Temporal dynamics of sensorimotor integration in speech perception and production: independent component analysis of EEG data

David Jenson[1], Andrew L. Bowers[2], Ashley W. Harkrider[1], David Thornton[1], Megan Cuellar[3] and Tim Saltuklaroglu[1]*

[1] *Department of Audiology and Speech Pathology, University of Tennessee Health Science Center, Knoxville, TN, USA*
[2] *Department of Communication Disorders, University of Arkansas, Fayetteville, AR, USA*
[3] *Speech-Language Pathology Program, College of Health Sciences, Midwestern University, Chicago, IL, USA*

Activity in anterior sensorimotor regions is found in speech production and some perception tasks. Yet, how sensorimotor integration supports these functions is unclear due to a lack of data examining the timing of activity from these regions. Beta (~20 Hz) and alpha (~10 Hz) spectral power within the EEG μ rhythm are considered indices of motor and somatosensory activity, respectively. In the current study, perception conditions required discrimination (same/different) of syllables pairs (/ba/ and /da/) in quiet and noisy conditions. Production conditions required covert and overt syllable productions and overt word production. Independent component analysis was performed on EEG data obtained during these conditions to (1) identify clusters of μ components common to all conditions and (2) examine real-time event-related spectral perturbations (ERSP) within alpha and beta bands. 17 and 15 out of 20 participants produced left and right μ-components, respectively, localized to precentral gyri. Discrimination conditions were characterized by significant (pFDR < 0.05) early alpha event-related synchronization (ERS) prior to and during stimulus presentation and later alpha event-related desynchronization (ERD) following stimulus offset. Beta ERD began early and gained strength across time. Differences were found between quiet and noisy discrimination conditions. Both overt syllable and word productions yielded similar alpha/beta ERD that began prior to production and was strongest during muscle activity. Findings during covert production were weaker than during overt production. One explanation for these findings is that μ-beta ERD indexes early predictive coding (e.g., internal modeling) and/or overt and covert attentional/motor processes. μ-alpha ERS may index inhibitory input to the premotor cortex from sensory regions prior to and during discrimination, while μ-alpha ERD may index sensory feedback during speech rehearsal and production.

Keywords: speech perception, speech production, EEG, mu rhythm, independent component analysis

INTRODUCTION

It remains critical to disentangle the neural networks that allow an infinite array of co-articulated vocal tract gestures to be produced by a speaker and effortlessly sensed, recognized, and understood by a listener. Though these two complimentary and highly integrated processes often are examined independently, considerable recent effort has focused upon understanding how classical production mechanisms (e.g., the motor system) are involved in speech perception (D'Ausilio et al., 2012; Mottonen and Watkins, 2012; Murakami et al., 2013) and classical perception regions (i.e., auditory and somatosensory systems) are involved in production (Burnett et al., 1998; Stuart et al., 2002; Purcell and Munhall, 2006). Sensorimotor integration (SMI) provides an interface for speech perception and production and is fundamental to efficient verbal communication (e.g., Perrier et al., 1996; Rogalsky et al., 2011; Tourville and Guenther, 2011; Guenther and Vladusich, 2012; Moulin-Frier and Arbib, 2013). However, questions regarding the nature and timing of SMI prevail and relatively few studies address SMI in both speech perception and production within the same experiment (Wilson et al., 2004; Pickering and Garrod, 2007; Hickok et al., 2011; Adank, 2012).

Neuroimaging techniques have identified the auditory dorsal pathway (posterior temporal lobe, inferior parietal lobe, premotor cortex; PMC) as playing a role in both speech perception and production. In production, there is clear evidence of cooperation between feedforward and feedback systems for motor control (Houde and Nagarajan, 2011). Within speech perception, SMI is explained through independent yet convergent "dual streams" of neural activity (Scott and Johnsrude, 2003; Hickok and Poeppel, 2004; Hickok, 2009, 2012; Rauschecker, 2012; Specht, 2013). The ventral stream (predominantly within auditory regions) provides speech decoding and comprehension. The dorsal stream

(including activity from sensorimotor regions) is thought to provide an audio-motor interface linking auditory to articulatory goals in speech perception. Though dorsal stream activity has been reported to be left-lateralized, there is recent evidence of bilateral organization (Cogan et al., 2014; Simmonds et al., 2014).

Despite evidence that the motor system is relatively inactive in non-degraded passive listening conditions (Scott et al., 2009; Szenkovits et al., 2012) and lesions demonstrating that damage to motor regions has little effect on the ability to perceive speech (Rogalsky et al., 2011), a number of perception tasks have been identified in imaging studies in which motor regions are recruited. These conditions typically have been those in which task demands are increased and include categorical discrimination of foreign phonemes (Callan et al., 2006), phoneme segmentation (Burton et al., 2000; Locasto et al., 2004; Burton and Small, 2006), and speech in noise (Osnes et al., 2011; Alho et al., 2012; D'Ausilio et al., 2012). Thus, motor system activity in speech perception may be context dependent, in addition to being variable across individuals (Szenkovits et al., 2012).

Given equivocal findings, the role of the motor system in speech perception is hotly debated. Perhaps a more pertinent question is the extent to which dorsal stream motor activity functionally enhances the perceptual process (Gallese et al., 2011). Hickok et al. (2011) maintain that contributions of the motor system are strictly modulatory and depend on the cognitive demands associated with a particular task. Others support a more functional role (Binder et al., 2004; Meister et al., 2007; D'Ausilio et al., 2009; Sato et al., 2009; Osnes et al., 2011; Grabski et al., 2013; Mottonen et al., 2013). In these studies, dorsal stream articulatory-motor based speech representations are associated with accurate speech perception in some tasks. However, to bolster understanding of the functional contributions of dorsal stream motor activity in speech perception, it is necessary to address the time-course of activity relative to acoustic stimulation in addition to task performance.

For example, Callan et al. (2010) used a combination of fMRI and magnetoencephalography (MEG), measuring PMC activity in a forced-choice, syllable discrimination in noise task. For correct discriminations, activity in the PMC preceded and immediately followed acoustic stimulation. These findings were interpreted as PMC activity functionally aiding in speech perception and were explained from a Constructivist perspective. That is, previous sensorimotor experiences bestow the motor system with the capacity to provide early top-down influences (in the form of predictive internal models) to help constrain sensory analysis and aid in perception (Sohoglu et al., 2012). In a manner that is also consistent with earlier analysis-by-synthesis theories (Stevens and Halle, 1967), these data suggest motor activity should be maintained while the internal model (i.e., hypothesis) is compared to the sensory consequences. Had motor activity in this study been found in a different time frame, other explanations might arise. For example, if motor activity only coincided directly with the occurrence of acoustic stimuli and was not related to functional performance, it might be interpreted from a Direct Realist viewpoint (Fowler, 1986), as a motor reflection of sensory stimulation. Similarly, motor activity that followed acoustic offset by 200 ms or

more might be interpreted as covert rehearsal while the acoustic stimuli are kept in working memory (Callan et al., 2010).

