Dynamic auditory contributions to error detection revealed in the discrimination of Same and Different syllable pairs

David Jenson\textsuperscript{a,}\textsuperscript{*}, Tim Saltuklaroglu\textsuperscript{b}

\textsuperscript{a} Washington State University, Elson S. Floyd College of Medicine, Department of Speech and Hearing Sciences, Spokane, WA, USA
\textsuperscript{b} University of Tennessee Health Science Center, College of Health Professions, Department of Audiology and Speech Pathology, Knoxville, TN, USA

**A R T I C L E   I N F O**

**Keywords:**
EEG
Sensorimotor
Internal models
Working memory
Speech-induced suppression
Analysis by synthesis

**A B S T R A C T**

During speech production auditory and motor regions within the sensorimotor dorsal stream operate in concert to facilitate online error detection. As the dorsal stream also is known to activate in speech perception, the purpose of the current study was to probe the role of auditory regions in error detection during auditory discrimination tasks as stimuli are encoded and maintained in working memory. \textit{A priori} assumptions are that sensory mismatch (i.e., error) occurs during the discrimination of Different (mismatched) but not Same (matched) syllable pairs. Independent component analysis was applied to raw EEG data recorded from 42 participants to identify bilateral auditory alpha rhythms, which were decomposed across time and frequency to reveal robust patterns of event related synchronization (ERS; inhibition) and desynchronization (ERD; processing) over the time course of discrimination events. Results were characterized by bilateral peri-stimulus alpha ERD transitioning to alpha ERS in the late trial epoch, with ERD interpreted as evidence of working memory encoding via Analysis by Synthesis and ERS considered evidence of speech-induced-suppression arising during covert articulatory rehearsal to facilitate working memory maintenance. The transition from ERD to ERS occurred later in the left hemisphere for Different trials than for Same trials, with ERD and ERS temporally overlapping during the early post-stimulus window. Results were interpreted to suggest that the sensory mismatch (i.e., error) arising from the comparison of the first and second syllable elicits further processing in the left hemisphere to support working memory encoding and maintenance. Results are consistent with auditory contributions to error detection during both encoding and maintenance stages of working memory, with encoding stage error detection associated with stimulus concordance and maintenance stage error detection associated with task-specific retention demands.

1. \textbf{Introduction}

Neuroimaging studies over the last two decades have clearly identified the dorsal stream (DS) as an essential network for providing sensorimotor control for speech production (Tourville and Guenther, 2011; Guenther, 2016) and perception (Hickok and Poeppel, 2004, 2007). In speech production, anterior DS (i.e., premotor cortex; PMC) encodes the sensory consequences of the forthcoming motor plan (forward models) in projections to posterior DS (i.e., middle and posterior superior temporal gyrus; mSTG and pSTG). Detected mismatches between forward models and sensory targets or available reafferent input give rise to sensory error signals (Blakemore et al., 2000; Heinks-Maldonado et al., 2005), with motor plans subsequently updated via sensory-to-motor feedback (i.e., inverse modeling). Despite the absence of a motor plan to be executed, it is becoming clearer that DS functions in speech perception and discrimination tasks in many ways parallel those of production, largely due to the cognitive demands of the tasks (Wostmann et al., 2017). Pre-stimulus PMC activity (Callan et al., 2010) suggests that attentional mechanisms are supported by forward modeling, though to help focus attention and constrain forthcoming auditory analyses (Schroeder et al., 2010; Zion Golumbic et al., 2012). Post-stimulus activity in both PMC (Jenson et al., 2014a, 2019a) and pSTG (Jenson et al., 2015) suggests that phonological working memory is supported by instantiations of forward and inverse models (Wilson, 2001). However, when comparing anterior DS functions in speech perception and production, one question remains largely unanswered. In the absence of motor plans and self-generated re-afferent feedback from production, can auditory regions engage in error detection when...
processing and retaining stimuli in working memory? To successfully interrogate this notion, it is first essential to consider the specific patterns of neural activity related to error detection in auditory cortex.

Error detection occurs simultaneously with speech production, allowing for online monitoring of co-articulated speech. During normal, error-free speech, sensory predictions coded in the forward model align with auditory targets and available reafference, reducing the need for corrective feedback and suppressing auditory responses. This ‘speech-induced-suppression’ is commonly observed as a net inhibition of activity in auditory regions (Greenlee et al., 2011; Chang et al., 2013). However, under conditions of altered feedback (Watkins et al., 2005; Parkinson et al., 2012; Behroozmand et al., 2015), the sensory mismatch arising from the comparison of prediction to altered reafference leads to elevated auditory activity, with feedback returned to anterior motor regions via an inverse (i.e., sensory to motor) model for integration into ongoing motor planning (Houde and Chang, 2015; Guenther, 2016).

Interestingly, auditory suppression is also observed during covert (i.e., imagined) speech (Kauramaki et al., 2015; Balk et al., 2015), suggesting that covert production is not necessary to elicit error detection processes. Covert speech contains a motor plan that is not executed and thus does not give rise to a reafferent signal against which forward models can be compared. Consequently, the paired forward and inverse internal models instantiating covert production (Pickering and Garrod, 2013) support error detection by comparing forward model predictions to sensory targets in auditory regions (Tian and Poeppel, 2010; Jenson et al., 2015). As internal modeling is thought to be involved in the cognitive mechanisms scaffolding perception (Skipper et al., 2006, 2017), it is proposed that error detection in auditory cortex also may occur when multiple sounds are being held in working memory for comparison purposes.

Net sensorimotor activity from the DS in speech perception typically correlates with the cognitive demands of the task (Alho et al., 2012, 2014; Deng et al., 2012; Peschke et al., 2012; Wostmann et al., 2017). Electroencephalographic (EEG) time-frequency studies also clearly demonstrate that, in contrast to production, DS activity in perception is highly variable across the time course of the task (Bowers et al., 2013; Jenson et al., 2014a; Jenson et al., 2014b; Saltuklaroglu et al., 2018). This is evident through evaluating patterns of event-related synchronization (ERS) and desynchronization (ERD) over the time course of perceptual tasks (see Table 1 for a fuller explanation). Within these oscillatory frameworks, ERS refers to an increase in spectral power (Pluetscheller and Lopes Da Silva, 1999) reflecting the functional deactivation of local neural assemblies by the alignment of oscillatory phase (Jensen and Mazaheri, 2010; Schroeder et al., 2010), while ERD refers to a decrease in spectral power, indicating a desynchronization of local neural assemblies as they prepare for and engage in active processes (Pluetscheller, 1992; Wiarda and Ross, 2019). Specifically, patterns of alpha and beta ERS/ERD vary greatly within a perception task as cognitive demands shift from early attention (pre-stimulus) to later working memory (post-stimulus; Jenson et al., 2019b). Unlike passive perception tasks, syllable discrimination tasks require participants to retain sounds while they are compared in working memory, making them particularly well-suited for evaluating the possibility of sensorimotor-based error detection processes. In discrimination studies that measure activity in the anterior DS mu rhythm, the most robust EEG activity is typically found post-stimulus and characterized by concurrent alpha and beta ERD suggesting paired inverse (alpha) and forward (beta) modeling, respectively (Jenson et al., 2014a, 2019b; Saltuklaroglu et al., 2017, 2018; Jenson, 2021).

