Journal of Neurolinguistics 60 (2021) 101006

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

Research paper

Mu rhythm dynamics suggest automatic activation of motor and premotor brain regions during speech processing

Daniela Santos Oliveira, Tim Saltuklaroglu, David Thornton, David Jenson, Ashley W. Harkrider, M. Blake Rafferty, Devin M. Casenhiser [*]

University of Tennessee Health Science Center, USA

A R T I C L E I N F O

Keywords:
Speech perception
Electroencephalography
Dual stream model
Mu rhythm

A B S T R A C T

Results are mixed regarding the question of whether listening to speech activates automatic speech specific neural processes, or whether such activity is due to general cognitive mechanisms such as attention and memory. We suggest that analyzing the dynamics of mu rhythms over the course of a trial using event-related spectral perturbations (ERSPs) is useful for disentangling the competing hypotheses. Our findings suggest that there is automatic activation of motor and premotor areas of the brain that occurs while passively listening to speech, and that this activation is not present while listening to noise. Moreover, this activation is more strongly present in the low alpha frequencies, which are associated with attention, when participants engage in active discrimination of stimuli, and late in the trial prior to a button press.

1. Introduction

For years, speech scientists have investigated how neural activity in premotor and motor regions relates to speech perception. Much of the earlier work was used to support or refute strong theoretical orientations of the critical role for the motor system in speech perception (e.g., Motor Theory, Direct Realist Theory). However, since the identification of a sensorimotor dorsal stream that includes premotor and motor regions (e.g., Hickok & Poeppel, 2004), the scientific focus has shifted towards understanding why and in what context sensorimotor activity is observed during speech perception. This research largely falls into two groups. One suggesting that the observed activity is reflective of general cognitive mechanisms (e.g., attention, prediction, working memory), and one suggesting that activity is specific to the speech-perception process. Below, we briefly review the research cited in favor of each position, and then present our own research suggesting that examining the timecourse of neural oscillations enables researchers to better separate changes in neural activity that are due to general cognitive mechanisms from those that are due to automatic speech processing.

The notion that activation in motor and premotor regions increases as the general cognitive demands of the perception task increase is fairly well established. For example, several studies have demonstrated that active speech discrimination tasks are associated with greater motor and/or premotor activity in comparison to listening to speech passively (Alho et al., 2012, 2014; Peschke et al., 2012; Wostmann et al., 2017). Similarly, within active discrimination paradigms, increasing the difficulty of the discrimination task also increases motor and premotor neural activation. Processing non-native speech phonemes, for example, evidences greater activation than does processing native phonemes (Callan et al., 2004, 2014; Wilson & Iacoboni, 2006). Greater activation is found for speech embedded in noise as compared with clear speech (Alho et al., 2012; Binder et al., 2004; Scott et al., 2004; Zekveld et al., 2006), and for

---

[*] Corresponding author. 600 Henley Street, Suite 119, Department of Audiology & Speech Pathology, Knoxville, TN, 37996, USA.
E-mail address: dcasenhi@uthsc.edu (D.M. Casenhiser).

https://doi.org/10.1016/j.jneuroling.2021.101006
Received 27 October 2020; Received in revised form 13 April 2021; Accepted 10 May 2021
Available online 28 May 2021
0911-6044/© 2021 Elsevier Ltd. All rights reserved.

speech that is acoustically masked or distorted (Adank & Devlin, 2010; Du et al., 2014; Osnes et al., 2011; Poldrack, 2000). Thus, many researchers argue that observed differences in activation index involvement of general cognitive mechanisms, such as those related to the greater attention required when performing a difficult task. Dorsal stream activity may contribute to attentional mechanisms by providing early motor-sensory transformations that help predict forthcoming stimuli and filtering signal from noise in degraded listening conditions. Similarly, actively discriminating or categorizing syllables requires that after hearing sounds, the listener covertly replay the syllables within working memory (Jacquemot & Scott, 2006); a process that recruits the dorsal stream for instantiating sensorimotor transformations (Jenson et al., 2019). As such, these cognitive-based dorsal stream activations may precede (e.g., during prediction) or follow (e.g., via activation of working memory buffer) speech perception and cannot be interpreted as being representative of direct sensory-motor activations that may occur at the time of speech perception (Binder et al., 2004; Buchsbaum & D'Esposito, 2008; Krawczyk, 2002).

While it is difficult to ignore these findings, some argue that at least some motor and premotor activity is integral to the perceptual process itself (Alho et al., 2014; Cuellar et al., 2012; Iacoboni, 2008; Jenson et al., 2014; Pulvermuller & Fadiga, 2010; Skipper et al., 2017; Wilson et al., 2004). Several findings are commonly cited in support of this. For example, passively listening to speech has been found to elicit motor and premotor activity (Pulvermuller et al., 2006; Wilson & Iacoboni, 2006), and work by Liebenthal et al. (2013) demonstrated that there is sensorimotor activity during passive listening to speech in a duplex perception task that was not present during passively listening to non-speech. This finding indicates a process specific to speech processing rather than general auditory processing. However, a number of negative findings offer contradictory evidence (Belin et al., 2000, 2002). Furthermore, a review of findings by Scott, McGettigan, and Eisner (2009) concluded that whole brain imaging studies of speech perception rarely demonstrate motor activity, and a recent report by Schomers and Pulvermuller (2016) suggest that when functional magnetic resonance imaging (fMRI) is used to detect motor activity, positive findings may be masked by the presence of scanner noise.

That the debate continues is hardly surprising given the complexity of the signal being processed and numerous possible explanations of findings that are often driven by theoretical orientation. Part of the issue lies in the nature of the speech signal itself (Heald & Nusbaum, 2014). Speech and the neuronal activations that accompany speech are transient signals. Researchers attempting to study such signals are required to isolate and correlate location and morphology of the neural activations with timing of the speech signals. As a result, methods that lack temporal precision, or that average over the course of a trial or trials may fail to capture important changes in the neurological signal as it unfolds over time. Activations that occur prior to, during and following the presentation of speech stimuli are likely to index different processes, thus averaging across a trial may obscure the changes. Consequently, between-condition differences in neural activity may be due to consistent changes across the entire trial, or due to differences occurring only at distinct points in a trial.

In this study we directly address this need by analyzing the event-related spectral perturbations (ERSPs) of neural oscillations from electroencephalography (EEG). This method is well-suited to meeting the challenges of the temporal and spatially sensitive nature of measuring neural activation during speech processing. Using Independent Components Analysis (ICA) together with source modelling we are able to identify spatially fixed and temporally independent sources of neural activity that are linearly mixed across several sensors thereby increasing the spatial accuracy of measurements (Delorme & Makeig, 2004; Makeig et al., 2004; Onton et al., 2006). While the spatial accuracy does not match that of fMRI measurements, it is a great improvement over the accuracy common to traditional ERP studies, and is aimed at minimizing the volume conduction errors that can be made when using scalp electrode positions as source estimates. Moreover, like ERP research, ERSPs afford the researcher millisecond accuracy – accuracy not achievable with standard fRMI methods, and unlike fMRI, EEG offers the advantage of a quiet testing environment, free from potentially confounding influences of scanner noise.

Finally, a supplemental advantage of applying ICA to EEG data is that the resulting neural sources can be decomposed across time and frequency through the analysis of ERSPs. This is a critical advantage since the frequency, morphology and spectral power of neural oscillations, together with their source location, are thought to differentiate neurological functions. Decreases in spectral power reflect desynchronization of cortical neural network oscillations (called event-related desynchronization or ERD), while increases in spectral power reflect event-related synchronization (ERS) of cortical networks (Pfurtscheller & Lopes da Silva, 1999). ERD in the alpha band (~8–13 Hz) is thought to be associated with attentional mechanisms, and has been found to index an increase in attentional or cognitive demands (Klimesch, 2012). Weisz and colleagues' (2011) study indicated that alpha ERD across a wide network of brain regions (prefrontal, temporal, parietal) was related both to reductions in speech intelligibility and working memory set size, supporting its sensitivity to both speech perception and working memory. In speech production, alpha ERD may provide an index of sensory feedback in audio-vocal monitoring (Jenson et al., 2014, 2015; Tamura et al., 2012). Alpha ERS, on the other hand, has been interpreted as having inhibitory or gating effects which can be very important to optimize energy demands and control over excitatory processes (Klimesch, 2012; Klimesch et al., 1996). Research has additionally connected ERS with gating of noise in speech discrimination tasks (Jenson et al., 2014). Beta ERD (~13–30 Hz), on the other hand, is often associated with predictive 'top-down' coding (hypothesis testing) for sensory analysis (Buschman & Miller, 2007; Kilavik et al., 2013; Siegel et al., 2012).