Oscillatory models offer a time-sensitive means of examining neural processing of speech. These models posit a strong relation among phases of delta, theta, and gamma oscillations, and the temporal envelope of speech with respect to the encoding of discrete speech units (e.g., syllables). This relation reflects further evidence of auditory-motor coupling grounded in evolutionary adaptation for efficiency (Ghitza et al., 2012; Giraud and Poeppel, 2012). Measuring changes in spectral power across beta (15–25 Hz) and alpha (8–13 Hz) frequency bands may offer an additional method for understanding sensorimotor processing. Beta suppression is often associated with the anticipation of performance (Gladwin et al., 2006; Arnal, 2012; Bickel et al., 2012; Zaepffel et al., 2013) of motor activity and predictive (i.e., a priori) top-down coding for sensory analysis. Alpha bands dominate the human brain and the enhancement or suppression of alpha band power often is considered an indicator of cortical activation/inhibition (Klimesch, 2012). Event-related alpha desynchronization (ERD) is considered a release from inhibition for sensory gating and may also contribute to predictive coding. In addition, alpha power generally is suppressed with increased attentional and cognitive demands. Weisz and colleagues provide evidence of an independent auditory alpha generator, implicating a link to speech perception (Weisz et al., 2011; Obleser and Weisz, 2012). Additionally, in support of alpha sensitivity to speech perception, they found that magnitude of alpha suppression across a broad (prefrontal, temporal, parietal) network corresponded with reductions in speech stimulus intelligibility.

The rolandic mu (μ) rhythm is characterized by an arc-shape, alpha and beta band peaks, and typically localized to sensorimotor regions (Pineda, 2005; Hari, 2006). Spectral power within the μ-rhythm is often considered a down-stream measure of motor activity from the PMC (Pineda, 2005). Suppression of the power in the alpha band of the μ-rhythm (μ-alpha) has been used to measure sensorimotor activity in response to viewing biologically relevant (i.e., reproducible) vs. non-relevant visual stimuli such as hand (Oberman et al., 2005; Perry and Bentin, 2010) and face (Muthukumaraswamy and Johnson, 2004) movements, visually presented speech (Crawcour et al., 2009), and motor imagery tasks (Tamura et al., 2012; Holler et al., 2013). μ-alpha also suppresses to action-based sounds (Pineda et al., 2013), speech stimuli in segmentation tasks, and when identifying speech in noise (Cuellar et al., 2012). Additionally, Tamura et al. (2012) reported μ-alpha suppression to overt and imagined speech production under various types of auditory feedback. Their findings suggest that μ-alpha suppression in speech provides an index of feedback in audio-vocal monitoring. This interpretation seems logical considering that μ-alpha suppression is thought to arise from somatosensory activity when guidance is needed for ongoing movement (Hari, 2006). Considering also that μ-beta suppression is indicative of motor activity, identifying patterns of μ-alpha and μ-beta ERS/ERD across speech tasks is likely to reveal further important information about the timing of motor and sensory contributions to SMI in speech processing.

To this end, Bowers et al. (2013) recently employed an EEG technique to study SMI during speech perception, adapting a

similar design from an fMRI and MEG study (Callan et al., 2010). Specifically, participants passively listened to and actively discriminated (i.e., forced choice, same or different) between pairs of syllables (/ba/ and /da/) and tone sweeps presented in different signal-to-noise ratios (SNRs). Raw data from 30 EEG recording channels were analyzed via independent component analysis (ICA). ICA is blind-source separation (i.e., linear decomposition) tool that can be used both as a strong filter and a means of independent and spatially fixed sources of neural activity (Delorme and Makeig, 2004; Makeig et al., 2004; Onton et al., 2006). Left and right μ-rhythm component clusters with characteristic spectral peaks at ∼10 Hz and ∼20 Hz (Hari, 2006), maximally localized to the sensorimotor cortex with activation extending into the PMC, were identified in most participants. Time-frequency analysis of μ components using event-related spectral perturbation (ERSP) analysis showed ERD in the beta band that was strongest when speech was accurately (>95% correct) discriminated in noise with a SNR of + 4 dB. Most importantly, in this condition only, μ-beta suppression (i.e., motor activity) began prior to speech perception and peaked immediately following stimulus offset. The findings were interpreted in accord with Callan et al. (2010) and others, suggesting that PMC/sensorimotor regions can readily contribute to speech perception (e.g., Skipper et al., 2007). From an oscillatory perspective, they were interpreted as evidence of early top-down influences from the motor system (i.e., internal models), helping to constrain auditory analysis in shared channels between sensorimotor and auditory regions (Arnal and Giraud, 2012).

Bowers et al. (2013) demonstrated that this event-related EEG technique with subsequent ICA/ERSP analysis is suitable for measuring SMI in speech perception. However, it is important to note that all their conditions employed background noise. Mottonen et al. (2013) used rTMS to impair motor representations from the lips and found impaired speech discrimination, suggesting that auditory-motor dorsal stream activity is important for speech discrimination in normal as well as degraded conditions. Alho et al. (2012) reported similar findings using evoked potentials. More generally, μ suppression has been found in anticipation of correctly predicted visual targets suggesting functional support in the task from attentional/motor networks (Bidet-Caulet et al., 2012). Hence, a salient question that remains pertains to the extent to which patterns of beta suppression in noisy speech discrimination tasks can disassociate the influences of a degraded listening environment from those functionally related to accurate categorical perception (Specht, 2014). To further understand how μ rhythms respond in speech discrimination tasks, it is necessary to examine the time-course of μ ERS/ERD in a quiet discrimination condition.

Regions within the dorsal stream help mediate sensorimotor control in speech production (Houde and Nagarajan, 2011; Hickok, 2012; Rauschecker, 2012). Feed-forward motor plans are generated that, once properly trained, allow for fluid generation of co-articulated speech gestures at an appropriate speaking rate (Tourville and Guenther, 2011). In addition, inverse forward models (i.e., efference copies) of predicted sensory consequences are sent from motor regions (i.e., premotor/motor cortex) to higher order auditory (e.g., superior temporal sulcus) and somatosensory (e.g., inferior parietal lobe) sites for dynamic comparisons with auditory and somatosensory production, providing the neurophysiological basis for an ongoing feedback loop. As forward prediction is compared with the intended targets and subsequently integrated with the true sensory (i.e., acoustic) and somatosensory consequences of production, corrective feedback is returned to motor control centers in a manner so efficient as to allow for online correction should a mismatch occur between predicted consequence and the articulatory goal. According to dynamic state feedback control (SFC) models, across the time course of any given speech production, complex dynamic interplay can exist between feedforward and feedback control in response to ongoing changes in vocal tract configurations and acoustic/somatosensory perturbations (Ventura et al., 2009; Houde and Nagarajan, 2011; Golfinopoulos et al., 2011; Hickok, 2012, 2014). Hence, as in speech perception, the addition of temporal data from regions within dorsal stream networks is likely to help foster a better understanding of the feedforward and feedback dynamics in speech production.