Interestingly, this pattern of mu activity is similar to that observed in both covert and overt speech production (Jenson et al., 2014a, 2018), supporting notions that in discrimination tasks, syllables are held in phonological working memory by covert replay.

While error detection in speech production is well understood from altered feedback studies that induce sensory mismatches (Watkins et al., 2005; Parkinson et al., 2012; Behroozmand et al., 2015), it remains unclear how the mechanism operates to support discrimination as syllables are held in working memory. Initial acoustic traces are subject to rapid decay (Wilsch and Obleser, 2016) and therefore, perceived stimuli must be encoded and maintained in working memory to allow successful discrimination performance (Jacquemot and Scott, 2006). The encoding stage unfolds during the peri- and early post-stimulus window and supports the extraction of an articulatory representation of the stimulus (Wilsch and Obleser, 2016). The maintenance stage, in contrast, emerges in the post-stimulus window following the encoding stage and refers to the refreshing of working memory contents via covert articulatory rehearsal (Buchsbaum and DeJong, 2005). Working memory encoding has been associated with Analysis by Synthesis (Stevens and Halle, 1967; Poeppel and Nonnenmacher, 2011), whereby a coarse sketch of the auditory signal is relayed to anterior motor regions for mapping onto an articulatory (i.e., motor-based) hypothesis (Skipper et al., 2006, 2017). The sensory representation of this articulatory hypothesis is then returned to auditory regions for validation against the full signal, with any mismatch returned to motor regions for hypothesis revision (Skipper, 2014). Following the validation of the articulatory representation, working memory contents are refreshed through covert articulatory rehearsal (i.e., paired forward and inverse modeling) in the maintenance stage. However, while this notion is consistent with the results of studies probing the mu rhythm in discrimination tasks (Jenson et al., 2014a, 2019b), it remains unclear how error detection mechanisms operate as pairs of stimuli are encoded and held for comparison in working memory.

One means of clarifying this point is by examining differences in neural activity between Same (i.e., matched) and Different (i.e., unmatched) trials with a priori assumptions that a sensory mismatch only occurs in trials that are Different. In this vein, Jenson and Saltuklaroglu (2021) compared mu rhythm (anterior DS) activity in Same versus Different trials of correctly discriminated syllable pairs. Differences were found late in trials, consistent with encoding and maintenance of stimuli in working memory. Specifically, post-stimulus ERD was stronger in Different trials in the left hemisphere and Same trials in the right hemisphere. Data were interpreted through the framework of Analysis by Synthesis to suggest that the articulatory representation of the first syllable was used as a predictive template for the second syllable, allowing for rapid and metabolically-efficient detection of matched pairs. It was further proposed that the use of predictive templates elicited error signals in Different trials only, requiring further processing in the speech-specialized left hemisphere (Specht, 2014) to identify stimuli prior to the engagement of covert rehearsal mechanisms for stimulus retention. However, by only examining data from the anterior DS it was not possible to entirely rule out the possibility that increased late ERD in Different trials was solely due to an increased maintenance load (i.e., maintaining two distinct syllables in working memory rather than one reduplicated pair).

| Abbreviation | Definition | Physiologic process | Cognitive implication | Working memory implication |
|--------------|------------|---------------------|-----------------------|---------------------------|
| ERS          | Event related synchronization | Increase in spectral power relative to a baseline window | Deactivation of local neural assemblies by alignment of oscillatory phase | Associated with working memory maintenance (post-stimulus) |
| ERD          | Event related desynchronization | Decrease in spectral power relative to a baseline window | Local neural assemblies becoming independent to prepare for and execute active processes | Associated with working memory encoding (peri-stimulus) |
To fully interrogate error detection accounts of the mu findings reported in Jenson and Saltuklaroglu (2021), it is essential to consider posterior DS activity during the same tasks. Specifically, examination of auditory regions holds the potential to differentiate contributions of the encoding and maintenance stages of working memory. In their review of contemporary working memory models, Buchsbaum and D’esposito (2019) argue for a sensorimotor framework in which interactions between posterior sensory (i.e., auditory) and anterior motor regions instantiate the perceptual reactivation processes allowing information to be retained in working memory. Within this sensorimotor framework, posterior DS is identified as a region active during both the encoding and maintenance stages of working memory tasks, suggesting its sensitivity to both stages. Alpha is considered to be the oscillatory signature of posterior DS (i.e., auditory) regions (Tiihonen et al., 1991; Lehtelä et al., 1997), with several studies reporting a distinct auditory alpha rhythm distinguishable from other alpha generators by its sensitivity to auditory, but not visual or motor processes (Tiihonen et al., 1991; Weisz et al., 2011; Bastarrrika-Iriarte and Caballero-Gaudes, 2019). Of relevance to the current investigation, Wianda and Ross (2019) proposed distinct neural signatures for working memory stages, with encoding characterized by low alpha (7–10 Hz; Klimesch et al., 2006; 2007) ERD and maintenance characterized by high alpha (10–13 Hz) ERS, providing a robust framework for interpretation. Jenson et al. (2015) probed the auditory alpha rhythm during speech discrimination, identifying high alpha ERS in the late post stimulus window during which anterior DS regions were active (Jenson et al., 2014c). Hence, when mu (Jenson et al., 2014a) and auditory data were considered together, there appeared to be clear evidence of covert rehearsal and error detection in discrimination tasks as syllables are held in working memory (Jenson et al., 2015). In the current study, we will probe auditory alpha activity during the discrimination of Same/Different syllable pairs to characterize auditory contributions to error detection with peri-stimulus low alpha ERD and post-stimulus high alpha ERS considered indices of working memory encoding and maintenance, respectively (Wianda and Ross, 2019).