In the present study, we analyze a particular wave morphology called mu waves, which are defined by a pattern of co-occurring alpha and beta oscillations over motor and premotor regions (Hari, 2006; Pineda, 2005). Recent evidence suggests that mu alpha and beta oscillations can be reliable indices of sensory and motor processes occurring during speech perception (Jenson et al., 2014, 2015). Since the mu rhythm is localized to primary motor/PMC regions, and since the PMC is localized within the auditory dorsal stream (Hickok & Poeppel, 2004), mu oscillations are thought to reflect maintenance of bidirectional communication between auditory and somatosensory regions via the arcuate and longitudinal fasciculi. As such, investigation of mu alpha and beta oscillations present a unique system for the analysis of speech processing. In fact, a recent body of research analyzing the time course of mu rhythm activity in active syllable discrimination tasks (Bowers et al., 2013a, 2014a; Jenson et al., 2014, 2015; Saltuklaroglu et al., 2017; Thornton et al., 2018, 2019) has shown oscillatory changes beginning prior to syllable presentation and continuing well after the syllables are heard. These oscillatory changes were interpreted in light of attentional and working memory demands that surround perception. In fact, the most robust finding across these studies is strong mu-alpha and mu-beta ERD following syllable presentation. This pattern of activity is interpreted as arising from sensorimotor activity from within the working memory buffer since it increases with task demands that appear to rely on replay in a working memory buffer such as post-stimulus information processing and retrieval (Klimesch, 2012; Klimesch et al., 2007), is present during pre-speech planning, changing to ERS at speech onset ((Jenson et al., 2015), and is present during covert speech production (Jenson et al., 2014) among other reasons (for a review, see Saltuklaroglu et al., 2018).

In this study, we employ ERSP analysis of mu waves during a passive speech perception task to shed light on the nature of the sensorimotor activity as it unfolds over time while participants passively listen to speech. Only one study has investigated mu rhythms in this capacity. Bowers et al. (2013b) failed to detect significant mu oscillatory changes when passively listening to syllables (compared to control condition). However, Bowers et al. embedded their stimuli in background noise, raising the question of whether the noisy background reduced sensitivity and rendered a Type II error as has been suggested to arise from fMRI scanner noise (Schomers & Pulvermuller, 2016). As with previous studies, we compare an active speech discrimination task to a passive listening task in order to control for task-dependent processes. Based on previous findings, we hypothesize that the active task will show more robust overall mu oscillatory activity across trials than the passive task, consistent with increased cognitive demands. However, we expect that this difference will be found late in the trial, when active (i.e., not automatic) processing related to the discrimination task relies on working memory. In contrast, if passive listening in quiet background is sufficient to elicit sensorimotor activity, it is most likely to be found as the syllables are being presented and auditory-motor transformations are occurring. Using the current measures, this would be evident by mu alpha ERD during syllable presentation. Given the relative dearth of additional cognitive demands, other surrounding (early or late) mu activity is expected to be relatively absent in passive listening. The results follow largely in line with our predictions and reveal new temporally sensitive insights on the nature of sensorimotor activity in speech perception and its ties to general cognitive activity.

2. Method

2.1. Participants

24 female monolingual American English native speakers with a mean age of 25.3 (range 20–69) participated in this study. This is a sample of convenience, but also serves to control for recently reported sex differences in mu activity during speech perception (Thornton et al., 2019). As assessed by the Edinburgh Handedness inventory, 23 participants were right-handed and 1 was ambidextrous. No participants reported having a history of cognitive, language or hearing impairments. Prior to the experiment, informed consent approved by the University of Tennessee Health Science Center's Institutional Review Board was obtained for all participants.

2.2. Materials

The stimuli were monosyllabic pairs of English syllables composed of /f/ and /v/, followed by the central vowel /ʌ/, forming the monosyllables /fʌ/ and /vʌ/. The stimuli were all produced and recorded by a phonetically trained American English native speaker. The speaker produced thirty samples of each monosyllable at a regular speech rate and level, and maintaining a similar intonation. The best exemplars of each monosyllable, judged by similarities in overall duration, intonation curve, and vowel quality, were then selected to be used in the categorical discrimination task. The best exemplars were so judged by two linguistically trained listeners. The stimuli's intensity and duration were then normalized using PRAAT (Boersma & Weenink, 2017). The intensity of each monosyllable was manipulated to be approximately 72 dB, and the duration was set to be exactly 600 ms for each monosyllable. Pairs of syllables were then created and a 250 ms of silence was inserted as an interstimulus interval (ISI). A silence of 550 ms was also inserted after each pair of monosyllables, so that the total duration of each trial would be exactly 2000 ms (see Fig. 1).

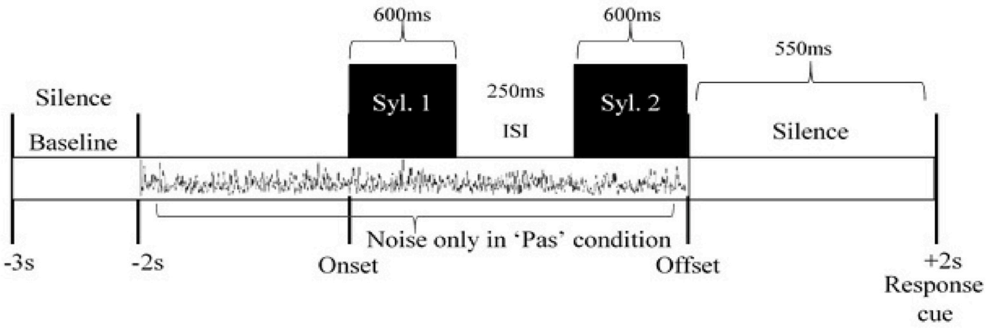

Fig. 1. Trial timeline.

2.3. Procedure

The experiment was conducted in an electronically and magnetically shielded, double-walled, sound-treated booth. Participants sat in a comfortable reclining armchair with their heads and necks well supported. Participants heard binaurally presented audio from three different conditions: passive listening to white noise (control condition), passively listening to pairs of syllables, and active discrimination of syllable pairs. Noise was used during the control condition because mu activity can be influenced by attentional mechanisms (Anderson et al., 2011; Jones et al., 2010), so even the low levels of attention required during the passive listening task may influence attention. Moreover, previous work on mu rhythms has indicated that mu suppresses to perceived biological (e.g., reproducible) stimuli but not non-biological stimuli. Listening to noise approximates the same amount of attention required during passive listening to speech, but should not induce mu activity related to processing speech (or biological stimuli).

The syllable stimuli were presented in four possible combinations, forming the following contrasts: /fʌ/–/vʌ/; /vʌ/–/fʌ/; /fʌ/–/fʌ/; /vʌ/–/vʌ/. Conditions were presented in two blocks consisting of 40 trials each in which there were 20 matching and 20 mismatched pairs to control for selection bias (Venezia et al., 2012), resulting in a total of 240 trials.

Stimuli were presented using Neuroscan's Stim 2 presentation software on a PC computer. Participants were always presented with the passive listening condition first, so that no tendency to covertly categorize the sounds would carry over. In the active task, participants were instructed to actively discriminate the syllable pairs by pressing a button. It has been shown that premotor planning occurs in repeated button press movements 1 s prior to muscle contraction (Graimann & Pfurtscheller, 2006; Hari, 2006) and sensorimotor suppression (ERD) peaks shortly after (200 ms) (Makeig et al., 2004). In order to control for the possibility that preparation for the response might confound motor activity related to stimulus processing, participants' manual response was cued by a 100 ms, 1000Hz, sawtooth wave tone, which appeared 2000 ms after stimulus onset. The passive listening condition was also followed by button press, in order to control for the required button press in the active conditions (see below). The button press in the passive listening condition also serves to ensure participants are paying attention to the stimuli presented. Furthermore, button press order (right or left hand) was counterbalanced across participants and experimental conditions. The presence of a button press in all conditions ensures that any differential activity is not a result of motor activity related to button press motion, but instead a result of the demands of the different conditions.

