervised clustering algorithm (NMF) to identify auditory and motor regions involved in speech production, which revealed a robust auditory error signal characterized by enhanced neural responses to increased feedback delays. We found that most auditory sites sensitive to DAF also showed suppressed neural activity during speech compared to listening. This significant correlation between speech-induced suppression and DAF sensitivity provides the first evidence in humans for a shared mechanism between auditory corollary discharge and error monitoring. We further identified four subregions of the speech network that are centrally engaged in the processing of auditory feedback: STG, supramarginal gyrus, dorsal precentral gyrus and IFG. The exquisite resolution of ECoG provided us with precise spatiotemporal evolution of feedback processing in these distinct regions. Neural responses were enhanced in amplitude and extended in duration for large delays reflecting the error signal caused by altered feedback and the subsequent longer articulation. To dissociate the error signal from the effect of prolonged articulation, we used dynamic time warping algorithm and temporally aligned the neural signals with the patients' speech acoustics. Our results highlighted dorsal precentral gyrus as a critical region for processing DAF. Dorsal precentral gyrus was engaged only when DAF had a stronger disruptive effect on speech during sentence-reading and its engagement started early suggesting a critical role in auditory-motor integration and correction of vocalization.

We demonstrated that when subjects read aloud single words during DAF, neural response amplitude in auditory sites increased as a function of feedback delay. This result suggests that auditory error signal does not simply encode a mismatch between intended and perceived speech but is sensitive to the amount of mismatch. Interestingly, the sensitivity to DAF across auditory sites varied, some sites showing stronger sensitivity compared to others. To explain the variability of DAF sensitivity across auditory sites, we considered the role of speech-induced auditory suppression, which has been reported in previous studies in the form of a response reduction in auditory cortex during speaking compared with listening (Eliades and Wang 2003, Flinker, Chang et al. 2010, Greenlee, Jackson et al. 2011). Current models of speech motor control predicted a shared mechanism between auditory suppression and sensitivity to DAF (Houde and Nagarajan 2011, Tourville and Guenther 2011), suggesting a role for auditory suppression in feedback monitoring. In support of this prediction, single-unit activity in auditory neurons of marmoset monkeys showed that neurons that were suppressed during vocalization exhibited increased activity during frequency-shifted feedback (Eliades and Wang 2008). However, this finding has not been replicated in humans. In contrary, a previous ECoG study that used frequency-shifted feedback during production of a vowel showed that suppressed auditory sites do not overlap with sites that are sensitive to feedback alterations (Chang, Niziolek et al. 2013). It is possible that a larger auditory neural population is highly sensitive to temporal rather than spectral features of the auditory feedback. Indeed, using DAF instead of frequency-shifted feedback, we demonstrated a significant correlation between auditory suppression and sensitivity to feedback. This provides evidence for an overlap of the two neural populations as well as a tight link between auditory cortex suppression and feedback monitoring, as in the non-human primates (Eliades and Wang 2008).

Reading single words during DAF elicited a nominal behavioral effect and increased neural responses only in auditory but not in motor regions. However, when subjects read aloud sentences during DAF, this longer and more complex stimulus elicited a much stronger behavioral effect and increased neural responses in motor as well as auditory regions. STG, supramarginal gyrus, dorsal precentral gyrus and IFG showed enhanced neural responses to DAF, which are typically modeled as the main components of the dorsal pathway for speech that is responsible for sensorimotor integration and auditory feedback processing (Hickok and Poeppel 2007). Analysis of the time course of responses revealed that response enhancement started early in STG, dorsal precentral gyrus and supramarginal gyrus providing further evidence for a functional correspondence between these regions during altered feedback. In support of theoretical models of speech processing, clinical reports demonstrated that posterior STG and supramarginal gyrus damage are implicated in conduction aphasia (Fridriksson, Kjartansson et al. 2010) and patients with conduction aphasia are less affected by DAF (Boller and Marcie 1978) indicating the involvement of these regions in feedback processing.

A surprising result was the early response enhancement to DAF in dorsal precentral gyrus, which occurred only during sentence-reading. Dorsal precentral gyrus is a complex functional region implicated in auditory, motor and visual speech processing (Pulvermuller, Huss et al. 2006, Callan, Jones et al. 2014, Cheung, Hamiton et al. 2016, Ozker, Yoshor et al. 2018). Predictive speech models suggest that dorsal precentral region encodes internal models of speech by mapping visual speech gestures or acoustic speech features onto the corresponding articulatory movements. During speech production, this internal prediction is used in combination with the auditory feedback from hearing's one's own voice to adjust the motor commands to produce intended speech (Houde and Nagarajan 2011, Tourville and Guenther

2011). For sentence-reading during DAF, when introduced delay renders auditory feedback less reliable, processing may rely more heavily on the internal prediction evidenced by increased dorsal precentral gyrus engagement. Dorsal precentral gyrus is critically recruited when speech production becomes more effortful due to less reliable auditory feedback.