The overarching purpose of the current study is to further elucidate sensorimotor error detection mechanisms in speech discrimination. To this end, we will determine the influence of trial type (i.e., Same vs. Different) on auditory activity during speech discrimination. Consistent with previous findings (Jenson et al., 2015) and in accordance with encoding and then maintenance of stimuli (Jacquemot and Scott, 2006), auditory activity following stimulus offset is expected to be characterized by low alpha ERD (encoding) followed by high alpha ERS (maintenance). If the first syllable serves as predictive template or sensory target for which the second syllable is to be compared, then we expect to find differences in auditory activity for Same vs Different trials. Specifically, immediately following stimulus offset, as syllables are encoded in working memory, Different trials that produce a sensory mismatch are expected to be characterized by a more protracted low alpha ERD as the target for which the second syllable is to be compared, then we expect to find differences in auditory activity following stimulus offset. Thus, the length of stimuli was 2 s. It should be noted that while segmentation (separation of a syllable into constituents) is not required for successful discrimination performance, segmentation is known to modulate sensorimotor activity in perception tasks (Locasto et al., 2004; Thornton et al., 2018). Consequently, to minimize the potential for segmentation effects to influence sensorimotor activity, syllable pairs only differed by the initial consonant (i.e., the vowel was always the same in a syllable pair). Subject to this constraint, 36 distinct syllable pairs were generated. The stimuli used for the control condition consisted of white noise presented at 70 dB SPL.

The stimuli for the current study were generated from recordings of a male native English speaker producing CV syllables on an AKG 520 microphone paired with a Mackie 402-VLZ3 pre-amp and a Krohn-Hite model 3384 amplifier. Syllables were comprised of a voiced consonant (i.e.,/b/,/d/,/g/,/l/) paired with a vowel (i.e.,/i/,/a/,/e/). Recordings were digitized at 44.1 kHz and bandpass filtered between 20 Hz and 20 kHz. In order to curtail the potential impact of lexicality (Pratt et al., 1994; Chiappe et al., 2001; Koz et al., 2010; Ostrand et al., 2016), two of the syllables (i.e.,/bi/,/li/) were excluded, resulting in ten distinct syllables. Ten tokens of each syllable were recorded, with the best exemplar of each syllable selected on the basis of consonant quality, vowel quality, and overall duration. Selected tokens were bandpass filtered from 300 to 3400 Hz (Callan et al., 2010) and normalized for duration (200 ms) and intensity (70 dB SPL) using Audacity 2.0.6.

Syllable pairs for the purpose of discrimination were generated by inserting 200 ms of silence between syllables, with 1400 ms of silence following the offset of the second syllable. Thus, the length of stimuli was 2 s. It should be noted that while segmentation (separation of a syllable into constituents) is not required for successful discrimination performance, segmentation is known to modulate sensorimotor activity in perception tasks (Locasto et al., 2004; Thornton et al., 2018). Consequently, to minimize the potential for segmentation effects to influence sensorimotor activity, syllable pairs only differed by the initial consonant (i.e., the vowel was always the same in a syllable pair). Subject to this constraint, 36 distinct syllable pairs were generated. The stimuli used for the control condition consisted of white noise presented at 70 dB SPL.

### 2.2. Stimuli

The stimuli for the current study were generated from recordings of a male native English speaker producing CV syllables on an AKG 520 microphone paired with a Mackie 402-VLZ3 pre-amp and a Krohn-Hite model 3384 amplifier. Syllables were comprised of a voiced consonant (i.e.,/b/,/d/,/g/,/l/) paired with a vowel (i.e.,/i/,/a/,/e/). Recordings were digitized at 44.1 kHz and bandpass filtered between 20 Hz and 20 kHz. In order to curtail the potential impact of lexicality (Pratt et al., 1994; Chiappe et al., 2001; Koz et al., 2010; Ostrand et al., 2016), two of the syllables (i.e.,/bi/,/li/) were excluded, resulting in ten distinct syllables. Ten tokens of each syllable were recorded, with the best exemplar of each syllable selected on the basis of consonant quality, vowel quality, and overall duration. Selected tokens were bandpass filtered from 300 to 3400 Hz (Callan et al., 2010) and normalized for duration (200 ms) and intensity (70 dB SPL) using Audacity 2.0.6.

Syllable pairs for the purpose of discrimination were generated by inserting 200 ms of silence between syllables, with 1400 ms of silence following the offset of the second syllable. Thus, the length of stimuli was 2 s. It should be noted that while segmentation (separation of a syllable into constituents) is not required for successful discrimination performance, segmentation is known to modulate sensorimotor activity in perception tasks (Locasto et al., 2004; Thornton et al., 2018). Consequently, to minimize the potential for segmentation effects to influence sensorimotor activity, syllable pairs only differed by the initial consonant (i.e., the vowel was always the same in a syllable pair). Subject to this constraint, 36 distinct syllable pairs were generated. The stimuli used for the control condition consisted of white noise presented at 70 dB SPL.

### 2.3. Design

The data for the current study consist of a subset of the conditions reported in Jenson et al. (2019b), which employed a 2 x 3 (set size x signal clarity) within-subjects design referenced to a control condition (7 conditions total). The levels of Set Size were Small (/ba//da/only; 4 possible pairings) and Large (full stimulus complement; 36 possible pairings) and the levels of signal clarity were Quiet (silent background), Noise-masked, and Narrow-band filtered. The data for the current study were drawn from both levels of set size and a single level of signal clarity (i.e., Quiet). The conditions presented to subjects were therefore:

1. Passive listening to white noise
2. Discrimination of /ba//da/pairs in quiet (4 possible pairings)
3. Discrimination of full complement of syllable pairings in quiet (36 possible pairings)

Condition 1 was used as a control condition, while conditions 2 and 3 were active discrimination conditions in which subjects pressed one of two buttons on a keypad to indicate whether syllables were judged to be the same or different. In order to control for motor activity arising from the button press in discrimination conditions, a button press was included in the control condition. To minimize the potential for response bias (Venezia et al., 2012), an equal number of Same and Different trials were included in both of the discrimination conditions. To enable hypothesis testing in the current study, data from all Same trials (from conditions 2 and 3) were extracted and combined into a new condition, with no known history of hearing, attentional, cognitive, or communicative impairment participated in the current study. Handedness was assessed with the Edinburgh Handedness Inventory (Oldfield, 1971), and all subjects provided informed consent prior to participation. Based on reports that sensorimotor processing strategies may differ between males and females (Popovich et al., 2010; Kumari, 2011; Thornton et al., 2019), it was deemed necessary to restrict the subject pool to a single sex, and females were selected as a sample of convenience.