The passive condition was presented first, followed by the active discrimination condition and finally the noise condition. All stimuli were presented at an absolute intensity of 72 dB. An example timeline is displayed in Fig. 1. Presentation of audio stimuli was routed through insert ER-1-14A ear tips.

2.4. EEG data acquisition

Methods applied to EEG data acquisition follow methods employed in Bowers et al. (2013b) and Jenson et al. (2014). Electroencephalogram (EEG) data were acquired and analyzed using a Neuroscan 4.3.3 system. The total duration of EEG data collection was approximately 60 min. Sixty-eight electrode channels were used to acquire EEG data based on the extended international 10–20 method of electrode placement (Klem et al., 1999) using an unlinked, sintered NeuroScan Quik Cap. Electrodes were referenced to a midline electrode anterior to FZ1 during recording and were referenced to the common linked left (M1) and right (M2) mastoids during preprocessing. The electro-oculogram (EOG) was recorded by placing electrodes on the left superior orbit and the left inferior orbit (VEOG) as well as the lateral and medial canthi of the left eye (HEOG) to monitor vertical and horizontal eye movements, respectively.

EEG data were collected using Compumedics NeuroScan 4.3.3 software and Synamps 2 system. The raw EEG data were filtered (0.15–100Hz) and digitized via a 24-bit analog-to-digital converter at a sampling rate of 500Hz. Data were time-locked to the onset of acoustic stimuli (time = 0), while time zero represents the midpoint of white noise in the control condition.

2.5. Data processing

2.5.1. Individual processing

Following data collection, the recorded EEG data were processed using EEGLAB 12 software (Delorme & Makeig, 2004). Individual participants' data files were appended to make a single 80 trial data set for each condition. Files were then down sampled to 256Hz, in order to decrease computational requirements for ICA processing. Data were then epoched into 5000 ms segments with individual epochs spanning from −3000 to +2000 ms around time zero. A bandpass filter between 3 and 34Hz was applied to ensure that alpha and beta could be identified while filtering muscle movement and line noise from surrounding frequencies. Data were then re-referenced to mastoid electrodes and visually inspected for gross artifacts, which were manually removed together with incorrect responses or response latencies that were more than twice the mean latency per subject. Only correctly discriminated trials were included in the following analysis to ensure participants were attending to the stimuli and task in hand.

To ensure a stable ICA decomposition, a minimum contribution of 40 epochs per participant per condition was required for inclusion in the experiment. The average number of useable trials across participants per condition, however, exceeded the threshold for inclusion; thus, no participants were lost due to this criterion.

After following the above described steps to pre-process the individual data, and prior to ICA training, the pre-processed EEG data for each individual participant were concatenated across conditions, in order to obtain one single set of ICA weights. This allowed for comparison of activity across conditions within spatially fixed ICs, i.e., all conditions for each subject bear the same set of component weights, which allows for a valid comparison between conditions. An extended Infomax algorithm (Lee et al., 1999) was then used to decorrelate the data matrix, prior to ICA rotation. The data matrix was subsequently subjected to ICA training using the 'extended *runica*' algorithm in EEGLAB 12. The initial learning rate was set to 0.001 and stopping weight of 10^{-7}. After decomposition, 66 ICs were yielded for each participant, reflecting the total number of recording electrodes (68–2 reference electrodes). Scalp maps for each IC were obtained by projecting the inverse weight matrix (W^{-1}) back onto the spatial EEG channel configuration.

After ICA decomposition, equivalent current dipole (ECD) models for each component were computed, using a standard 4-shell spherical model in the DIPFIT toolbox (Oostenveld & Oostendorp, 2002), freely available at https://sccn.ucsd.edu/eeglab/. Standard 10–20 electrode coordinates were warped to the head model followed by coarse and fine-fitting to the spherical wire matrix, yielding a single dipole model for each 1584 ICs (66 ICs x 24 participants). Dipole localization requires back-projecting the signal to a source that may have generated the scalp distribution for a given IC, and then computing the best forward model to explain the highest percentage of scalp map variance (Delorme et al., 2012). The residual variance (RV) in dipole localizations was also computed, referring to the potential mismatch between the initial scalp map and the forward projection of the ECD model.

2.5.2. *Group EEG data processing/analyses*

The EEGLAB STUDY module was used to conduct group data analyses. This module allows for the analyses of ICA data from multiple participants across conditions, using specified designs relevant for the hypothesis in test. For the purposes of the current study, several contrasts were run to analyze possible significant differences between passive and active conditions. Only components that localized within the cortical volume with a residual variance (RV) < 25% were included in the group analysis as higher RV levels are likely indicative of artifact. The mean RV% for all subjects included in bilateral mu clusters was <15%.

After the STUDY analysis was complete, component pre-clustering was performed, using the K-means statistical toolbox, part of the EEGLAB toolbox (Delorme & Makeig, 2004), which uses information based on common scalp maps, dipoles and spectra to group similar components from each participant via PCA. The resulting neural clusters were then individually inspected in terms of their spectra, scalp maps and dipoles, with special attention paid to the clusters of interest: left and right mu clusters and neighboring clusters. Inclusion criteria for bilateral mu clusters included a characteristic mu spectrum with peaks in alpha and beta frequency ranges, RV < 25%, and localization to accepted mu rhythm generator sites (i.e., BA1 – BA4, and BA6; i.e., somatosensory, premotor, and primary motor regions).

Source localization of bilateral mu clusters consisted of the mean Talairach (x, y, z) coordinates of all contributing components. These cluster coordinates were then mapped to hypothesized cortical sources with the Talairach Client (Lancaster et al., 1997, 2000), freely available at http://www.talairach.org/.

Following the identification and inspection of the mu clusters, an ERSP analysis was employed to compute changes (scaled in normalized dB units) in power across time (i.e., time-frequency analysis) within the spectral range of interest (4–33Hz). Time-frequency transforms were derived using a family of Morlet sinusoidal wavelet set at 3 cycles at 3Hz, rising linearly to 20 cycles at 33Hz. The present study used a 900 ms pre-stimuli period, selected from −2900 to −2000 ms time interval, as the baseline for each trial. This baseline was constructed from a surrogate distribution based on estimates of spectral power from 200 randomly selected latency windows from within the 1000 ms inter-trial interval (Makeig et al., 2004). Subsequent individual ERSP changes from baseline over time were computed using a bootstrap resampling method ($p < 0.025$ corrected for false discovery rate (FDR)). The single trial current for all experimental conditions for frequencies between 7 and 27Hz and times from −500 to 2000 ms were entered in the time-frequency analyses.

3. Results

3.1. *Discrimination accuracy*

Participants' average discrimination accuracy was 98.8% (range: 95–100%). Only correct trials were used in the EEG analysis to ensure that participants were paying attention to the stimuli and task in hand.

3.2. *mu clusters*

Right and left mu-components, localized to sensorimotor regions, were found across all speech perception conditions. This pattern of bilateral activity reflecting sensory-motor transformations during speech perception is expected and has been reported elsewhere (Cogan et al., 2014). Nineteen participants contributed with thirty-nine IC's presenting distinctive markers of the left mu-rhythm and 22 participants contributed with forty-seven IC's presenting distinctive markers of the right mu-rhythm (see Fig. 2). The mean source location of the mu clusters was identified within the frontal lobe, precentral gyrus, and Broadmann area 6, according to the following Taliarach coordinates: [(x, y, z) −40, −5, 51] in the left hemisphere; [(x, y, z) 41, −6, 49] in the right hemisphere. The mean residual variance (RV) in both the left and right mu clusters was less than 9% (8.42% and 8.55% respectively).

3.3. *Event-related spectral perturbations*

Similar patterns of results were found in both the left and right hemispheres, the left generally presenting slightly stronger power.