The largest obstacle to deploying imaging techniques with high temporal resolution such as EEG and MEG to speech production is signal contamination from muscle artifact. It is well known that myogenic activity from eyes (e.g., blinking), lips, head, and jaw produces robust electrical activity in frequency ranges broad enough to spuriously influence most neural activity. In addition, due to volume conduction, the effects of myogenic activity are not focal and can influence recordings from all cranial electrodes (McMenamin et al., 2011). Due to this limitation, EEG and MEG studies targeting language production networks have employed a variety of experimental designs intended to circumvent overt speech production. These designs have typically involved delayed or covert speech production. As evidence exists showing similarities in neural activity in overt and covert production tasks, Tian and Poeppel (2010, 2012) including the generation of internal models (Sams et al., 2005; Tian and Poeppel, 2010), covert production often provides a viable substitute for overt production tasks. However, in terms of SMI, the two tasks are different and may not share all the same neurophysiology (Ganushchak et al., 2011), especially in some pathological conditions with compromised sensorimotor control such as stuttering (Max et al., 2003; Loucks and De Nil, 2006; Watkins et al., 2008; Hickok et al., 2011; Cai et al., 2014; Connally et al., 2014). Other studies have stopped short of measuring activity during production, instead relying on oscillatory data from a time window prior to actual production. In this vein, a number of ERP studies have measured 'lexical' access and morphological encoding strategies (Hirschfeld et al., 2008; Costa et al., 2009; Dell'acqua et al., 2010; Strijkers et al., 2010). Whole head MEG data have revealed patterns of μ-alpha suppression in auditory regions with μ-beta suppression in auditory-motor (i.e., dorsal stream) integration regions (Gehrig et al., 2012). Similarly, Herman et al. (2013) measured real-time changes in oscillatory data from syllable encoding and pre-production time periods to identify discrete input/output operations within the dorsal stream phonological loop, again highlighting the value of temporal information.

Improvements in source estimations and data analysis techniques, along with continued widespread availability appear to

be contributing to a resurgence of EEG. ICA has been suggested as an effective technique for separating neural from myogenic activity on the basis of the assumption of temporal and spatial independence of components. Therefore, especially when stereotypical in nature, myogenic activity can be separated from neural activity in the unmixing process following ICA training on sufficient data (Delorme and Makeig, 2004; Onton et al., 2006; Gwin et al., 2010). The use of ICA in this capacity has been demonstrated to remove movement artifact while performing hand movements (Shou et al., 2012), walking and running (Gwin et al., 2010; Lau et al., 2014), and in distinguishing distinct patterns of electro-cortical activity in knee vs. ankle movements (Gwin and Ferris, 2011). However, its application to speech production has been limited. Though seemingly daunting, Tran et al. (2004) reported successfully using ICA to remove artifact from stuttered speech in children. In addition, other studies have demonstrated that ICA can be used to reveal neural activity not evident in univariate analyses (Geranmayeh et al., 2012; Simmonds et al., 2014).

Though there is reason for optimism regarding the potential use of EEG with ICA for measuring neural activity in speech production tasks (Ganushchak et al., 2011), concerns remain regarding the potential for ICA to adequately remove all muscle artifact. These include reduced validity for localization in intracerebral space, the fact that muscle artifact is often non-stereotypical and therefore not always suited for identification via ICA, and that a substantial portion of the variance in the whole EEG signal (i.e., up to 67% of components) can be accounted for by pure myogenic activity, reducing spectral power in neural components of interest (McMenamin et al., 2009; Shackman et al., 2009; McMenamin et al., 2010, 2011). It is clear that preliminary investigations using ICA in speech production should proceed cautiously using simple productions.

As a launching point, the current study focuses on activity from μ components for the following reasons. First, they are ubiquitously found in EEG recordings, particularly when identified via ICA. Thus, the possibility of yielding μ components in ICA decomposition remains high even when muscle components predominate. Second, μ rhythms typically are localized to primary motor/PMC regions, which are key sites within the dorsal stream. The PMC, in particular, is bi-directionally connected to higher-level auditory and somatosensory regions via the arcuate and longitudinal fasciculi. Its location and connectivity allow it to serve as an important intermediary for integrating forward prediction (internal modeling) and sensory feedback in both perception and production (Houde and Nagarajan, 2011; Rauschecker, 2012). Third, μ-ERS/ERD already has revealed real-time data interpreted as predictive coding in speech perception (Bowers et al., 2013). Alpha and slow beta bands which are contained within the μ rhythm are the only frequency domains that display ERS/ERD sensitivity to stimulus and/or task (Klimesch, 2012). Therefore, further time-frequency analyses of μ rhythms potentially may reveal important information about SMI in production.

There are two main goals in this study. The first is to bolster understanding of the timing and function of dorsal stream activity in speech perception by examining ERS/ERD patterns

in quiet and noisy discrimination conditions. The second is to provide initial evidence that, via the application of ICA/ERSP, the use of EEG can be extended into the realm of speech production. Collectively, the intention is to show that ICA can be used accurately to identify dorsal stream sensorimotor μ components common to both speech perception and production. It is first hypothesized that right and left μ components, localized to sensorimotor/PMC regions, will be found across perception and production conditions. By placing EMG electrodes on the upper and lower lip, it also is anticipated that ICA will identify prominent perilabial muscular activity. Once μ components are identified, the second hypothesis is that ERSP analyses will provide differential time-frequency measures of alpha and beta ERS/ERD. Real-time oscillatory changes in the spectral power of alpha and beta bands of the μ rhythm are expected to provide novel information regarding the timing of SMI in speech perception and production that may be interpreted via dual stream/SFC models. Additionally, significant activity from perilabial components is expected only in overt production, allowing it to be mapped in real-time to SMI activity.

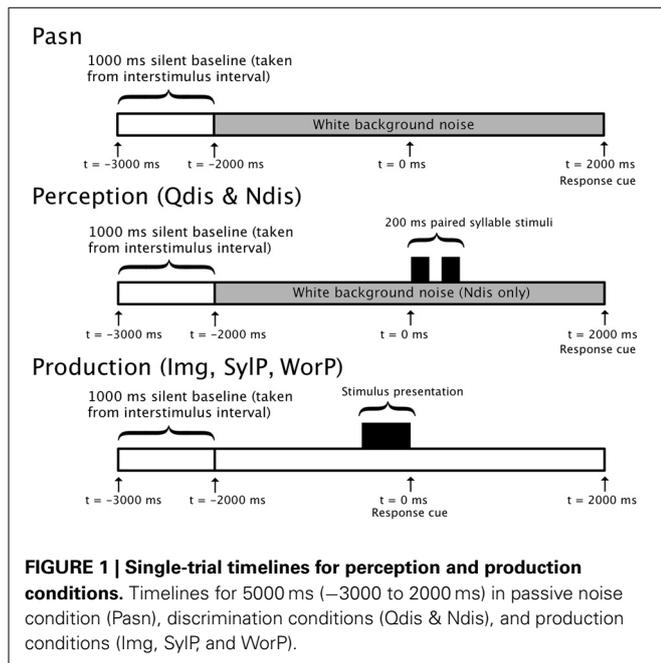## MATERIALS AND METHODS

PARTICIPANTS

Twenty right-handed English-speaking adults (17 female and 3 males) with a mean age of 23.94 years (range 21–39 years) were recruited from audiology and speech pathology classes at the University of Tennessee Health Science Center. Participants reported no diagnosed history of communicative, cognitive, or attentional disorders. Handedness dominance was assessed using the Edinburgh Handedness Inventory (Oldfield, 1971). This study was approved by the Institutional Review Board of the University of Tennessee Health Science Center. Prior to the experiment, participants provided signed informed consent on a document approved by the Institutional Review Board.