### 2. Methods

#### 2.1. Participants

42 female native English speakers (mean age = 24.1; 3 left handed)
and the same process was repeated for Different trials. The conditions used for the purpose of hypothesis testing in the current study were therefore:

a. Passive listening to white noise
b. Discrimination of Same trials
c. Discrimination of Different trials

d. Loading processed data files from all subject and conditions into the EEGLAB STUDY module;
e. Principal Component Analysis (PCA) to identify clusters of components exhibiting common patterns of activity across subjects;
f. Identification of bilateral auditory alpha clusters from the results of PCA;
g. ERSP decomposition of auditory alpha clusters; and
h. Source localization of bilateral auditory alpha clusters.

2.6.1. Individual pre-processing

All six raw data files for each subject were appended to create a single data file containing 240 trials. This aggregate data file was re-referenced to linked mastoid channels (M1, M2) for the reduction of common-mode noise and down-sampled to 256 Hz to reduce the
computational demands of subsequent processing steps. Correlation coefficients were then calculated between all pairs of channels, with correlations exceeding 0.99 considered evidence of salt-bridging (Alschuler et al., 2014). For channel pairs exceeding this threshold, the channel closest to midline was discarded to minimize signal redundancy while retaining the overall distribution of activity across the scalp. Visual inspection was then performed on all remaining channels, with channels judged to be noisy on the presence of high frequency noise or large signal discontinuities excluded from further processing stages. Five second epochs ranging from −3000 ms to +2000 ms around time zero were then extracted from the continuous data file, yielding a single epoched file containing 240 trials. This data file was then divided into three smaller files according to the experimental conditions (Control, Same, Different) containing 80 trials each. Individual trials had to meet two sets of criteria to be retained in the analysis. First, Same and Different trials were discarded if they were not discriminated correctly, with the number of trials retained at this step (from the original 80 per condition) being; Same [mean = 79.05; sd = 1.33] and Different [mean = 78.34; sd = 4.04]. Second, retained trials containing gross artifact (exceeding 200 µV), or in which the response latency exceeded 2000 ms were discarded. The mean number of trials retained after satisfying both criteria were; Control [mean = 64.98; sd = 8.41], Same [mean = 60.88; sd = 7.85], and Different [mean = 62.59; sd = 7.01], with all remaining trials submitted to further processing.

2.6.2. Independent component analysis
Prior to ICA training, pre-processed data files for each subject were concatenated so that ICA decomposition would generate a single set of channel weights uniform across conditions. This consistency of channel weights is essential to allow comparison of component activations across conditions for the purpose of hypothesis testing. Data were decorrelated with the extended Infomax algorithm (Lee et al., 1999) and then submitted to ICA training with the extended “runica” algorithm with an initial learning weight of $10^{-3}$ and a stopping weight of $10^{-7}$. The number of components resulting from ICA decomposition is determined by the number of channels submitted, with a maximum of 66 components (68 recording channels – 2 reference channels) per subject. However, as the number of channels excluded during pre-processing differed across subjects, the number of resulting components per subject was variable. Following decomposition, inverse weight matrices ($W^{-1}$) were projected back onto the recording montage to yield scalp topographies for each component, corresponding to coarse estimates of scalp distribution.

2.6.3. Dipole localization
Source localization was performed with the DIPFIT 2.3 toolbox (Oostenveld and Oostendorp, 2002) to generate equivalent current dipole (ECD) models (corresponding to point source estimates) for all independent components. Subject-specific channel locations were referenced to the 10–20 montage (Jasper, 1958), then warped to the BESA (i.e., spherical) heal model. Channel warping retains the relative configuration of recording channels on the scalp while minimizing the mean distance between the digitized locations and the 10–20 montage. Individual channel locations were unavailable for four subjects due to an equipment error, and standardized channel locations were used for this subset of participants. Automated multi-step (i.e., coarse, then fine) fitting to the head model resulted in dipole models for each of the 2205 components, representing physiologically plausible solutions to the inverse problem. Dipole models were then validated by projecting them onto the recording montage (Delorme et al., 2012), with the degree of mismatch between this projection and the original scalp-recorded signal (residual variance) considered a measure of the “goodness of fit” for each ECD model.

2.6.4. Study module
All 126 condition datasets (42 subjects x 3 conditions) were loaded into EEGLAB’s STUDY module to allow for the comparison of component activations across conditions and subjects. Components with a dipole model that could not be localized within the cortical volume or with residual variance exceeding 20% were excluded from group analyses. A residual variance threshold of 20% was selected as higher levels are likely indicative of either artifact or noise.

2.6.5. Principal Component Analysis clustering
Component clustering was performed using the K-means algorithm (Teknomo, 2006) in the Matlab statistical toolbox to group components based on similarities in ECD localization, spectra, and scalp maps. Components were allocated to 43 clusters, from which bilateral auditory alpha clusters were identified. While final allocation to clusters was based primarily on the results of PCA, clusters were visually inspected to ensure that all components met inclusion criteria and that no components meeting inclusion criteria had been omitted. Inclusion criteria for auditory alpha clusters were a characteristic alpha spectrum, ECD localization to STG or MTG (Jenson et al., 2015), and residual variance...
Components deemed to be mis-allocated during visual inspection were re-allocated to correct clusters. Any participant for whom PCA (and subsequent visual inspection) allocated a component that met inclusion criteria to either left or right auditory clusters was determined to have ‘contributed’ to that cluster.

2.6. Source localization

Following final allocation to component clusters, bilateral auditory alpha clusters were localized via ECD. The ECD cluster localization is the arithmetic mean of \((x, y, z)\) Talairach coordinates of all contributing components. These Euclidean coordinates were then submitted to the Talairach Client for mapping onto anatomic locations, yielding likely estimates of Brodmann’s areas and cortical locations.

2.6.7. Event Related Spectral Perturbations

ERSPs were employed to evaluate fluctuations in spectral power between 7 and 30 Hz (normalized dB units) over the time course of speech discrimination events. Signals were decomposed with a family of Morlet wavelets with a starting width of 3 cycles at 7 Hz and an expansion factor of 0.5. Spectral fluctuations were compared against a surrogate distribution generated from 200 randomly sampled time points in the inter-trial interval (i.e., \(-3000 \text{ to } -2000 \text{ ms}\)). All single trial data were submitted to time-frequency decomposition, with individual changes over time computed with a bootstrap resampling method (\(p < .05\), uncorrected).

Condition differences were evaluated with permutation statistics (2000 permutations) paired with cluster-based corrections for multiple comparisons (Maris and Oostenveld, 2007). A pair of 1 × 3 Repeated Measures ANOVAs (one per hemisphere) were performed to evaluate the presence of an omnibus effect, with paired \(t\)-tests to decompose the omnibus effects.