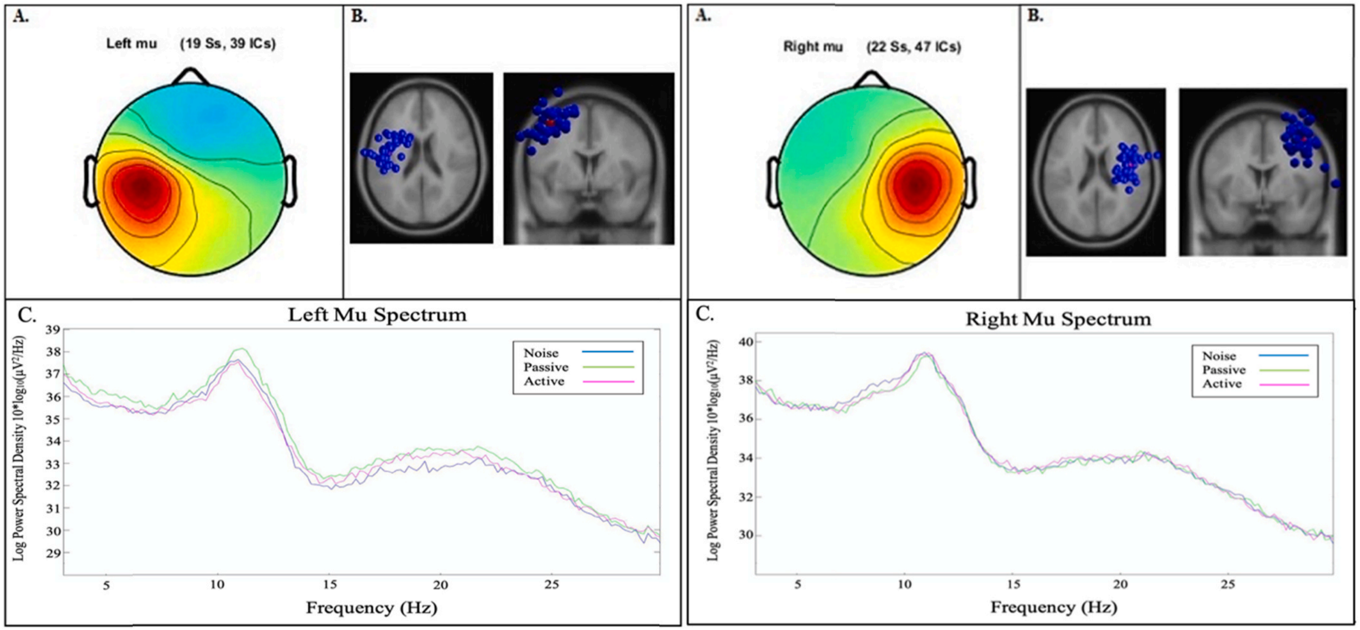

Fig. 2. Left panel: Cluster results for the left hemisphere mu components. Right panel: Cluster results for the right hemisphere mu components. A) Mean scalp potential distribution (W-1) scaled to RMS microvolts and individual scalp distributions for each participant. B) Average equivalent current dipole location. C) Mean spectra of the component as a function of condition.

For this reason, although always presenting figures for both the right and left hemispheres (cf. Cogan et al., 2014), the discussion will focus on the left mu mean ERSP's.

Left and right mu mean ERSP's across subjects and conditions are depicted in a time-frequency graph with bootstrapped ($p < 0.05$),[1] FDR corrected significance values for condition outlined in red (see Fig. 3). The data revealed significantly ($p^{FDR}<0.05$) higher power for mu alpha and beta ERD in active speech perception, when compared to passive speech perception. This effect was evident both during and especially after stimuli presentation with the greatest areas of activation beginning approximately 300 ms after syllable offset. Although weak alpha and beta ERD is visible in the passive task, during the time between the presentation of the two syllables (~600–800 ms), its presence is significantly stronger in the active task.

A separate ANOVA was conducted to further explore differences in sensorimotor activation between the passive speech perception task and the baseline noise condition. The passive speech perception task presented significantly ($p^{FDR}<0.05$) stronger alpha and beta ERD during and between syllable presentation (600–800 ms), as well as significantly stronger alpha ERD right after stimuli offset (1450 ms) (Fig. 4). In addition, when compared to the passive speech perception task, the noise condition presents overall greater increases in power (ERS), both in alpha and beta bands, consistent with inhibition of active phonological processing.

4. Discussion

The main aim of the current project was to disentangle contributions from general cognitive mechanisms and automatic speech-processing to sensorimotor dorsal auditory stream activity observed in speech perception tasks. To accomplish this, we used ERSP analyses to temporally decompose oscillatory activity from EEG mu rhythms in passive and active speech perception tasks. Independent component analysis (ICA) blindly identified bilateral mu rhythms that localized to BA 6 (premotor cortex) which is consistent with known mu rhythm generators (Hari, 2006; Nystrom, 2008) and the anterior hub of the dorsal stream (Hickok & Poeppel, 2004). Across all conditions, ERSP analyses revealed similar patterns of oscillatory activity in left and right mu components. In addition, similar condition-related differences in oscillatory activity were detected in both hemispheres; findings that are consistent with bilateral sensorimotor contributions to speech processing (Cogan et al., 2014). It is important to note that all conditions required a button press response at the end of the trial to help ensure that sufficient attention was being paid to each trial, even during passive listening when the cognitive demands of the task were otherwise low (Mottonen et al., 2014). Furthermore, by employing a button press in all conditions, we controlled for any oscillatory activity due to anticipation of the movement. Observed differences between conditions cannot, therefore, be attributed to the button press response.

We hypothesized that we would observe greater overall sensorimotor activity in active discrimination condition as compared to passive listening condition, and that increased activity would be observed late in the trial (following stimulus offset). In agreement with our hypotheses, the largest difference between the active and passive conditions was found at the end of the trial, with active discrimination being characterized by significantly stronger mu-alpha and mu-beta ERD, following stimulus offset, peaking around 1600 ms. Such broadly extended desynchronization, indicating sensorimotor activity, does not occur immediately after each syllable is heard as would be expected if the activity were related to automatic speech processing. Instead, its late occurrence suggests that it reflects the processes of maintaining or replaying speech stimuli in working memory while the same/different discrimination is made – a cognitive demand that is absent in passive perception. This finding complements previous literature showing that active discrimination elicits stronger anterior dorsal activity than passive listening (Alho et al., 2014; Callan et al., 2014; Meister et al., 2007), and literature suggesting that dorsal stream activity observed in speech perception is largely a result of sensorimotor contribution to working memory (Hickok & Poeppel, 2004). This late mu ERD has been found in similar studies using active discrimination tasks (Jenson et al., 2014; Saltuklaroglu et al., 2017; Thornton et al., 2018) and interestingly, it is the same pattern that is observed in speech production (Gehrig et al., 2012; Mandel et al., 2016; Salmelin & Sams, 2002). Thus, it is possible to surmise that internal modeling mechanisms, similar to those that provide sensorimotor control for speech production, may be operational when holding and perhaps replaying sounds within the phonological working memory buffer (Alho et al., 2012; Saltuklaroglu et al., 2018).

To test the hypothesis that passive speech perception (relatively free of cognitive demands) is sufficient to elicit dorsal stream activity, we compared differences in mu oscillations in two conditions: passively listening to speech syllables and passively listening to white noise. White noise was used as a control condition because previous work (Bowers et al., 2013; Jenson et al., 2014) has demonstrated that it helps maintain attention to the task while incurring a low level of mu oscillatory activity due to maintenance of attention similar to that of passive listening to speech. Supporting our hypothesis, the largest differences found between the passive and noise conditions were associated with mu-alpha ERD beginning approximately 300 ms following the onset of syllables when participants passively listened to speech. This activity is present through syllable presentation and two distinct ERD peaks can be observed, especially in the left hemisphere. The two peaks may correspond with direct sensory-to-motor transformations occurring during automatic processing of each syllable and the consequent activation of premotor/motor cortices. According to the dual stream model, such activation is due to articulatory representations resulting from auditory input, and these mechanisms should therefore happen immediately after acoustic analysis of the speech signal in superior temporal regions. Consistent with this time course, mismatched negativity (MMN) studies indicate that a time window of about 200 ms–300 ms after the stimuli onset is necessary to process an auditory/acoustic speech signal (Näätänen, 2001; Zhang et al., 2005). Furthermore, in a study examining the role of the dorsal auditory stream in phonological processing, using event-related potentials and fMRI, Liebenthal and colleagues (2013) found early and

---

[1] The FDR alpha value in EEGLAB was set to 0.025 for an effective alpha value of 0.05.