STIMULI

Perception

/ba/ and /da/ syllables were created using AT&T naturally speaking text-to-speech software which employs synthetic analogs of a human male speaker. Syllable pairs were generated such that half of the stimuli were composed of different syllables (e.g., /ba/ and /da/) and the other half were identical (e.g., /ba/ and /ba/). The stimuli were low-pass filtered below 5 KHz and normalized for root-mean-square (RMS) amplitude. Each syllable was 200 ms in duration. Each syllable pair was also separated by 200 ms, resulting in a total of 600 ms from the first syllable onset to the second syllable offset (**Figure 1**).

For one condition (discrimination in noise; Ndis), syllable pairs to be discriminated were embedded in white noise with a SNR of +4 dB. This SNR was chosen as it has been shown previously (Bowers et al., 2013) to produce discrimination accuracies > 95% in a similar group of participants. In the other discrimination condition (discrimination in quiet; Qdis), syllable pairs were presented without background noise. To prevent discrimination response bias (Venezia et al., 2012), in both Qdis and Ndis stimuli sets, there were equal numbers of syllable pairs that were identical as there were different.

FIGURE 1 | Single-trial timelines for perception and production conditions. Timelines for 5000 ms (−3000 to 2000 ms) in passive noise condition (Pasn), discrimination conditions (Qdis & Ndis), and production conditions (Img, SylP, and WorP).

Production

Speech targets were syllable pairs, similar to those used in the discrimination tasks above, and tri-syllable nouns (initiated with /b/ or /d/ and followed by a vowel). They were displayed centered in Microsoft PowerPoint slides with plain black backgrounds in large white Arial font (size 56). **Figure 1** shows the timelines for epochs in both perception and production conditions.

DESIGN

A 6-condition within-subject design was employed. Based on extant literature, the conditions were created to increase motoric demands incrementally (i.e., from the perceiving of white noise to overtly producing tri-syllable words). Participants were required to:

(1) passively listen to white noise (Pasn).
(2) discriminate (same or different) between pairs of syllables in a quiet background (Qdis).
(3) discriminate (same or different) between pairs of syllables in a noisy background (Ndis).
(4) imagine producing a pair of syllables (Img).
(5) overtly produce (i.e., say) a pair of syllables (SylP).
(6) overtly produce (i.e., say) tri-syllable nouns initiated by /b/ or /d/ and followed by a vowel (WorP).

Thus, condition 1 (Pasn) required no discrimination and was a control task for the two discrimination conditions (Qdis and Ndis). Conditions 2–5 employed /ba/ and /da/ syllables. Conditions 2 and 3 (Qdis and Ndis) required same/different discriminations of random /ba/ and /da/ combinations, while conditions 4 and 5 required covert (Img) and overt (SylP) production of randomly selected /ba/ and /da/ combinations. Condition 6 (WorP) also required overt production, but in this condition

tri-syllable nouns were used as opposed to the 2-syllable combinations employed in the SylP condition. In the WorP condition, words meeting these criteria were selected from Blockcolsky et al. (2008). Examples of these words include "dialog," "butterscotch," "daffodil," and "buffalo."

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. Compumedics NeuroScan Stim 2 version 4.3.3 software was used to present stimuli to participants via a PC computer and record button-press responses. A button-press response was required for all three perception conditions because anticipation of a button-press has previously been known to elicit μ-rhythm ERD (Makeig et al., 2004; Graimann and Pfurtscheller, 2006; Hari, 2006). Hence, in the Pasn condition, the button-press was used as a control for the required button-press response in the discrimination conditions and to ensure that participants were paying attention in each trial. The cue to respond was a 100 ms, 1000 Hz tone that was presented at the end of the epoch (i.e., +2000 ms). In the Pasn condition, participants were instructed simply to listen passively to the noise and press the designated button after hearing a pure tone cue in each trial. Designation of button-press responses (right or left hand) was counterbalanced across all subjects and experimental conditions. Performance in the discrimination conditions was evaluated by calculating the percentage of correct trials.

In the production conditions, stimuli appeared on a 69.5 × 39.0 cm display placed 132 cm in front of the reclining chair. The stimuli appeared on the screen for 1 s. Participants were instructed to begin their production response immediately when the stimulus disappeared from the monitor. In the Img condition participants were told to imagine saying (i.e., covertly producing) the pair of syllables while refraining from making any overt articulatory movements or vocalization. In the SylP and WorP condition, participants were instructed to speak the syllable pair or word in their normal speaking voice. All overt speech productions were easily completed in the time window (2 s) following the cue to speak. All conditions were presented in two blocks of 40 trials each. The order of the 12 blocks (6 conditions × 2 blocks) was randomized for each participant.

EEG ACQUISITION

Sixty-eight electrode channels were used to acquire whole-head EEG data. These included two electromyography (EMG) and two electrocardiogram (ECG) electrodes. Electrode configuration was based upon the extended international standard 10–20 (Jasper, 1958) method using an unlinked, sintered NeuroScan Quik Cap (Towle et al., 1993). All recording electrodes were referenced to the common linked left (M1) and right (M2) mastoids. 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. The two surface electromyography (EMG) electrodes were placed at midline above the upper lip and below the lower lip for the

purposes of collecting perilabial EMG data related to speech production.

EEG data were collected using Compumedics NeuroScan Scan 4.3.3 software and the Synamps 2 system. The raw EEG data were filtered (0.15–100 Hz) and digitized via a 24-bit analog-to-digital converter at a sampling rate of 500 Hz. Data collection was time-locked to time point zero at the onset of acoustic stimuli delivery in speech perception trials and the cue to begin speaking in production trials. Thus, in the perception conditions, time zero referenced the acoustic onset of the first syllable. In the production conditions, syllable and word stimuli were orthographically displayed on the monitor between times -1000 ms and zero. Hence, disappearance of the text at time zero was the cue for participants to begin speaking (**Figure 1**).

EEG DATA PROCESSING

EEGLAB 12 open source software (Delorme and Makeig, 2004) was used to process all EEG data by performing the following steps for individual and group processing/analysis.

(1) Individual processing/analysis:

    (a) 12 raw EEG files (6 conditions x 2 blocks) were preprocessed for each participant.
    (b) Independent component analysis (ICA) was performed on all concatenated files across all conditions for each participant.
    (c) All neural and non-neural dipoles were localized for each independent component (IC) identified.

(2) Group analysis:

    (a) Using the STUDY module of EEGLAB 12, two separate analyses were performed using 'in head' only (neural) and 'all' (neural and non-neural) ICs.
    (b) Principal component analysis (PCA) subsequently was used to identify and cluster common components across participants.
    (c) Left and right μ clusters were identified from the 'in-head' STUDY, whereas the EMG cluster representing perilabial muscle activity was identified from the 'all' STUDY.
    (d) μ clusters were localized using equivalent current dipole (ECD) and current source density (CSD) analyses.
    (e) Time-frequency analyses (measuring changes in spectral power across time) were performed by measuring event-related spectral perturbations (ERSP) in the left and right μ clusters as well as in the EMG cluster.

Details of each step in the data processing / analyses are described below.