3. Results

Data from one participant was removed the study for failure to follow instructions as she indicated that she used a single hand for all button-press responses. Consequently, the results reported herein are based on data from the 41 remaining subjects.

3.1. Cluster characteristics

Of the 41 participants whose neural data were submitted for analysis, 35 contributed to auditory clusters, with contribution determined by the allocation of a component meeting cluster inclusion criteria to left or right auditory clusters by PCA or subsequent visual inspection. Specifically, 22 subjects contributed to both clusters, while 7 contributed only to the left cluster and 6 contributed only to the right cluster. This differential contribution to clusters is commonly reported in the literature (Zhu et al., 2020; Van Der Cruysen et al., 2021) and is not considered problematic for statistical testing given the similar number of components allocated to left and right auditory clusters (29 vs 28, respectively) in the current study. The left cluster was localized to Talairach \([-51, -37, 4]\) in the posterior middle temporal gyrus (BA-21) with a residual variance of 5.5%, and the right cluster was localized to Talairach \([51, -41, 7]\) in the posterior middle temporal gyrus (BA-21) with a residual variance of 4.08%. Fig. 2 displays the distribution of components allocated to left and right auditory clusters, respectively. Table 2 displays the contribution of participants to bilateral auditory clusters.

3.2. Condition accuracy

Subjects contributing to auditory clusters discriminated the syllables with a high degree of accuracy in both Same [mean = 98.8%, SD = 1.3] and Different [mean = 97.9%, SD = 4.0] conditions. A generalized linear mixed model [fixed effect = condition; random effect = subjects] with a gamma distribution and a log-link function implemented in SPSS (version 28.0) revealed no accuracy differences across discrimination conditions for contributing subjects \([F(1,68) = 0.37; p = .545]\), suggesting that the neural differences observed in the current study are not driven by task difficulty.

Fig. 2. Auditory cluster characteristics. The top row displays data from the left hemisphere auditory cluster and the bottom row displays data from the right hemisphere auditory cluster. A) Auditory spectra. B) Mean scalp distribution for components contributing to auditory clusters. C) Dipole model localization for component cluster members. D) Dipole density function for contributing components.
Multiple comparisons revealed significant differences from ~200 to 850 ms in the low alpha band (cluster-corrected ANOVA employing permutation statistics with cluster corrections for the remainder of the trial epoch in both discrimination conditions. A 1 × 3 ANOVA employing permutation statistics with cluster corrections for multiple comparisons revealed significant differences from ~200 to 950 ms in the low alpha band [cluster-corrected \( p = .0007 \)] and from ~900 ms to the end of the trial epoch in the higher alpha band spreading across alpha and beta bands [cluster-corrected \( p = .0005 \)]. Given the presence of significant activations when compared to the control condition bilaterally, it was possible to test experimental hypotheses regarding the effect of Same and Different trials.

### 3.4.2. Same/different comparison

A paired t-test with cluster-based corrections for multiple comparisons revealed significant left hemisphere differences between Same and Different trials from ~800 to 1200 ms in the alpha band [cluster-corrected \( p = .0195 \) spanning 7–14 Hz. A similar comparison between Same and Different trials in the right hemisphere did not reveal significant differences at any time/frequency point. Bilateral ERSP results for the Same/Different contrast are shown in Fig. 4.

### 4. Discussion

In the current study we employed independent component analysis of EEG data to identify bilateral component clusters with alpha spectral signatures over posterior temporal regions. While previous investigations of the auditory alpha rhythm have often reported source localizations to pSTG (Weisz et al., 2011; Muller and Weisz, 2012; Bowers et al., 2014; Jenson et al., 2015), the alpha clusters in the current study localized slightly more inferiorly to pMTG bilaterally. It should be noted, however, that a number of studies have also reported auditory source localizations to pMTG (Christoffels et al., 2007; Herman et al., 2013; Jenson et al., 2015), consistent with the findings of the current study. Furthermore, the Euclidean distance between alpha cluster source localizations in the current study and pSTG was ~5 mm bilaterally, within the typical margin of error for EEG source localization (Bradley et al., 2016). ERSP decompositions of alpha clusters further revealed robust fluctuations of activity over the time course of perceptual events, suggesting a response profile sensitive to auditory stimulation in line with previous descriptions of the auditory alpha rhythm (Tiihonen et al., 1991; Lehtela et al., 1997; Weisz et al., 2011; Muller and Weisz, 2012). In light of these factors, we are confident that identified bilateral alpha clusters represent indices of posterior DS (i.e., auditory) activity, such that it was possible to test experimental hypotheses regarding auditory dynamics during speech discrimination events.

Across all discrimination conditions and in both hemispheres, auditory activity was characterized by low alpha ERD during the peri and early post-stimulus period, followed by high alpha ERS in the late post-stimulus period. These findings are consistent with the results of Jenson et al. (2015) who interpreted peri-stimulus auditory alpha ERD through the framework of Analysis by Synthesis as evidence of coarse stimulus processing to relay an inverse model to anterior DS for mapping onto articulatory hypotheses and validation against the full signal. This
interpretation was supported by the presence of concurrent alpha and beta ERD over anterior DS (i.e., mu rhythm) during the peri-stimulus window, suggesting reciprocal interactions between motor and auditory regions to facilitate stimulus processing (Liebenthal and Möttönen, 2018). We propose that the presence of a similar pattern of auditory ERD in the current study with concurrent anterior DS alpha and beta ERD (Jenson and Saltuklaroglu, 2021) may be tentatively interpreted as evidence of stimulus decoding through Analysis by Synthesis. This interpretation is consistent with the assertions of Wilsch and Obleser (2016), who propose that auditory sensory memory operates at the pre-categorical level while auditory working memory operates at the post-categorical (i.e., phonological) level, and we suggest that Analysis by Synthesis supports this transition from sensory to working memory during encoding. Specifically, the cessation of ERD may reflect hypothesis confirmation resulting in stimulus identification and encoding to working memory with the emergence of auditory alpha ERS reflecting the engagement of a subsequent stage of the processing hierarchy.