DAF leads to disfluency and reduced speech rate in fluent speakers (Lee 1950, Black 1951, Fairbanks 1955, Stuart, Kalinowski et al. 2002). We measured articulation duration as the total articulation time and demonstrated that it increased with increasing amount of delays. Speech paradigms in previous DAF studies used various amounts of delays ranging from 25 to 800 milliseconds and consistently reported that the strongest disruption of speech occurred at 200 millisecond delay (Lee 1950, Black 1951, Fairbanks 1955, Stuart, Kalinowski et al. 2002, Yamamoto and Kawabata 2014). In line with previous reports, we showed that the strongest disruption occurred at 200 milliseconds delay for both word-reading and sentence-reading tasks. This time interval is thought to be critical for sensorimotor integration during speech production because it is of about the same order of average syllable duration. Given that the temporal distance between two consecutive stressed syllables is roughly 200 milliseconds, it has been suggested that delaying auditory feedback by this amount of time causes a rhythmical interference that results in the maximal disruption of speech fluency (Kaspar and Rübeling 2011).

The dynamics of the cortical speech network can also provide further explanation for the maximal disruption of speech at 200 ms feedback delay. Previous ECoG studies showed that IFG is activated before articulation onset and remained silent during articulation, while motor cortex is activated both before and during articulation. These studies suggested that IFG produces an articulatory code that is subsequently implemented by the motor cortex and reported a ~200 ms

temporal lag between IFG and motor cortex activation (Flinker, Korzeniewska et al. 2015, Magrassi, Aromataris et al. 2015). A feedback delay in the same order of this temporal lag likely interrupts propagation of the articulatory code from IFG to motor cortex, thereby disrupting speech. Unlike prior reports, we found sustained IFG activity throughout speech production, however this was during DAF where sustained IFG recruitment may be necessary to support compensatory speech correction. The onset of IFG activation was seen in conjunction with dorsal precentral gyrus, however sensitivity to DAF was seen at two distinct time periods with an early recruitment of dorsal precentral gyrus and much later involvement of IFG.

DAF behavioral paradigms have been widely used for decades to understand vocal monitoring, however the cortical dynamics underlying this process remained largely unknown. We elucidate the magnitude, timing and spatial distribution of the error signal between produced and perceived speech, controlling for articulation. Our results highlight dorsal precentral gyrus as a selective region which is recruited immediately when speech feedback becomes unreliable and production more effortful, implicating it in auditory-motor mapping that underlies vocal monitoring of human speech.

Acknowledgement

We thank Zhuoran Huand and Qingyang Zhu for their assistance in analyzing voice recordings of the subjects.

Funding

This study was supported by grants from the NIH (F32 DC018200 Ruth L. Kirschstein postdoctoral fellowship from the National Institute on Deafness and Other Communication

Disorders to M.O. and R01NS109367 from the National Institute of Neurological Disorders and Stroke to A.F.) and by the Leon Levy Foundation Fellowship (to M.O.).

Data availability

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Competing interests

The authors report no competing interests.

References

Ashburner, J. (2007). "A fast diffeomorphic image registration algorithm." Neuroimage **38**(1): 95-113.

Behroozmand, R., L. Karvelis, H. Liu and C. R. Larson (2009). "Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation." Clin Neurophysiol **120**(7): 1303-1312.

Behroozmand, R., R. Shebek, D. R. Hansen, H. Oya, D. A. Robin, M. A. Howard, 3rd and J. D. Greenlee (2015). "Sensory-motor networks involved in speech production and motor control: an fMRI study." Neuroimage **109**: 418-428.

Black, J. W. (1951). "The effect of delayed side-tone upon vocal rate and intensity." Journal of Speech and Hearing Disorders **16**(1): 56-60.

Blanchet, P. and P. Hoffman (2014). "Factors influencing the effects of delayed auditory feedback on dysarthric speech associated with Parkinson's disease." Commun Disord Deaf Stud Hearing Aids **2**(106): 2.

Bloodstein, O. (1969). "A handbook on stuttering."
Boller, F. and P. Marcie (1978). "Possible role of abnormal auditory feedback in conduction aphasia." Neuropsychologia **16**(4): 521-524.

Callan, D. E., J. A. Jones and A. Callan (2014). "Multisensory and modality specific processing of visual speech in different regions of the premotor cortex." Front Psychol **5**: 389.