Processing/analysis of EEG data from each participant

Data pre-processing. Raw data from both 40-trial blocks in each condition were: (1) appended to make a single 80 trial data set for each condition; (2) downsampled to 256 Hz to decrease computational requirements for ICA processing; (3)
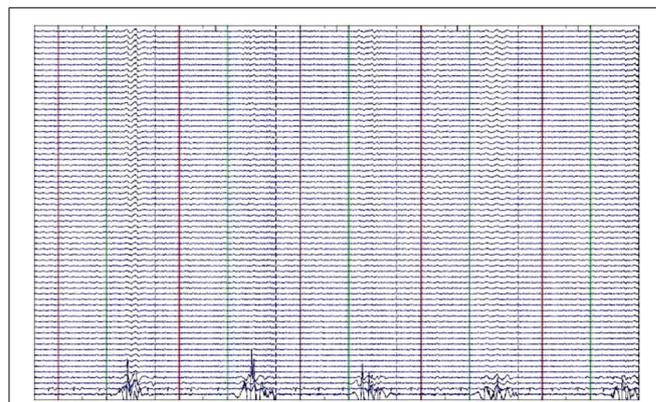
FIGURE 2 | Stereotypic muscle activity from perilabial EMG electrodes. Example of filtered and epoched EEG data from the WorP condition showing stereotypical EMG activity during speech production.

epoched into 5000 ms segments with individual epochs spanning from −3000 to +2000 ms around time zero; (4) bandpass filtered (3–34 Hz) to ensure that alpha and beta could be identified while filtering muscle movement from surrounding frequencies; (5) re-referenced to mastoid electrodes; (6) visually inspected for gross artifact (> 200 μV), which was manually removed; and (7) pruned to remove trials with incorrect responses or response latencies greater than 2 s in the Qdis and Ndis conditions, although few trials were removed (see below). A minimum contribution of 40 epochs per participant per condition was required for inclusion in the experiment. However, the average number of usable trials across participants per condition far exceeded the minimum of 40 required for inclusion. **Figure 2** shows an example of EEG activity from 5 trials in one participant in the WorP condition following filtering and epoching. Critically, the muscle activity from the EMG component appears to be relatively stereotypical in nature (e.g., **Figure 2**), thereby facilitating ICA efforts to separate the neural activity from the muscle activity in the subsequent ICA signal decomposition.

Independent component analysis (ICA). Prior to ICA training, pre-processed EEG data for each participant were concatenated across all 6 conditions so that a single set of ICA weights could be obtained. This allowed for a comparison of activity to be made across conditons within spatially fixed ICs. An extended Infomax algorithm (Lee et al., 1999) was used to decorrelate the data matrix prior to ICA rotation. ICA training was provided using the "extended *runica*" algorithm in EEGLAB 12 with an initial learning rate set to 0.001 and a stopping weight of 10–7. Following decomposition, 66 ICs were yielded for each participant reflecting the total number of recording electrodes (68 − 2 reference electrodes, M1 and M2). Scalp maps for each IC were obtained by projecting the inverse weight matrix (W^{-1}) back onto the spatial EEG channel configuration.

Following ICA decomposition, equivalent current dipole (ECD) models for each component were computed using a boundary element model (BEM) in the DIPFIT toolbox, freely

available at sccn.ucsd.edu/eeglab/dipfit.html (Oostenveld and Oostendorp, 2002). Standard 10–20 electrode coordinates were warped to the head model followed by automated coarse-fitting to a BEM, yielding a single dipole model for each of 1320 ICs (66 ICs × 20 participants). Dipole localization requires back-projecting the signal to a source that may have generated the scalp potential 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). Residual variance (RV) in dipole localizations were also computed, referring to the potential mismatch between the initial scalp map and the forward projection of the ECD model.

Group data analyses

EEGLAB STUDYs. Group data analyses were conducted via the EEGLAB STUDY module. The STUDY module allows ICA data from multiple participants across conditions to be analyzed using specified designs. In the current study, the designs specified were dictated by the within-subjects conditional differences of interest. The STUDY module allows further filtering to be applied with respect to the RV in dipole localization and inclusion vs. exclusion of out-of head dipoles. Thus, ICA files with dipole information from each individual (see above) were applied to the two separate STUDY modules. For the purposes of measuring neural activity, only "in-head" dipoles with RV < 20% were analyzed.

For the purposes of identifying perilabial EMG activity, a second STUDY was conducted that included "all" dipoles from in head and outside the head. In this second STUDY, the RV criterion was raised to 50% (Gramann et al., 2010) dipoles because EMG activity emanates from outside the head and by nature, muscular movement incurs higher unexplained RV.

Principal component clustering of ICs. In both the "in head" and "all" STUDYs, component pre-clustering was performed on the basis of common scalp maps, dipoles, and spectra. The K-means statistical toolbox (implemented in EEGLAB; Delorme and Makeig, 2004) then used these criteria to group similar components from each participant via PCA. After removal of outliers (3 SD from any cluster mean), components from the "in head" STUDY were assigned to 20 possible neural clusters, which included left and right sensorimotor μ clusters. Components in the "all" STUDY were assigned to 66 possible clusters and included one non-neural cluster depicting perilabial EMG activity.

Final component designation to left and right μ clusters was based primarily on the PCA followed by individual inspection of spectra, scalp maps, and dipoles of all components within those clusters and neighboring clusters. Final inclusion criteria for membership to μ clusters included localization to BA 1–4, and 6 (i.e., somatosensory regions, primary motor and premotor regions) and characteristic μ spectra, though over 90% of components emanated from BA 6.

Components in the "all" STUDY were assigned to 66 possible clusters, most of which, as expected, depicted non-neural activity. The cluster characterizing perilabial EMG activity was found on

the basis of dipole location and ERSP analysis showing activity only in overt production tasks (see below).

μ cluster source localization. ECD source localization is simply from the average (x, y, z) coordinate of all the IC dipoles (identified via the DIPFIT module) within a given cluster. Alternatively, standardized low-resolution brain electromagnetic tomography (sLORETA) uses current source density (CSD) distribution from electrical potential measures across the scalp to address the inverse problem and provide an estimate of source localization (Pascual-Marqui, 2002). The head model uses a Talairach cortical probability brain atlas, digitized at the Montreal Neurological Institute (MNI). EEG electrode locations are cross-registered between spherical and realistic head geometry (Towle et al., 1993). Spatial resolution of 5 mm is achieved by sampling 6239 voxels in 3-D brain space. For each IC that contributed to the two μ clusters, the inverse weight projections on the original EEG channels were exported to the sLORETA. Cross-spectra were computed and mapped to the standard Taliarach brain atlas cross-registered with the Montreal Neurological Institute (MNI) coordinates, yielding sLORETA estimates of CSD for left and right μ dipoles in the "in-head" STUDY. To evaluate the statistical significance of dipole locations across participants, statistical comparisons relative to zero (i.e., no activation) were computed (Grin-Yatsenko et al., 2010). Paired (Student) t-tests were conducted on frequencies between 4 and 33 Hz (1000 frames) with the smoothing parameter set to 1 (single common variance for all variables), using 5000 random permutations yielding corrected t-value thresholds and statistical significance ($p < 0.001$) for all 6239 voxels.

While these two methods of EEG source localization were expected to produce similar results (Bowers et al., 2013), for reliability purposes it was deemed useful to use both techniques.