In line with notions of alpha ERS reflecting a net inhibition of cortical regions (Neuper et al., 2006; Jensen and Mazaheri, 2010), it may be proposed that auditory regions are functionally deactivated during the
late post-stimulus window of the current study, and it is therefore essential to consider the mechanism by which this deactivation may occur. Sensorimotor activity in the post-stimulus window of discrimination tasks is commonly interpreted as covert articulatory rehearsal (Jenson et al., 2014a, 2019b; Saltuklaroglu et al., 2018), and it is thought to give rise to the alpha ERS observed in the current study through the delivery of forward models. Specifically, covert rehearsal is instantiated in the paired execution of forward and inverse models (Pickering and Garrod, 2013) to refresh working memory contents, with the delivery of forward models exerting a net inhibitory effect on auditory regions (Kawato and Wolpert, 1998; Guenther, 2016). While originally considered within the framework of overt production (Numminen et al., 1999; Curio et al., 2005; Houde et al., 2002), speech-induced suppression of SIS of auditory regions has also been observed in covert (i.e., imagined) speech tasks (Kauramaki et al., 2010; Balk et al., 2013). In further support of this interpretation, oscillatory coherence between motor and auditory regions increases during the post-stimulus window (Bowers et al., 2019), consistent with notions of forward model delivery. Thus, overall ERSP patterns are compatible with interpretations of working memory encoding through Analysis by Synthesis and maintenance through covert articulatory rehearsal. However, to evaluate experimental hypotheses it remains essential to consider how these processes are differentially influenced by Same and Different trials.

The first hypothesis that Different trials would exhibit a protracted time course of low alpha ERD compared to Same trials was well-supported by data from the left hemisphere, with ERD persisting ~400 ms longer in Different trials. In line with interpretations through Analysis by Synthesis of low alpha ERD representing the encoding phase (Wianda and Ross, 2019), we suggest that Different trials elicit a prolonged encoding phase compared to Same trials. A prolonged encoding phase of working memory has been observed in anterior sensorimotor regions during the discrimination of degraded stimuli (Jenson et al., 2019b), and delayed processing has also been reported to noise-masked stimuli in auditory regions (Koener and Zhang, 2015; Benitez-Barrera et al., 2021). However, as the stimuli employed in the current study were undegraded and presented in a clear background, it is necessary to consider alternative explanations for a prolonged encoding phase. Jenson and Saltuklaroglu (2021) proposed that in discrimination trials, an articulatory representation of the first syllable is used as a predictive hypothesis for the second, leading to an error signal in Different trials only which requires hypothesis revision through internal modeling across the sensorimotor network, a process that would be expected to elicit a protracted phonological encoding stage. Under this interpretation, we propose that the prolonged low alpha ERD observed in Different trials indexes deeper processing necessary to resolve detected error signals, occurring in sustained communication between anterior motor and posterior auditory regions. Considered within the framework of error detection, this may be interpreted to suggest that at least a portion of auditory error detection processes in speech discrimination occur during the encoding phase.

It should also be noted that the left hemisphere differences span low and high alpha ranges, with both low ERD and high ERS co-occurring from ~900 to 1150 ms. Given the distinct spectral signatures of encoding and maintenance stages proposed by Wianda and Ross (2019), results may be interpreted to suggest a previously unobserved temporal overlap of encoding and maintenance stages. This could represent an effect of low alpha ERS (considered recently given the high degree of accuracy across discrimination tasks) or may alternatively suggest that encoding and maintenance stages are not always discrete. While overlap between these two stages has previously been reported in tasks employing multiple sensory modalities (Bashivan et al., 2014), to our knowledge this is the first study to report a potential overlap of encoding and maintenance stages within a single sensory modality. We suggest that the observed overlap of low ERD and high ERS may reflect the detection and resolution of a prediction error elicited by the use of the first syllable as a predictive template for the second. Since the timeline of high alpha ERS emergence was largely consistent between Same and Different trials, we argue that covert rehearsal mechanisms engage upon identification of the first syllable, potentially driven by the need to avoid degradation of the original sensory trace (Jacquemot and Scott, 2006; Wilsch and Obleser, 2016). In the event of a confirmed prediction, in which the second syllable matches the first, there is a seamless transition from encoding to maintenance stages. However, the violation of this prediction in Different trials leads to deeper processing, reflected in prolonged alpha ERD, with the resulting overlap of alpha ERD and ERS potentially reflecting the detection and resolution of an error. We propose that error detection processes engage during the encoding phase in all discrimination trials, while the overlap of ERD and ERS observed in the left hemisphere in Different trials reflects the detection and resolution of an error. However, this interpretation is made cautiously, and further work is necessary.

While broadly consistent with patterns observed in the left hemisphere, right hemisphere auditory alpha activity did not exhibit differences between Same and Different trials at any time-frequency point. It should be noted that a lack of right hemisphere differences is compatible with the results of Jenson and Saltuklaroglu (2021) who interpreted weaker right hemisphere mu ERD and stronger left hemisphere mu ERD in Different trials as evidence that in the case of detected mismatches, the right hemisphere relinquishes further processing to the speech-specialized left hemisphere. Such a notion aligns both with the well-established left hemisphere dominance for speech and language (Hickok and Poeppel, 2000, 2004, 2007; Specht, 2014) and proposals that the left hemisphere performs a more fine-grained level of analysis than the right hemisphere during speech processing (Poeppel, 2003; Ylinen et al., 2015). Observed results are consistent with the interpretation that only the left hemisphere participates in encoding stage error detection processes. However, because the stimulus set for the current study was purely speech-based and did not include non-speech stimuli, it is not possible to preclude the possibility of Analysis in Synthesis in non-speech auditory discrimination (Thornton et al., 2018). Additionally, in order to clearly demonstrate a left-hemispheric specialization for the error detection processes under investigation, it would be necessary to observe a differences in hemispheric activity (Nieuwenhuis et al., 2011), and results should be interpreted with caution.

It should also be considered that the lack of right hemisphere differences argues against more simplistic explanations of repetition priming in which recently presented items are processed more efficiently and quickly than novel items (Schacter and Buckner, 1998; Henson, 2003). Under priming-based accounts, the shorter duration of ERD in Same trials could be interpreted as evidence of repetition priming for matched syllable pairs. However, since differences are not observed in the right hemisphere and the timeline of right hemisphere ERD in both Same and Different trials aligns with the timeline of left hemisphere Same trials, such an interpretation is not tenable. Furthermore, oscillatory responses to repetition priming are typically associated with spectral power reductions (Tavabi et al., 2011) rather than the timeline differences observed herein. Consequently, we propose that data are consistent with interpretation as evidence of a protracted phonological encoding phase in Different trials necessary to resolve error signals arising from the use of the first syllable as a predictive template for the second.