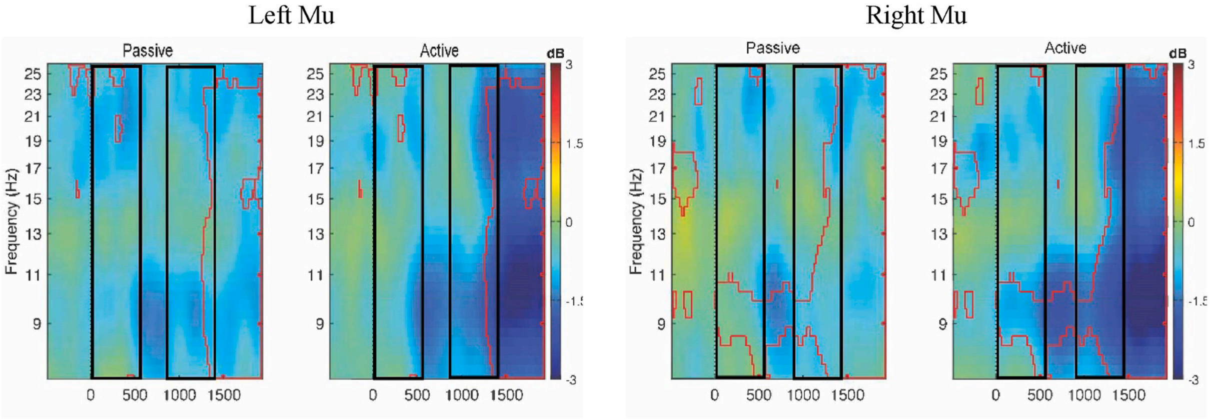

Fig. 3. Mean left and right hemisphere ERSP's for passive vs. active perception tasks. ERSP's are scaled in the same RMS decibel units as a function of condition (1 × 2) and random effects analysis in the traditional alpha (8–13Hz) and beta (13–25Hz) ranges. Areas of significant differences are outlined in red. Event-related decreases in spectral power (ERD) are represented in a blue to teal scale (−3 dB) and indicate an increase in sensory-motor activity. Event-related synchrony (ERS) is represented in a yellow to red scale (3 dB) and indicate a decrease in sensory-motor activity. Boxes mark the time during which each syllable was being presented.

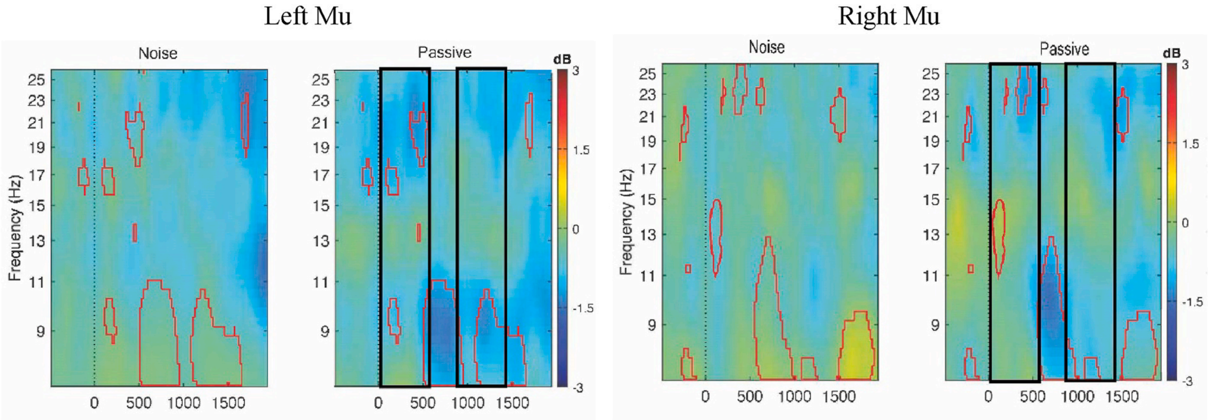

Fig. 4. Mean left and right hemisphere ERSP's for listening to noise vs. passive speech perception task. ERSP's are scaled in the same RMS decibel units as a function of condition (1 × 2) and random effects analysis in the traditional alpha (8–13Hz) and beta (13–25Hz) ranges. Areas of significant differences are outlined in red. Event-related decreases in spectral power (ERD) are represented in a blue to teal scale (−3 dB) and indicate an increase in sensory-motor activity. Event-related synchrony (ERS) is represented in a yellow to red scale (3 dB) and indicate a decrease in sensory-motor activity. Boxes mark the time during which each syllable was being presented.

almost simultaneous activity in superior temporal regions (80–100 ms after stimulus onset) as well as in somatomotor brain areas (95–230 ms after stimulus onset), indicating that the activation of internal articulatory models of speech occur quite early in the neural processing of the speech signal, and may, in fact, be automatic and independent of working memory maintenance.

It is worth mentioning that there is a low level of mu suppression present in the passive listening to noise condition. It is possible that this activity could indicate mu motor activation. The presence of MRI scanner noise, for example, during tasks has been associated with increased amplitude in measures of motor activity (Cho et al., 1998; Schomers & Pulvermuller, 2016). However, we believe that the activity observed during the passive listening to noise condition is unlikely to be due to motor activation. First, previous work has suggested that mu activity is observed in neurotypical participants during the perception of biological stimuli (e.g., speech or biologically reproducible movement), but not during perception of non-biological stimuli (Cuellar et al., 2012; Ulloa & Pineda, 2007). Second, the observed energy is spread across the trial and inconsistent in amplitude across times and frequencies rather than falling into fairly discrete bands of activity as are observed during processing of biological stimuli. Thus, we believe it is more likely that this activity fluctuations in attention across the trial. We also note that work on the effects of MRI scanner noise does not demonstrate mu activity during a noise only condition; rather it shows how the *addition* of scanner noise affects measures of motor activity. Indeed, we are aware of only one study demonstrating that a noise only condition significantly suppressed mu rhythms (Saltuklaroglu et al., 2017), but crucially mu suppression was only observed in a group of participants who stutter, and not in the neurotypical group who did not exhibit changes from baseline. Moreover, the pattern of suppression observed in the stuttering group was localized to clearly defined alpha and beta bands, and was not diffusely spread across time and frequencies unlike what we observe in the present study. Finally, these fluctuations in the passive noise condition are in the same direction (desynchronization, albeit much smaller) as those observed in the experimental conditions, thus increasing the likelihood of Type II error (i.e., making it more difficult to observe a significant difference from our experimental condition) when between condition comparisons are made. Because we were able to detect statistically significant time-frequency voxels in spite of the small oscillatory fluctuations in the control condition, we maintain that this error was not made and the findings are indeed valid.

Although not directly related to our hypotheses, we also observe a small area of significant ERSP mu-beta ERD in the passive listening condition that occurs about 250 ms after the onset of the first syllable. This activity does not occur after the second syllable, and does not occur during the active discrimination task. We therefore do not believe it is uniquely due to the speech perception process vis à vis, for example sensory-motor transformation. Instead, in the context of the block design, whereby participants were presented with syllable pairs in 80 consecutive trials, we surmise this mu-beta activity reflects a temporal expectation of the next syllable occurring. This is consistent with a proposed role of beta oscillations in the prediction of quasi-periodic auditory events, even in the absence of direct attention (Arnal & Giraud, 2012; Fujioka et al., 2015; Todorovic et al., 2015).

As far as we know, this is the first experiment providing data that compares mu sensorimotor activity related to passive speech perception in comparison to passively listening to noise. The results observed are consistent with previous findings indicating that the mu-rhythm is responsive to biologically relevant/reproducible sounds (Wilson et al., 2004) and with fMRI studies showing that passively listening to speech can elicit motor activation (Murakami et al., 2015; Wilson et al., 2004; Wilson & Iacoboni, 2006). Although significantly weaker mu sensorimotor activity was found in passive speech tasks when compared to active tasks, it is extremely interesting to find that passively listening to speech elicits a certain level of sensorimotor activity which is not present when listening to noise. Further, it is crucial to note the timing of those differences so that they are able to be disentangled from later occurring working memory activity. It is worth noting as well that passive conditions were always presented before active ones in order to eliminate the possibility of priming the participant to discriminatory behavior.