Time-frequency analysis (change in spectral power across time). ERSP analyses were used to compute changes (scaled in normalized dB units) in power across time (i.e., time-frequency analysis) within the spectral range of interest (4–33 Hz). Time-frequency transforms were derived using a Morlet sinusoidal wavelet set at 3 cycles at 3 Hz, rising linearly to 20 cycles at 40 Hz. The 1000 ms pre-stimulus period was selected from the silent inter-trial interval to serve as a baseline for each trial. These baselines were 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 ($p < 0.05$ uncorrected). The single trial current for all experimental conditions for frequencies between 7 and 30 Hz and times from −500 to 1500 ms were entered in the time-frequency analyses.

In the "in-head" STUDY, differences in cross-conditional ERSPs in right and left μ clusters were computed using permutation statistics (2000 permutations) with a 95% confidence interval ($p < 0.05$). The random distribution represents the null hypothesis that no condition differences exist. Type I error was controlled by correcting conservatively for false discovery

rates (*p*FDR; Benjamini and Hochberg, 2000). Statistical analysis in the perception conditions used a 1 × 3 (Pasn, Qdis, Ndis) repeated measures ANOVA design. *Post-hoc* comparisons examined differences between Pasn vs. Qdis and Pasn vs. Ndis conditions. In the production conditions, a 1 × 3 repeated measure ANOVA design examined differences in ERSP activity across the Img, SylP, and WorP conditions. A *post-hoc* paired comparison examined differences between SylP and WorP conditions. In the "all" STUDY, cross-conditional ERSPs were computed in the production conditions using a 1 × 3 repeated measure ANOVA design.

RESULTS

DISCRIMINATION ACCURACY
In participants that contributed to μ clusters, the average number of useable trials (out of 80) across participants in each condition were: Pasn = 73.8 (*SD* = 7.2); Qdis = 74.8 (*SD* = 4.6); Ndis = 69.0 (*SD* = 11.4); Img = 75.0 (*SD* = 5.8); SylP = 71.1 (*SD* = 7.4); WorP = 69.9 (*SD* = 8.0). In the Qdis condition, all participants discriminated with 91–100% accuracy. In the Ndis condition, all except one participant discriminated with 84–100% accuracy. The remaining participant discriminated with 65% accuracy. The average discrimination accuracies in the Qdis and Ndis conditions were 97.3 and 94.4%, respectively. A paired *t*-test indicated that mean discrimination performance was not significantly different (*p* > 0.05) in these conditions. The average response latencies in the Qdis and Ndis conditions were 504 and 545 ms, respectively. A paired *t*-test again indicated that these latencies were not significantly different (*p* > 0.05). Together, these findings suggest that both discrimination

tasks were performed with similar high levels of accuracy and efficiency. It should be noted again that trials with incorrect discriminations were eliminated from the data so the EEG analysis was limited to correct productions only.

μ AND EMG CLUSTER CHARACTERISTICS
As predicted by the first hypothesis, 17/20 and 15/20 participants produced components with < 20% unexplained RV that contributed to left and right μ clusters, respectively. For the left μ cluster, the average Talaraich ECD location was [−41, 4, 46], while on the right it was [46, 0, 39]. The percentage of unexplained RV in these single dipole models was 10.1 and 8.7% for the left and right hemispheres, respectively. sLORETA analyses revealed significantly activated voxels (*p* < 0.001) associated with μ clusters. Maximum current source densities were found at Talairach [−45, −10, 45] on the left vs. [45, −5, 40] on the right. In accord with findings by Bowers et al. (2013), the two localization techniques produced similar results, here allowing sources of μ activity to be maximally localized within the precentral gyri with activity spreading across the PMC and sensorimotor regions. **Figures 3** and **4** respectively display the scalp maps (A), spectra (B), ECD dipole clusters (C) and CSD maxima (D) for left and right μ clusters, respectively. The EMG cluster was characterized by non-neural ICs with an average of 21.3% unexplained RV.

### TIME-FREQUENCY ANALYSES IN PERCEPTION (Pasn, Qdis AND, Ndis) CONDITIONS
Figure 5 shows Van Essen maps (generated using sLORETA) of significant voxels contributing to left (A) and right (B) μ clusters, followed by time-frequency (ERSP) analyses within
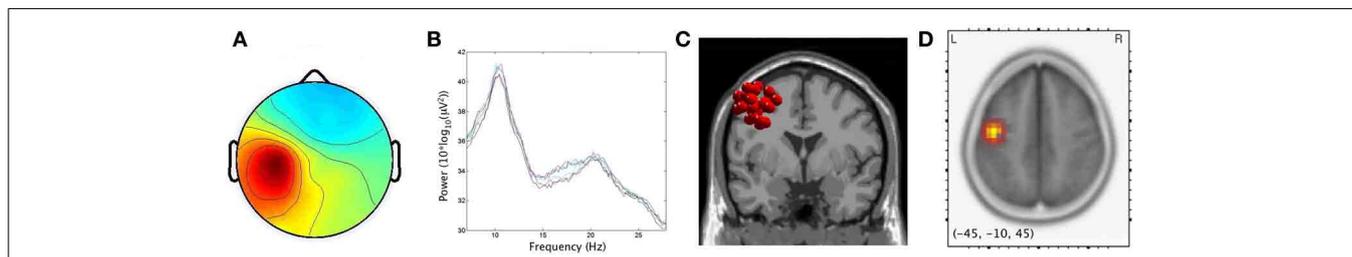