Chang, E. F., C. A. Niziolek, R. T. Knight, S. S. Nagarajan and J. F. Houde (2013). "Human cortical sensorimotor network underlying feedback control of vocal pitch." Proc Natl Acad Sci U S A **110**(7): 2653-2658.

Cheung, C., L. S. Hamiton, K. Johnson and E. F. Chang (2016). "The auditory representation of speech sounds in human motor cortex." Elife **5**.

Civier, O., S. M. Tasko and F. H. Guenther (2010). "Overreliance on auditory feedback may lead to sound/syllable repetitions: simulations of stuttering and fluency-inducing conditions with a neural model of speech production." J Fluency Disord **35**(3): 246-279.

Crapse, T. B. and M. A. Sommer (2008). "Corollary discharge across the animal kingdom." Nature Reviews Neuroscience **9**(8): 587-600.

Ding, C., X. He and H. D. Simon (2005). On the equivalence of nonnegative matrix factorization and spectral clustering. Proceedings of the 2005 SIAM international conference on data mining, SIAM.

Eliades, S. J. and X. Wang (2003). "Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations." Journal of neurophysiology **89**(4): 2194-2207.

Eliades, S. J. and X. Wang (2008). "Neural substrates of vocalization feedback monitoring in primate auditory cortex." Nature **453**(7198): 1102-1106.

Fairbanks, G. (1955). "Selective vocal effects of delayed auditory feedback." Journal of speech and hearing disorders **20**(4): 333-346.

Flinker, A., E. F. Chang, H. E. Kirsch, N. M. Barbaro, N. E. Crone and R. T. Knight (2010). "Single-trial speech suppression of auditory cortex activity in humans." Journal of Neuroscience **30**(49): 16643-16650.

Flinker, A., A. Korzeniewska, A. Y. Shestyuk, P. J. Franaszczuk, N. F. Dronkers, R. T. Knight and N. E. Crone (2015). "Redefining the role of Broca's area in speech." Proc Natl Acad Sci U S A **112**(9): 2871-2875.

Fridriksson, J., O. Kjartansson, P. S. Morgan, H. Hjaltason, S. Magnusdottir, L. Bonilha and C. Rorden (2010). "Impaired speech repetition and left parietal lobe damage." J Neurosci **30**(33): 11057-11061.

Goldberg, T. E., J. M. Gold, R. Coppola and D. R. Weinberger (1997). "Unnatural practices, unspeakable actions: a study of delayed auditory feedback in schizophrenia." Am J Psychiatry **154**(6): 858-860.

Greenlee, J. D., A. W. Jackson, F. Chen, C. R. Larson, H. Oya, H. Kawasaki, H. Chen and M. A. Howard III (2011). "Human auditory cortical activation during self-vocalization." PloS one **6**(3): e14744.

Hamilton, L. S., E. Edwards and E. F. Chang (2018). "A spatial map of onset and sustained responses to speech in the human superior temporal gyrus." Current Biology **28**(12): 1860-1871. e1864.

Hardy, C. J. D., R. L. Bond, K. Jaisin, C. R. Marshall, L. L. Russell, K. Dick, S. J. Crutch, J. D. Rohrer and J. D. Warren (2018). "Sensitivity of Speech Output to Delayed Auditory Feedback in Primary Progressive Aphasias." Front Neurol **9**: 894.

Hashimoto, Y. and K. L. Sakai (2003). "Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: an fMRI study." Hum Brain Mapp **20**(1): 22-28.

Hickok, G., J. Houde and F. Rong (2011). "Sensorimotor integration in speech processing: computational basis and neural organization." Neuron **69**(3): 407-422.

Hickok, G. and D. Poeppel (2007). "The cortical organization of speech processing." Nature reviews neuroscience **8**(5): 393-402.

Hirano, S., H. Kojima, Y. Naito, I. Honjo, Y. Kamoto, H. Okazawa, K. Ishizu, Y. Yonekura, Y. Nagahama, H. Fukuyama and J. Konishi (1997). "Cortical processing mechanism for vocalization with auditory verbal feedback." Neuroreport **8**(9-10): 2379-2382.

Houde, J. F. and M. I. Jordan (1998). "Sensorimotor adaptation in speech production." Science **279**(5354): 1213-1216.

Houde, J. F. and S. S. Nagarajan (2011). "Speech production as state feedback control." Frontiers in human neuroscience **5**: 82.

Jones, J. A. and K. G. Munhall (2000). "Perceptual calibration of F0 production: evidence from feedback perturbation." J Acoust Soc Am **108**(3 Pt 1): 1246-1251.