Our findings, however, seem to be inconsistent with those of Bowers et al. (2013b; 2014b), who used a similar ICA/ERSP technique. They failed to find significant mu rhythm oscillatory differences between a similar control condition and passively listening to syllable. However, there are two important methodological differences that may account for this discrepancy. First, Bowers and colleagues' stimuli were only 200 ms long whereas the stimuli in the present experiment were 600 ms. It may be that the 200 ms stimuli were too short to engage automatic phonological processing. Second, Bowers and colleagues' speech stimuli were embedded in noise whereas the speech stimuli in the current experiment were not. Schomers and Pulvermuller (2016) addressed similar discrepancies in findings in fMRI data. They suggested that the presence of acoustic noise (e.g., from a scanner when sparse sampling is not used) in the background is likely to mask the stimuli and reduce sensitivity for detecting sensorimotor activity in speech perception. In the current study, passive listening to speech with a quiet background elicited significantly greater mu ERD than the control condition, possibly because the quiet background allowed for more salient auditory and therefore, articulatory representation of the syllables.

It may also be worthwhile to note that Schomers and Pulvermuller (2016) suggested that motor responses such as a button press may provide an additional source of contamination in passive speech perception studies. However, the temporal resolution afforded by event-related EEG allows any mu ERD (sensorimotor activity) resulting from a button press to be excluded from a trial and activity related to anticipation of a button press to be observed only at the end of trials. Thus, this activity (that may be present in all conditions) did not obscure the passive speech-noise differences observed during stimulus presentation. As such, although lacking the spatial resolution of fMRI, this EEG technique eliminates background noise, effectively controls for activity from motor responses, and appears to be extremely well suited for detecting sensorimotor activity elicited even in perceptual tasks with low cognitive demands.

4.1. Conclusions and future directions

Findings from the present study contribute to a body of research highlighting the multiple applications and benefits of using ERSP analyses to understand speech processing. Our findings suggest that there is automatic activation of motor and premotor areas of the brain in response to listening to speech. We further suggest that this activation can be distinguished, at least in part, from task-related or general cognitive neural activity observed in more challenging tasks by examining the dynamics of ERSPs over the course of a trial. These findings are consistent with very recent ECOG research demonstrating activation in motor pathways while participants listened to sentences, and that these activations were later transformed for use in short-term memory using the pre-motor and sensory-motor cortices as bridges between representations (Musch et al., 2020). Future research to further elucidate the nature of speech processing in both normal and clinical populations is nonetheless necessary. In particular, even in the passive perception tasks in our experiment, it remains unclear if the recorded activity was due to the stimulus being speech. The present study does not compare speech sounds to non-speech sounds, for example. Liebenthal et al. (2013) do just this using an event-related fMRI paradigm. They report greater motor and premotor activation when participants identified syllables as compared to when they identified chirps and concluded that their findings supported the dual stream model. The study, however, uses an active rather than a passive paradigm, and is subject to the effects of cognitive processing and perhaps scanner noise described above. It nonetheless suggests that we would expect to find a difference between tone and speech processing. Somewhat contrarily, in active discrimination tasks, differences between speech and tone sweep mu ERSP activity have not been observed (Bowers et al., 2013; Saltuklaroglu et al., 2017). This may be due to the speech-like nature of the tone sweeps used or the cognitive processes in active discrimination overriding any speech-tone differences. Accordingly, it remains to be determined whether such effects would present during passive listening without noise, and whether the degree of motor and premotor activation differs when passively listening to speech and non-speech stimuli.

Applications of this study to different populations would likewise be useful. The present study examined only female adults, and thus it would be useful to determine whether males process speech in the same way, and to what degree such automatic processing might change or develop over the lifespan. Finally, applications to clinical populations such as apraxia patients or individuals with phonological processing difficulties may help to illuminate the nature of the disorders by uncovering differences in automatic activation of brain areas within the dorsal stream.

CRediT authorship contribution statement

Daniela Santos Oliveira: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft. **Tim Saltuklaroglu:** Conceptualization, Methodology, Supervision, Writing – review & editing. **David Thornton:** Formal analysis, Writing – review & editing, Software. **David Jenson:** Formal analysis, Writing – review & editing, Software. **Ashley W. Harkrider:** Supervision, Methodology, Resources. **M. Blake Rafferty:** Software, Visualization, Writing – review & editing. **Devin M. Casenhiser:** Conceptualization, Methodology, Supervision, Writing – original draft, Project administration.

Declaration of competing interest

The authors have no conflict of interest to declare.

References

Adank, P., & Devlin, J. T. (2010, Jan. 1). On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. NeuroImage, 49(1), 1124-1132. https://doi.org/10.1016/j.neuroimage.2009.07.032.

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. *Frontiers in Psychology, 5*, 394. https://doi.org/10.3389/fpsyg.2014.00394

Alho, J., Sato, M., Sams, M., Schwartz, J. L., Tiitinen, H., & Jaaskelainen, I. P. (2012). May, 1). Enhanced early-latency electromagnetic activity in the left premotor cortex is associated with successful phonetic categorization. NeuroImage, 60(4), 1937-1946. https://doi.org/10.1016/j.neuroimage.2012.02.011.

Anderson, E. B., Mitchell, J. F., & Reynolds, J. H. (2011, Jul, 27). Attentional modulation of firing rate varies with burstiness across putative pyramidal neurons in macaque visual area V4. Journal of Neuroscience, 31(30), 10983-10992. https://doi.org/10.1523/JNEUROSCI.0027-11.2011.

Arnal, L. H., & Giraud, A. L. (2012, Jul). Cortical oscillations and sensory predictions. Trends in Cognitive Sciences, 16(7), 390-398. https://doi.org/10.1016/j.tics.2012.05.003.

Belin, P., Zatorre, R. J., & Ahad, P. (2002, Feb). Human temporal-lobe response to vocal sounds. Brain Res Cogn Brain Res, 13(1), 17-26. https://doi.org/10.1016/s0926-6410(01)00084-2.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000, Jan, 20). Voice-selective areas in human auditory cortex. Nature, 403(6767), 309-312. https://doi.org/10.1038/35002078.

Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., & Ward, B. D. (2004, Mar). Neural correlates of sensory and decision processes in auditory object identification. Nature Neuroscience, 7(3), 295-301. https://doi.org/10.1038/nn1198.

Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer. In *(Version 5)*. http://www.praat.org/.

Bowers, A., Saltuklaroglu, T., Harkrider, A., & Cuellar, M. (2013a). Suppression of the micro rhythm during speech and non-speech discrimination revealed by independent component analysis: Implications for sensorimotor integration in speech processing. *PloS One, 8*(8), Article e72024. https://doi.org/10.1371/journal.pone.0072024

Bowers, A., Saltuklaroglu, T., Harkrider, A., & Cuellar, M. (2013b). Suppression of the mu rhythm during speech and non-speech discrimination revealed by independent component analysis: Implications for sensorimotor integration in speech processing. *PloS One, 8*(8), Article e72024. https://doi.org/10.1371/journal.pone.0072024

Bowers, A. L., Saltuklaroglu, T., Harkrider, A., Wilson, M., & Toner, M. A. (2014a). Dynamic modulation of shared sensory and motor cortical rhythms mediates speech and non-speech discrimination performance. *Frontiers in Psychology, 5*, 366. https://doi.org/10.3389/fpsyg.2014.00366

Bowers, A. L., Saltuklaroglu, T., Harkrider, A., Wilson, M., & Toner, M. A. (2014b). Dynamic modulation of shared sensory and motor cortical rhythms mediates speech and non-speech discrimination performance. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00366

Buchsbaum, B. R., & D'Esposito, M. (2008). Short-term and working memory systems. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (Vol. 3, pp. 237–260) (Elsevier.).

Buschman, T. J., & Miller, E. K. (2007, Mar, **30**). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science, 315(5820), 1860-1862. https://doi.org/10.1126/science.1138071.

Callan, D., Callan, A., & Jones, J. A. (2014). Speech motor brain regions are differentially recruited during perception of native and foreign-accented phonemes for first and second language listeners. *Frontiers in Neuroscience, 8*, 275. https://doi.org/10.3389/fnins.2014.00275

Callan, D. E., Jones, J. A., Callan, A. M., & Akahane-Yamada, R. (2004, Jul). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. NeuroImage, 22(3), 1182-1194. https://doi.org/10.1016/j.neuroimage.2004.03.006.