FIGURE 3 | Cluster results for left μ component. (A) mean scalp potential distribution (W⁻¹) scaled to RMS microvolts, **(B)** mean spectra of the components within the cluster for each condition, **(C)** distribution of equivalent current dipoles within the cluster, and **(D)** maximum current source density voxels (*t*-values) with greater values in darker colors and smaller values in lighter colors (at *p* < 0.001 corrected for multiple comparisons).
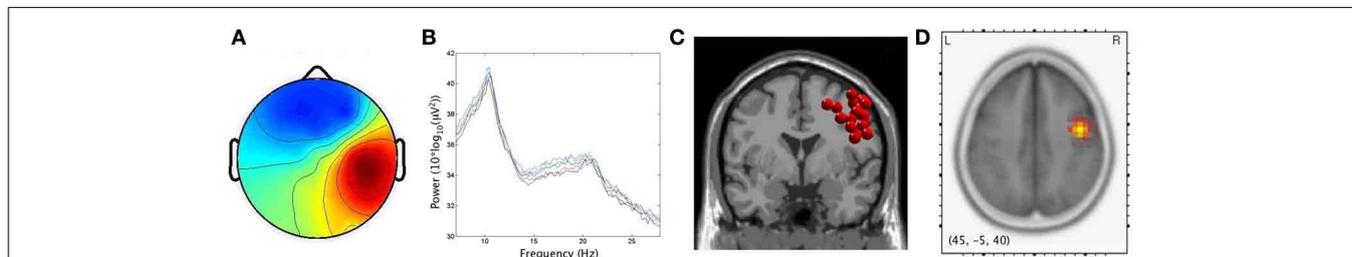
FIGURE 4 | Cluster results for right μ component. (A) mean scalp potential distribution (W⁻¹) scaled to RMS microvolts, **(B)** mean spectra of the components within the cluster for each condition, **(C)** distribution of equivalent current dipoles within the cluster, and **(D)** maximum current source density voxels (*t*-values) with greater values in darker colors and smaller values in lighter colors (at *p* < 0.001 corrected for multiple comparisons).
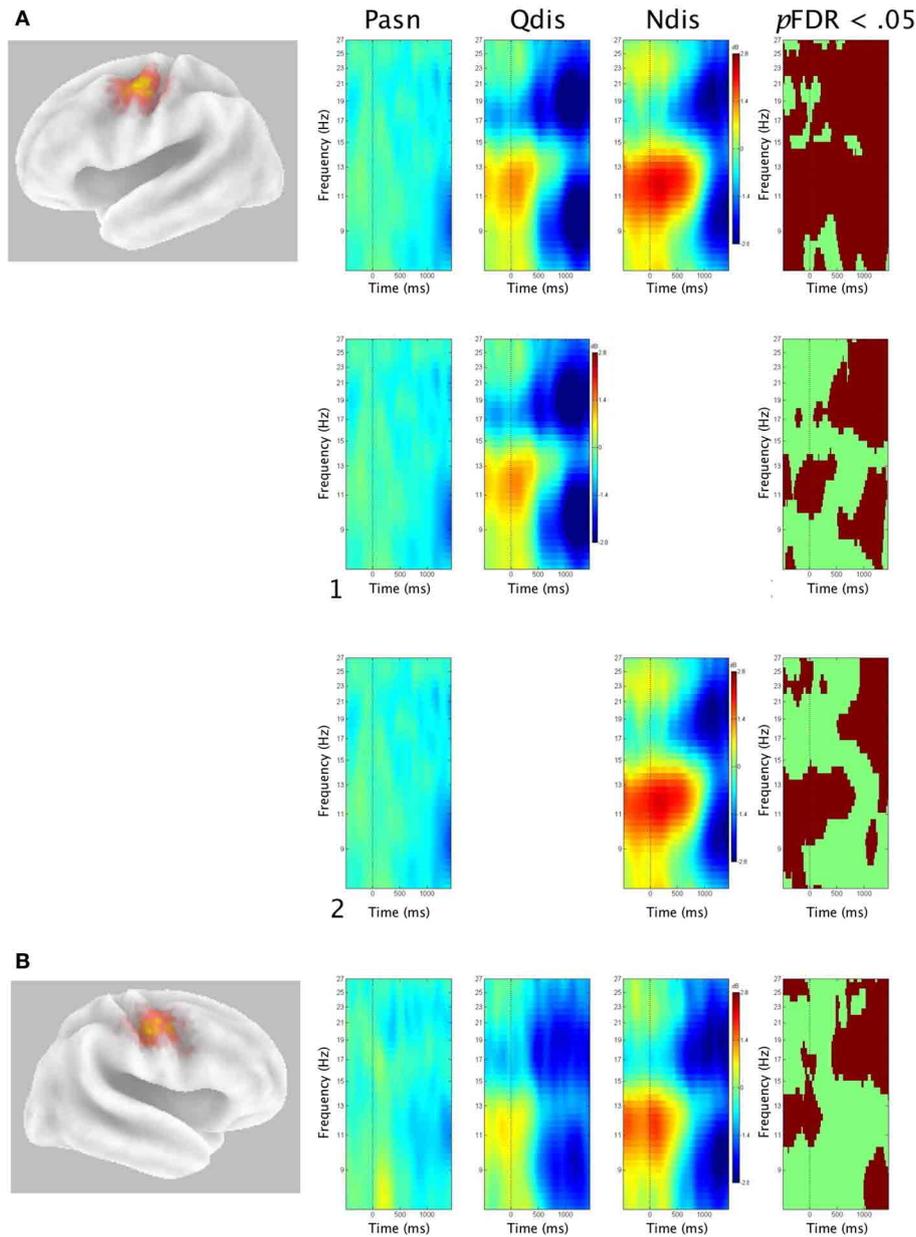
FIGURE 5 | Mean left and right ERSPs and sLORETA solutions for perception conditions. Rows A and B show sLORETA solutions for left and right μ clusters, respectively, depicted on a 3D Van Essen average template, followed by mean time-frequency ERSPs (event-related spectral perturbations) as a function of perception conditions, before, during, and after stimulus offset for **(A)** left μ clusters with (1) contrasts between Pasn and Qdis and (2) contrasts between Pasn and Ndis; and **(B)** right μ clusters (red, ERS, blue, ERD). The last frame in each row shows significant differences across conditions (pFDR< 0.05).

the 7–30 Hz bandwidth. The ERSP analyses show significant ERS/ERD changes from baseline in the Pasn, Qdis, and Ndis conditions. The last frame in each row shows statistical ERSP differences across conditions (pFDR < 0.05), thus supporting the second hypothesis.

For the left μ cluster, relative to the Pasn, alpha ERS began prior to acoustic stimulation and gradually gave way to alpha ERD beginning in low alpha frequencies (8–11 Hz) following acoustic offset in both discrimination conditions (Qdis and Ndis). Beta ERD in both discrimination conditions began in a narrow bandwidth (17–19 Hz), growing stronger and spreading across beta frequencies during and immediately following the acoustic stimulation condition. *Post-hoc* analyses (shown in **Figures 5A1,A2**) show differential patterns of significant beta ERD and alpha ERS/ERD in Pasn vs. Qdis comparisons and Pasn vs. Ndis comparisons.

Patterns of alpha/beta ERS/ERD were similar yet weaker and more diffuse in the right μ cluster compared to those on the left.

It followed that *post-hoc* ERSP comparisons of Qdis and Ndis to Pasn comparisons for right μ activity did not yield additional data of interest.