The second hypothesis that the magnitude of post-stimulus high alpha ERS would differ between Same and Different trials was not supported by data from either hemisphere, suggesting that auditory contributions to working memory maintenance are not influenced by trial type. Specifically, though late alpha ERS appeared weaker in Different trials, there were no significant differences between conditions in the magnitude of alpha ERS. This result was surprising given the phonological similarity effect (Camos et al., 2013; Chow et al., 2016), in which memory performance decreases for phonologically similar items as compared to dissimilar items, suggesting elevated maintenance load.
While the phonological similarity effect is known to reverse in non-words such as the syllable stimuli employed in the current study (Karlsen et al., 2007), the absence of condition differences is still surprising. It should also be noted that anterior sensorimotor regions demonstrate somatotopic specificity (Pulvermüller et al., 2006; Skipper et al., 2007; Bartoli et al., 2016), such that forward models in Different trials integrate information from a wider region of sensorimotor cortex, and would be expected to exert a larger inhibitory effect on auditory regions. We tentatively propose that the lack of ERS differences may have resulted from the low maintenance load (n = 2 syllables) compared to the typical maintenance load (3–6 items) in studies that report the phonological similarity effect (Karlsen et al., 2007; Sweet et al., 2008; Chow et al., 2016). Additionally, since syllables within a pair only differed by the initial consonant, all pairs rhymed, which is also known to enhance memory performance and potentially lowered maintenance demands (Karlsen and Lian, 2005). Thus, while the presence of post-stimulus auditory alpha ERS is consistent with a general role for auditory regions in working memory maintenance via covert articulatory rehearsal (Jenson et al., 2015), findings do not support differential engagement of error detection processes during working memory maintenance on the basis of trial type.

### 5. General discussion

While we have described the manner in which auditory alpha oscillatory power unfolds over the time course of speech discrimination tasks with a specific focus on how they differ during the discrimination of Same and Different syllable pairs, it remains to formulate a comprehensive account of auditory error detection during discrimination tasks. Due to the transient nature of acoustic signals and the rapid decay of sensory traces (Wilsch and Oleser, 2016), stimuli must be encoded and maintained in working memory to enable successful discrimination performance. We propose that auditory regions engage in two distinct error detection processes during speech discrimination tasks which are differentially active during encoding and maintenance stages. In line with previous reports of mu rhythm dynamics in the anterior DS (Jenson et al., 2014a, 2019b), we argue that auditory activity during the encoding stage supports the extraction of an articulatory representation (Jacquemot and Scott, 2006) and proceeds through Analysis by Synthesis (Stevens and Halle, 1967). Specifically, a coarse sketch of incoming acoustic stimuli is relayed to anterior motor regions by an inverse model (Stevens and Halle, 1967; Skipper et al., 2007; Bever and Poeppel, 2010) for mapping onto an articulatory representation. The sensory representation of this articulatory hypothesis is then returned to auditory regions for comparison against the full signal, with any detected error signal being returned to anterior motor regions for hypothesis revision. Results of the current study are consistent with the notion that during syllable discrimination, task-specific demands (i.e., the need to compare the two syllables) lead to the articulatory hypothesis of the first syllable being used as a predictive template for the second, allowing for rapid detection of matched trials. In contrast, the use of this predictive template in Different trials leads to the detection of sensory mismatch (i.e., error signals) in auditory regions, which must be resolved through iterative hypothesis-test-refine loops before stimulus identification is complete. We propose that the temporal overlap between auditory alpha ERD and ERS in Different trials reflects the detection and resolution of a prediction error, with the protracted timeline of ERD being interpreted as evidence of the deeper processing necessary to resolve the detected error. Findings align with the notion that auditory error detection processes are sensitive to syllable pair concordance. A similar sensitivity to stimulus match/mismatch has been revealed in anterior sensorimotor regions (Jenson and Saltuklaroglu, 2021), and the results of the current study are consistent with the proposal that this sensitivity is mirrored in posterior auditory regions.

Since the magnitude of post-stimulus auditory alpha ERS did not differ between Same and Different trials, we propose that it reflects a more general process emerging across discrimination tasks. Specifically, given the similarity of this activity to that reported during overt speech production (Jenson et al., 2015), we suggest that findings may be interpreted as evidence of speech-induced-suppression (Greenlee et al., 2011; Chang et al., 2013). Since covert production is supported by the same internal modeling processes as overt production (Tian and Poeppel, 2016; Pickering and Garrod, 2013), forward model delivery during the maintenance stage leads to the attenuation of auditory responses similar to the pattern observed during overt production (Greenlee et al., 2011; Chang et al., 2013). Such an interpretation is consistent with accounts of sensorimotor coding in working memory (Wilson, 2001), contemporaneous anterior sensorimotor activity consistent with covert production (Jenson and Saltuklaroglu, 2021), and oscillatory load effects (Grimault et al., 2009, 2014) over auditory regions during working memory tasks. This may suggest that auditory activity during covert rehearsal supports the same error detection processes active during overt speech production. Thus, we propose that while auditory oscillations support error detection during both encoding and maintenance stages of working memory, they do so through distinct mechanisms, with encoding stage error detection mediated by stimulus concordance and maintenance stage error detection driven by task-specific maintenance demands (i.e., the need to retain stimuli prior to responding).

### 6. Limitations

While the results of the current study provide evidence consistent with the involvement of auditory regions in error detection processes during speech discrimination tasks, some limitations should be acknowledged. First, based on the use of an exclusively female cohort, the applicability of results to the wider population remains unclear. This is particularly salient since sex differences have been reported in anterior sensorimotor regions (Popovich et al., 2010; Kumari, 2011; Thornton et al., 2019), and the current study explores related processes in posterior sensorimotor (i.e., auditory) regions. Future work should explore whether sex differences are also present in auditory oscillations during speech discrimination. Second, it should be noted that not all participants contributed components to auditory clusters. A reduced proportion of contributing subjects is common in EEG studies (Nystrom, 2008; Bowers et al., 2013) and has been linked to the use of standard head models. While individualized channel locations were employed in the current study, warping of these channel locations to standard head models during the localization stage still yielded reduced participant contribution. However, the proportion of contributing subjects in the current study (85% overall; 71% left; 68% right) is on par with previous reports (Cuellar et al., 2016; Kittilstved et al., 2018; Bowers et al., 2019; Oliveira et al., 2021), and we are confident that it does not meaningfully impact the results.

### 7. Conclusions and future directions

The current study revealed robust fluctuations of auditory oscillations over the time course of speech discrimination tasks, with Different trials eliciting a protracted time course of low alpha ERS compared to Same trials. Findings are consistent with interpretations of auditory contributions to error detection in both encoding and maintenance stages of working memory. We suggest that auditory error detection during working memory encoding proceeds through Analysis by Synthesis, with the use of an articulatory representation of the first syllable as a predictive template for the second eliciting prediction errors that require resolution during Different trials only. We further propose that auditory error detection during working memory maintenance arises from forward model (i.e., efference copy) delivery during covert articulatory rehearsal to facilitate stimulus retention, akin to error detection processes during overt production. Thus while error detection processes are engaged across working memory stages, error detection during encoding is associated with stimulus-specific features (i.e., stimulus
concordance) while error detection during maintenance is associated with task demands (i.e., stimulus retention). In addition to supporting an active role for auditory regions in error detection in speech discrimination tasks, the results of the current study further demonstrate the sensitivity of time-frequency EEG analyses to the cortical dynamics supporting sensorimotor processing in general and error detection in particular.