Cho, Z. H., Chung, S. C., Lim, D. W., & Wong, E. K. (1998, Feb). Effects of the acoustic noise of the gradient systems on fMRI: a study on auditory, motor, and visual cortices. Magnetic Resonance in Medicine, 39(2), 331-335. https://doi.org/10.1002/mrm.1910390224.

Cogan, G. B., Thesen, T., Carlson, C., Doyle, W., Devinsky, O., & Pesaran, B. (2014, Mar, 6). Sensory-motor transformations for speech occur bilaterally. Nature, 507(7490), 94-98. https://doi.org/10.1038/nature12935.

Cuellar, M., Bowers, A., Harkrider, A. W., Wilson, M., & Saltuklaroglu, T. (2012, Aug). Mu suppression as an index of sensorimotor contributions to speech processing: evidence from continuous EEG signals. International Journal of Psychophysiology, 85(2), 242-248. https://doi.org/10.1016/j.ijpsycho.2012.04.003.

Delorme, A., & Makeig, S. (2004, Mar, 15). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods, 134(1), 9-21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

Delorme, A., Palmer, J., Onton, J., Oostenveld, R., & Makeig, S. (2012). Independent EEG sources are dipolar. *PloS One, 7*(2), Article e30135. https://doi.org/10.1371/journal.pone.0030135

Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014, May, 13). Noise differentially impacts phoneme representations in the auditory and speech motor systems. Proceedings of the National Academy of Sciences of the United States of America, 111(19), 7126-7131. https://doi.org/10.1073/pnas.1318738111.

Fujioka, T., Ross, B., & Trainor, L. J. (2015, Nov, 11). Beta-Band Oscillations Represent Auditory Beat and Its Metrical Hierarchy in Perception and Imagery. Journal of Neuroscience, 35(45), 15187-15198. https://doi.org/10.1523/JNEUROSCI.2397-15.2015.

Gehrig, J., Wibral, M., Arnold, C., & Kell, C. A. (2012). Setting up the speech production network: How oscillations contribute to lateralized information routing. *Frontiers in Psychology, 3*, 169. https://doi.org/10.3389/fpsyg.2012.00169

Graimann, B., & Pfurtscheller, G. (2006). Quantification and visualization of event-related changes in oscillatory brain activity in the time–frequency domain. In N. Christa, & K. Wolfgang (Eds.), *Prog brain res, ume 159* pp. 79–97). https://doi.org/10.1016/S0079-6123(06)59006-5. Elsevier.

Hari, R. (2006). Action-perception connection and the cortical mu rhythm. *Progress in Brain Research, 159*, 253–260. https://doi.org/10.1016/S0079-6123(06)59017-X

Heald, S. L., & Nusbaum, H. C. (2014). Speech perception as an active cognitive process. *Frontiers in Systems Neuroscience, 8*, 35. https://doi.org/10.3389/fnsys.2014.00035

Hickok, G., & Poeppel, D. (2004, May-Jun). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition, 92 (1–2), 67-99. https://doi.org/10.1016/j.cognition.2003.10.011.

Iacoboni, M. (2008, Jan-May). The role of premotor cortex in speech perception: evidence from fMRI and rTMS. Journal of Physiology Paris, 102(1–3), 31-34. https://doi.org/10.1016/j.jphysparis.2008.03.003.

Jacquemot, C., & Scott, S. K. (2006, Nov). What is the relationship between phonological short-term memory and speech processing? Trends in Cognitive Sciences, 10 (11), 480-486. https://doi.org/10.1016/j.tics.2006.09.002.

Jenson, D., Bowers, A. L., Harkrider, A. W., Thornton, D., Cuellar, M., & Saltuklaroglu, T. (2014). Temporal dynamics of sensorimotor integration in speech perception and production: Independent component analysis of EEG data. *Frontiers in Psychology, 5*, 656. https://doi.org/10.3389/fpsyg.2014.00656

Jenson, D., Harkrider, A. W., Thornton, D., Bowers, A. L., & Saltuklaroglu, T. (2015). Auditory cortical deactivation during speech production and following speech perception: An EEG investigation of the temporal dynamics of the auditory alpha rhythm. *Frontiers in Human Neuroscience, 9*, 534. https://doi.org/10.3389/fnhum.2015.00534

Jenson, D., Thornton, D., Harkrider, A. W., & Saltuklaroglu, T. (2019, Dec). Influences of cognitive load on sensorimotor contributions to working memory: An EEG investigation of mu rhythm activity during speech discrimination. Neurobiology of Learning and Memory, 166, 107098. https://doi.org/10.1016/j.nlm.2019.107098.

Jones, S. R., Kerr, C. E., Wan, Q., Pritchett, D. L., Hamalainen, M., & Moore, C. I. (2010, Oct, 13). Cued spatial attention drives functionally relevant modulation of the mu rhythm in primary somatosensory cortex. Journal of Neuroscience, 30(41), 13760-13765. https://doi.org/10.1523/JNEUROSCI.2969-10.2010.

Kilavik, B. E., Zaepffel, M., Brovelli, A., MacKay, W. A., & Riehle, A. (2013, Jul). The ups and downs of beta oscillations in sensorimotor cortex. Experimental Neurology, 245, 15-26. https://doi.org/10.1016/j.expneurol.2012.09.014.

Klem, G. H., Luders, H. O., Jasper, H. H., & Elger, C. (1999). The ten-twenty electrode system of the international federation. The international federation of clinical neurophysiology. *Electroencephalography & Clinical Neurophysiology - Supplement, 52*, 3–6. https://www.ncbi.nlm.nih.gov/pubmed/10590970.

Klimesch, W. (2012, Dec). alpha-band oscillations, attention, and controlled access to stored information. Trends in Cognitive Sciences, 16(12), 606-617. https://doi.org/10.1016/j.tics.2012.10.007.

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007, Jan). EEG alpha oscillations: the inhibition-timing hypothesis. Brain Research Reviews, 53(1), 63-88. https://doi.org/10.1016/j.brainresrev.2006.06.003.

Klimesch, W., Schimke, H., Doppelmayr, M., Ripper, B., Schwaiger, J., & Pfurtscheller, G. (1996, Nov). Event-related desynchronization (ERD) and the Dm effect: does alpha desynchronization during encoding predict later recall performance? International Journal of Psychophysiology, 24(1–2), 47-60. https://doi.org/10.1016/s0167-8760(96)00054-2.

Krawczyk, D. C. (2002, Oct). Contributions of the prefrontal cortex to the neural basis of human decision making. Neuroscience & Biobehavioral Reviews, 26(6), 631-664. https://doi.org/10.1016/s0149-7634(02)00021-0.

Lancaster, J. L., Rainey, L. H., Summerlin, J. L., Freitas, C. S., Fox, P. T., Evans, A. C., Toga, A. W., & Mazziotta, J. C. (1997). Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. *Human Brain Mapping, 5*(4), 238–242. https://doi.org/10.1002/(SICI)1097-0193(1997)5:4<238::AID-HBM6>3.0.CO;2-4

Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999, Feb, 15). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. Neural Computation, 11(2), 417-441. https://doi.org/10.1162/089976699300016719.

Liebenthal, E., Sabri, M., Beardsley, S. A., Mangalathu-Arumana, J., & Desai, A. (2013, Sep, 25). Neural dynamics of phonological processing in the dorsal auditory stream. Journal of Neuroscience, 33(39), 15414-15424. https://doi.org/10.1523/JNEUROSCI.1511-13.2013.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004, May). Mining event-related brain dynamics. Trends in Cognitive Sciences, 8(5), 204-210. https://doi.org/10.1016/j.tics.2004.03.008.

Mandel, A., Bourguignon, M., Parkkonen, L., & Hari, R. (2016, Feb, 12). Sensorimotor activation related to speaker vs. listener role during natural conversation. Neuroscience Letters, 614, 99-104. https://doi.org/10.1016/j.neulet.2015.12.054.

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Lacoboni, M. (2007, Oct, 9). The essential role of premotor cortex in speech perception. Current Biology, 17 (19), 1692-1696. https://doi.org/10.1016/j.cub.2007.08.064.

Mottonen, R., van de Ven, G. M., & Watkins, K. E. (2014, Mar, 12). Attention fine-tunes auditory-motor processing of speech sounds. Journal of Neuroscience, 34(11), 4064-4069. https://doi.org/10.1523/JNEUROSCI.2214-13.2014.