Kalinowski, J., A. Stuart, S. Sark and J. Armson (1996). "Stuttering amelioration at various auditory feedback delays and speech rates." International Journal of Language & Communication Disorders **31**(3): 259-269.

Kaspar, K. and H. Rübeling (2011). "Rhythmic versus phonemic interference in delayed auditory feedback." Journal of Speech, Language, and Hearing Research.
Lee, B. S. (1950). "Effects of delayed speech feedback." The Journal of the Acoustical Society of America **22**(6): 824-826.

Lin, I. F., T. Mochida, K. Asada, S. Ayaya, S. Kumagaya and M. Kato (2015). "Atypical delayed auditory feedback effect and Lombard effect on speech production in high-functioning adults with autism spectrum disorder." Front Hum Neurosci **9**: 510.

Liu, H., E. Q. Wang, L. V. Metman and C. R. Larson (2012). "Vocal responses to perturbations in voice auditory feedback in individuals with Parkinson's disease." PLoS One **7**(3): e33629.

Lombard, E. (1911). "Le signe de l'elevation de la voix." Ann. Mal. de L'Oreille et du Larynx: 101-119.

Magrassi, L., G. Aromataris, A. Cabrini, V. Annovazzi-Lodi and A. Moro (2015). "Sound representation in higher language areas during language generation." Proc Natl Acad Sci U S A **112**(6): 1868-1873.

Mukamel, R., H. Gelbard, A. Arieli, U. Hasson, I. Fried and R. Malach (2005). "Coupling between neuronal firing, field potentials, and FMRI in human auditory cortex." Science **309**(5736): 951-954.

Nir, Y., L. Fisch, R. Mukamel, H. Gelbard-Sagiv, A. Arieli, I. Fried and R. Malach (2007). "Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations." Current biology **17**(15): 1275-1285.

Niziolek, C. A. and F. H. Guenther (2013). "Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations." J Neurosci **33**(29): 12090-12098.

Ozker, M., D. Yoshor and M. S. Beauchamp (2018). "Frontal cortex selects representations of the talker's mouth to aid in speech perception." Elife **7**.

Pulvermuller, F., M. Huss, F. Kherif, F. Moscoso del Prado Martin, O. Hauk and Y. Shtyrov (2006). "Motor cortex maps articulatory features of speech sounds." Proc Natl Acad Sci U S A **103**(20): 7865-7870.

Ray, S. and J. H. Maunsell (2011). "Different origins of gamma rhythm and high-gamma activity in macaque visual cortex." PLoS Biol **9**(4): e1000610.

Schneider, D. M. and R. Mooney (2018). "How movement modulates hearing." Annual review of neuroscience **41**: 553-572.

Stuart, A., J. Kalinowski, M. P. Rastatter and K. Lynch (2002). "Effect of delayed auditory feedback on normal speakers at two speech rates." J Acoust Soc Am **111**(5 Pt 1): 2237-2241.

Swink, S. and A. Stuart (2012). "The effect of gender on the N1–P2 auditory complex while listening and speaking with altered auditory feedback." Brain and language **122**(1): 25-33.

Takaso, H., F. Eisner, R. J. Wise and S. K. Scott (2010). "The effect of delayed auditory feedback on activity in the temporal lobe while speaking: a positron emission tomography study." J Speech Lang Hear Res **53**(2): 226-236.

Tourville, J. A. and F. H. Guenther (2011). "The DIVA model: A neural theory of speech acquisition and production." Lang Cogn Process **26**(7): 952-981.

Tourville, J. A., K. J. Reilly and F. H. Guenther (2008). "Neural mechanisms underlying auditory feedback control of speech." Neuroimage **39**(3): 1429-1443.

Toyomura, A., D. Miyashiro, S. Kuriki and P. F. Sowman (2020). "Speech-Induced Suppression for Delayed Auditory Feedback in Adults Who Do and Do Not Stutter." Frontiers in Human Neuroscience **14**: 150.

Yamamoto, K. and H. Kawabata (2014). "Adaptation to delayed auditory feedback induces the temporal recalibration effect in both speech perception and production." Exp Brain Res **232**(12): 3707-3718.

Yang, A. I., X. Wang, W. K. Doyle, E. Halgren, C. Carlson, T. L. Belcher, S. S. Cash, O. Devinsky and T. Thesen (2012). "Localization of dense intracranial electrode arrays using magnetic resonance imaging." Neuroimage **63**(1): 157-165.

Zollinger, S. A. and H. Brumm (2011). "The Lombard effect." Curr Biol **21**(16): R614-615.