**Credit author statement**

David Jenson: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Tim Saltuklaroglu: Conceptualization, Writing – original draft, Writing – review & editing.

**Interest statement**

The authors declare that they have no competing financial or non-financial interests.

**Funding statement**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Data availability**

The data for this project are freely available on the first author’s Harvard Dataverse profile at https://doi.org/10.7910/DVN/1HX6WJ.

**References**

Alho, J., Lin, F.H., Sato, M., Tiitinen, H., Sams, M., Jaaskelainen, I.P., 2014. Enhanced neural synchrony between left auditory and premotor cortex is associated with successful phonetic categorization. Front. Psychol. 5, 394.

Alho, K., Salonen, J., Rinne, T., Medvedev, S.V., Hugdahl, K., Hamalainen, H., 2014. Attention-related modulation of auditory-cortex responses to speech sounds during dichotic listening. Brain Res. 1644, 47–54.

Alschuler, D.M., Tenke, C.E., Bruder, G.E., Kayser, J., 2014. Identifying electrode agencies in the public, commercial, or not-for-profit sectors.

Buchsbaum, B.R., Driscoll, B.R., D’Esposito, M., 2019. A sensorimotor view of working memory. Cortex 112, 134–148.

Callan, D., Callan, A., Gamer, M., Sato, M.A., Kawato, M., 2010. Premotor cortex mediates perceptual performance. Neuroimage 51, 844–858.

Campana, V., Mora, G., Barrouillet, P., 2008. Phonological similarity effect in complex span task. J. Q. Exp. Psychol. 66, 1927–1950.

Chang, E.F., Niznik, C.A., Knight, R.T., Nagarajan, S., Hude, J.F., 2013. Human cortical sensorimotor network underlying feedback control of vocal pitch. Proc. Natl. Acad. Sci. U. S. A. 110, 2655–2658.

Chiappe, P., Chiappe, D.L., Siegel, L.S., 2001. Speech perception, lexicality, and writing skill. J. Exp. Child Psychol. 80, 58–74.

Chow, M., M.A., Ramara, B.N., Cogan, C.A., 2016. Phonological similarity in working memory tasks. Mem. Cogn. 44, 937–949.

Christoffers, I.K., Formisano, E., Schiller, N.O., 2007. Neural correlates of verbal feedback processing: an fMRI study employing overt speech. Hum. Brain Mapp. 28, 868–879.

Cuellar, M., Harkrider, A.W., Jenson, D., Thornton, D., Bowers, A., Saltuklaroglu, T., 2016. Time-frequency analysis of the EEG mu rhythm as a measure of sensorimotor integration in the later stages of swallowing. Clin. Neurophysiol. 127, 2625–2635.

Curio, G., Neuloh, G., Nuemann, J., Joumaki, V., Hari, R., 2000. Speaking modifies voice-evoked activity in the human auditory cortex. Hum. Brain Mapp. 9, 183–191.

Delorme, A., Palmer, J., Onten, I., Oostenveld, R., Makeig, S., 2012. Independent EEG sources are dipolar. PLoS One 7, e30135.

Deng, Y., Gao, R., Ding, G., Peng, D., 2012. Top-down modulations from dorsal stream in lexical recognition: an effective connectivity fMRI study. PLoS One 7, e33377.

Greenlee, J.D., Jackson, A.W., Chen, F., Larson, C.R., Oya, H., Kawasaki, H., Chen, H., Howard, M., 2011. Human auditory cortical activation during self-vocalization. PLoS One 6, e16474.

Grimaldi, S., Lefebvre, C., Vachon, F., Peretz, I., Zatorre, R., Robitaille, N., Jolicour, P., 2009. Lead-dependent brain activity related to auditory short-term memory for pitch: magnetoencephalography and fMRI. Ann. N. Y. Acad. Sci. 1169, 273–277.

Grimaldi, S., Nolden, S., Lefebvre, C., Vachon, F., Hyde, K., Peretz, I., Zatorre, R., Robitaille, N., Jolicour, P., 2014. Brain activity is related to individual differences in the number of items stored in auditory short-term memory for pitch: evidence from magnetoencephalography. Neuroimage 94, 96–106.

Guenther, F.H., 2016. Neural Control of Speech. Mit Press.

Heinks-Maldonado, T.H., Mathalon, D.H., Gray, M., Ford, J.M., 2005. Fine-tuning of auditory cortex during speech production. Psychophysiology 42, 180–190.

Henson, R.N., 2003. Neuroimaging studies of priming. Prog. Neurobiol. 70, 53–81.

Herman, A.B., Houde, J.F., Vinogradov, S., Nagarajan, S.S., 2013. Parsing the phonological loop: activation timing in the dorsal speech stream determines accuracy in speech reproduction. J. Neurosci. 33, 5439–5453.

Hickok, G., Poeppel, D., 2000. Towards a functional neuroanatomy of speech perception. Trends Cognit. Sci. 4, 131–138.

Hickok, G., Poeppel, D., 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92, 67–99.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402.

Houde, J.F., Chang, E.F., 2015. The cortical computations underlying feedback control in vocal production. Curr. Opin. Neurobiol. 33, 174–181.

Houde, J.F., Nagarajan, S.S., Sekihara, K., Merzenich, M.M., 2002. Modulation of the thalamocortical circuit during speech: an MEG study. J. Cognit. Neurosci. 14, 1125–1138.

Jacqueton, C., Scott, S.K., 2006. What is the relationship between phonological short-term memory and speech processing? Trends Cognit. Sci. 10, 480–486.

Jasper, H.H., 1958. The ten electrode system of the international federation. Electroencephalogr. Clin. Neurophysiol. 10, 371–379.

Jenson, D., Reilly, K.J., Harkrider, A.W., Thornton, D., Bowers, A., Saltuklaroglu, T., 2019b. Influences of sensorimotor oscillations on sensorimotor integration in the later stages of swallowing. Clin. Neurophysiol. 127, 2625–2635.

Jenson, D., Thornton, D., Harkrider, A.W., Saltuklaroglu, T., 2014b. Speech perception, and the sensorimotor mu rhythm. In: Proceedings Of the 2014 Biomedical Sciences And Engineering Conference, pp. 1–4.