Murakami, T., Kell, C. A., Restle, J., Ugawa, Y., & Ziemann, U. (2015, Jan, 28). Left dorsal speech stream components and their contribution to phonological processing. Journal of Neuroscience, 35(4), 1411-1422. https://doi.org/10.1523/JNEUROSCI.0246-14.2015.

Musch, K., Himberger, K., Tan, K. M., Valiante, T. A., & Honey, C. J. (2020, Feb, 11). Transformation of speech sequences in human sensorimotor circuits. Proceedings of the National Academy of Sciences of the United States of America, 117(6), 3203-3213. https://doi.org/10.1073/pnas.1910939117.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology, 38*(1), 1–21. https://doi.org/10.1111/1469-8986.3810001

Nystrom, P. (2008). The infant mirror neuron system studied with high density EEG. *Social Neuroscience, 3*(3–4), 334–347. https://doi.org/10.1080/17470910701563665

Onton, J., Westerfield, M., Townsend, J., & Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neuroscience & Biobehavioral Reviews, 30*(6), 808–822. https://doi.org/10.1016/j.neubiorev.2006.06.007

Oostenveld, R., & Oostendorp, T. F. (2002, Nov). Validating the boundary element method for forward and inverse EEG computations in the presence of a hole in the skull. Human Brain Mapping, 17(3), 179-192. https://doi.org/10.1002/hbm.10061.

Osnes, B., Hugdahl, K., & Specht, K. (2011, Feb, 1). connectivity analysis demonstrates involvement of premotor cortex during speech perception. NeuroImage, 54(3), 2437-2445. https://doi.org/10.1016/j.neuroimage.2010.09.078.

Peschke, C., Ziegler, W., Eisenberger, J., & Baumgaertner, A. (2012, Jan, **2**). Phonological manipulation between speech perception and production activates a parieto-frontal circuit. NeuroImage, 59(1), 788-799. https://doi.org/10.1016/j.neuroimage.2011.07.025.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999, Nov). Event-related EEG/MEG synchronization and desynchronization: basic principles. Clinical Neurophysiology, 110 (11), 1842-1857. https://www.ncbi.nlm.nih.gov/pubmed/10576479.

Pineda, J. A. (2005, Dec, 1). The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing". Brain Res Brain Res Rev, 50(1), 57-68. https://doi.org/10.1016/j.brainresrev.2005.04.005.

Poldrack, R. A. (2000, Jul). Imaging brain plasticity: conceptual and methodological issues–a theoretical review. NeuroImage, 12(1), 1-13. https://doi.org/10.1006/nimg.2000.0596.

Pulvermuller, F., & Fadiga, L. (2010, May). Active perception: sensorimotor circuits as a cortical basis for language. Nature Reviews Neuroscience, 11(5), 351-360. https://doi.org/10.1038/nrn2811.

Pulvermuller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006, May, 16). Motor cortex maps articulatory features of speech sounds. Proceedings of the National Academy of Sciences of the United States of America, 103(20), 7865-7870. https://doi.org/10.1073/pnas.0509989103.

Salmelin, R., & Sams, M. (, Jun). Motor cortex involvement during verbal versus non-verbal lip and tongue movements. Human Brain Mapping, 16(2), 81-91. https://doi.org/10.1002/hbm.10031.

Saltuklaroglu, T., Bowers, A., Harkrider, A. W., Casenhiser, D., Reilly, K. J., Jenson, D. E., & Thornton, D. (2018, Dec). EEG mu rhythms: Rich sources of sensorimotor information in speech processing. Brain and Language, 187, 41-61. https://doi.org/10.1016/j.bandl.2018.09.005.

Saltuklaroglu, T., Harkrider, A. W., Thornton, D., Jenson, D., & Kittilstved, T. (2017, Jun). EEG Mu (micro) rhythm spectra and oscillatory activity differentiate stuttering from non-stuttering adults. NeuroImage, 153, 232-245. https://doi.org/10.1016/j.neuroimage.2017.04.022.

Schomers, M. R., & Pulvermuller, F. (2016). Is the sensorimotor cortex relevant for speech perception and understanding? An integrative review. *Frontiers in Human Neuroscience, 10*, 435. https://doi.org/10.3389/fnhum.2016.00435

Scott, S. K., McGettigan, C., & Eisner, F. (2009, Apr). A little more conversation, a little less action–candidate roles for the motor cortex in speech perception. Nature Reviews Neuroscience, 10(4), 295-302. https://doi.org/10.1038/nrn2603.

Scott, S. K., Rosen, S., Wickham, L., & Wise, R. J. (2004, Feb). A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. Journal of the Acoustical Society of America, 115(2), 813-821. https://doi.org/10.1121/1.1639336.

Siegel, M., Donner, T. H., & Engel, A. K. (2012, Jan, 11). Spectral fingerprints of large-scale neuronal interactions. Nature Reviews Neuroscience, 13(2), 121-134. https://doi.org/10.1038/nrn3137.

Skipper, J. I., Devlin, J. T., & Lametti, D. R. (2017, Jan). The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. Brain and Language, 164, 77-105. https://doi.org/10.1016/j.bandl.2016.10.004.

Tamura, T., Gunji, A., Takeichi, H., Shigemasu, H., Inagaki, M., Kaga, M., & Kitazaki, M. (2012). Audio-vocal monitoring system revealed by mu-rhythm activity. *Frontiers in Psychology, 3*, 225. https://doi.org/10.3389/fpsyg.2012.00225

Thornton, D., Harkrider, A. W., Jenson, D., & Saltuklaroglu, T. (2018, Dec). Sensorimotor activity measured via oscillations of EEG mu rhythms in speech and non-speech discrimination tasks with and without segmentation demands. Brain and Language, 187, 62-73. https://doi.org/10.1016/j.bandl.2017.03.011.

Thornton, D., Harkrider, A. W., Jenson, D. E., & Saltuklaroglu, T. (2019, Jan, 23). Sex differences in early sensorimotor processing for speech discrimination. Scientific Reports, 9(1), 392. https://doi.org/10.1038/s41598-018-36775-5.

Todorovic, A., Schoffelen, J. M., van Ede, F., Maris, E., & de Lange, F. P. (2015). Temporal expectation and attention jointly modulate auditory oscillatory activity in the beta band. *PloS One, 10*(3), Article e0120288. https://doi.org/10.1371/journal.pone.0120288

Ulloa, E. R., & Pineda, J. A. (2007, Nov, 2). Recognition of point-light biological motion: mu rhythms and mirror neuron activity. Behavioural Brain Research, 183(2), 188-194. https://doi.org/10.1016/j.bbr.2007.06.007.

Venezia, J. H., Saberi, K., Chubb, C., & Hickok, G. (2012). Response bias modulates the speech motor system during syllable discrimination. *Frontiers in Psychology, 3*, 157. https://doi.org/10.3389/fpsyg.2012.00157

Weisz, N., Hartmann, T., Muller, N., Lorenz, I., & Obleser, J. (2011). Alpha rhythms in audition: Cognitive and clinical perspectives. *Frontiers in Psychology, 2*, 73. https://doi.org/10.3389/fpsyg.2011.00073

Wilson, S. M., & Iacoboni, M. (2006, Oct, 15). Neural responses to non-native phonemes varying in producibility: evidence for the sensorimotor nature of speech perception. NeuroImage, 33(1), 316-325. https://doi.org/10.1016/j.neuroimage.2006.05.032.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004, Jul). Listening to speech activates motor areas involved in speech production. Nature Neuroscience, 7 (7), 701-702. https://doi.org/10.1038/nn1263.

Wostmann, M., Lim, S. J., & Obleser, J. (2017, Jun, 1). The Human Neural Alpha Response to Speech is a Proxy of Attentional Control. Cerebral Cortex, 27(6), 3307-3317. https://doi.org/10.1093/cercor/bhx074.

Zekveld, A. A., Heslenfeld, D. J., Festen, J. M., & Schoonhoven, R. (2006, Oct 1). Top-down and bottom-up processes in speech comprehension. NeuroImage, 32(4), 1826-1836. https://doi.org/10.1016/j.neuroimage.2006.04.199.

Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. i. (2005, 7/1/). Effects of language experience: Neural commitment to language-specific auditory patterns. NeuroImage, 26(3), 703-720. https://doi.org/10.1016/j.neuroimage.2005.02.040.