TIME-FREQUENCY ANALYSES IN PRODUCTION (Img, SylP AND, WorP) CONDITIONS

Figure 6 shows Van Essen maps (generated using sLORETA) of significant voxels contributing to left (A) and right (C) μ clusters, followed by time-frequency (ERSP) analyses within the 7–30 Hz bandwidth. The ERSP analyses show significant ERS/ERD changes from baseline in the Pasn, Qdis, and Ndis conditions. The last frame in each row shows statistical ERSP differences across conditions (pFDR < 0.05), again supporting the second hypothesis. **Figure 6B** shows the average ECD dipole location for the EMG components followed by ERSP analyses with statistical differences across conditions.
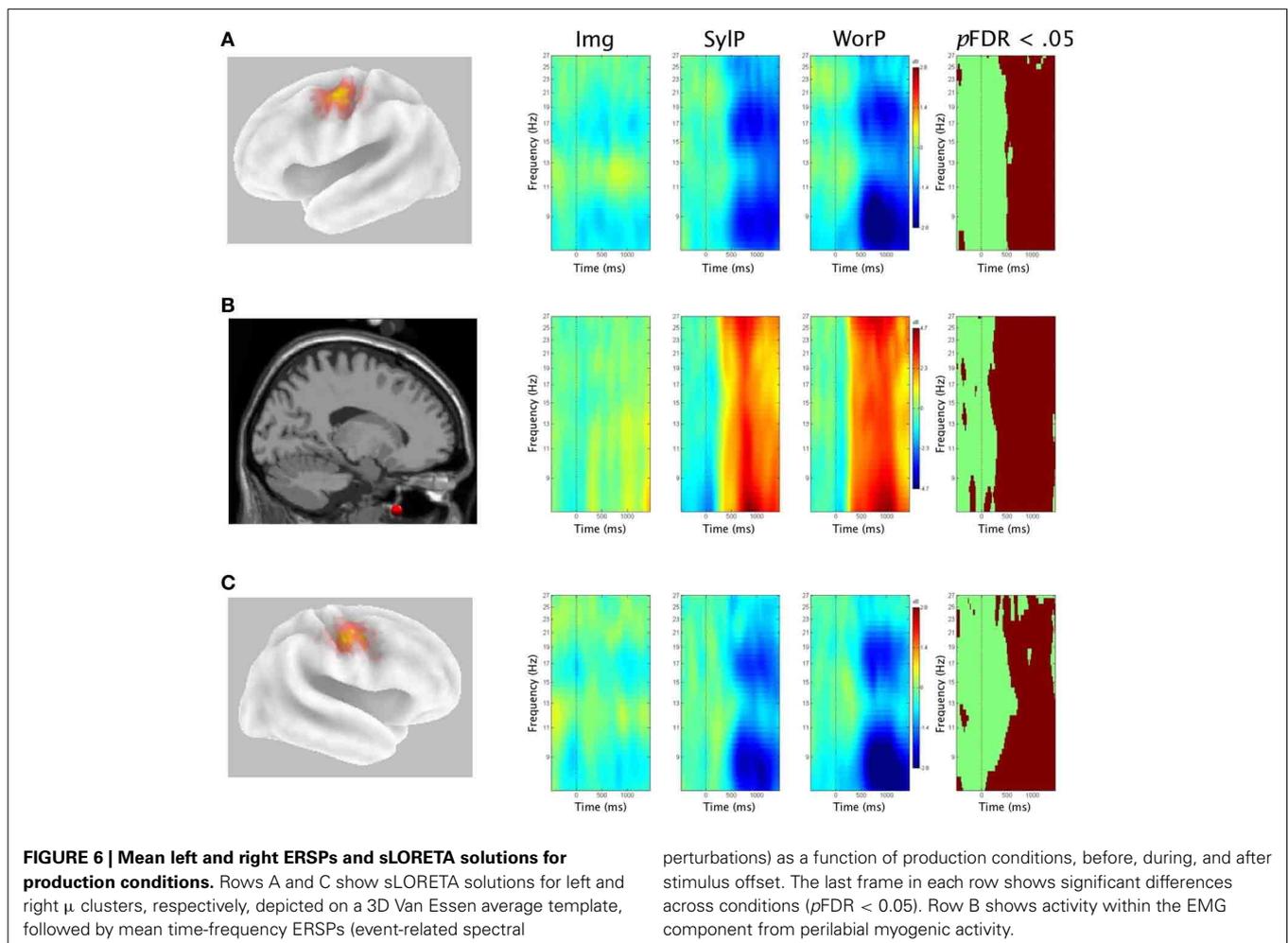
Significant EMG ERS (i.e., activity indicative of lip movement) in the SylP and WorP conditions began ∼300 ms after the cue to initiate speech. In both left and right μ clusters, alpha/beta ERD relative to baseline began in all production conditions up to 500 ms before the cue to speak. However, alpha/beta ERD in SylP and WorP conditions was significantly stronger (pFDR < 0.05) than in the Img condition during overt speech production

(i.e., coinciding with EMG activity). *Post-hoc* analyses in both left and right μ clusters showed no ERSP differences in SylP vs. WorP conditions.

As μ-ERD was significantly weaker in Img relative to overt production (SylP and WorP) conditions, ERSPs for all components contributing to left and right μ clusters were examined in the Img condition. On the left, only 8 of 17 participants displayed μ-ERD in this condition. The others either showed ERS or negligible change. On the right, 6 of 15 showed μ-ERD, 2 showed patterns of alpha ERS with beta ERD, and the others showed either ERS or negligible change.

DISCUSSION

In accord with the aims and hypotheses of this study, left and right μ components were identified across perception and production tasks. 85 and 75% of participants submitted components with ∼10% unexplained RV in left and right μ components, respectively. This proportion of useable μ components is similar to that found in other studies (e.g., Nystrom, 2008; Bowers et al., 2013), though the proportion of unexplained RV is slightly higher, possibly due to the inclusion of motor tasks. Bilateral localization of μ rhythm source maxima to the precentral gyrus with activity spreading across the premotor and

FIGURE 6 | Mean left and right ERSPs and sLORETA solutions for production conditions. Rows A and C show sLORETA solutions for left and right μ clusters, respectively, depicted on a 3D Van Essen average template, followed by mean time-frequency ERSPs (event-related spectral perturbations) as a function of production conditions, before, during, and after stimulus offset. The last frame in each row shows significant differences across conditions (pFDR < 0.05). Row B shows activity within the EMG component from perilabial myogenic activity.

sensorimotor cortices is consistent with accepted sources of the rhythm (Pineda, 2005; Hari, 2006) and important roles in speech perception/production (Skipper et al., 2007; Sato et al., 2009; Callan et al., 2010; Houde and Nagarajan, 2011; Tourville and Guenther, 2011). As these cortical sites are known to play important roles in SMI for both speech perception and production and μ rhythms are comprised of frequency bands that are sensitive to the demands of speech processing, this finding supports the subsequent examination of real-time activity within these clusters for better understanding the temporal dynamics of activity in the dorsal speech stream.

TIME-FREQUENCY ANALYSES IN PERCEPTION CONDITIONS

Similar to previous investigations (Makeig et al., 2004; Graimann and Pfurtscheller, 2006; Hari, 2006), anticipation of a button-press response in the Pasn condition yielded low-level increases from baseline in bilateral μ-ERD that gained slightly in strength with temporal proximity to the response. Discrimination conditions herein (Qdis and Ndis) also employed button-press responses, hence controlling for this effect in the statistical analysis. The Qdis and Ndis conditions produced similar highly accurate syllable discriminations and response reaction latencies such that differences between μ ERS/ERD in these conditions are attributable to the presence or absence of noise. Both conditions produced similar bilateral patterns of μ ERS/ERD that were generally stronger in the left hemisphere than the right, supporting left hemisphere dominance for SMI in speech perception (Hickok et al., 2011).

μ-alpha

Activity in the alpha band was characterized initially by ERS occurring prior to stimulus onset. ERS gradually gave way to ERD (**Figure 5**), with suppression first in low alpha (i.e., 8–10 Hz) and then high alpha (11–13 Hz). Alpha ERS was stronger and the transition occurred later in the Ndis condition than in the Qdis condition. Alpha rhythms are found globally across the cortex and their power can vary with numerous cognitive states and processes (Klimesch, 2012). Therefore, it is necessary to interpret alpha ERS/ERD relative to the tasks that induced them. Enhanced alpha (i.e., ERS) often is associated with cognitive load in working memory and attention tasks (Leiberg et al., 2006; Jensen et al., 2007; Haegens et al., 2010). It is thought to be an index of cortical inhibition of sensory information irrelevant to a given task, functioning to help sharpen attention to relevant information (Klimesch, 2012; Wilsch et al., 2014). In speech perception, this type of "active sensing" has been described in phenomena such as in the "cocktail party" effect, where specific attention to relevant speech cues helps filter similar competing background speech (Schroeder et al., 2010; Zion Golumbic et al., 2012).

Weisz et al. (2011) provide compelling evidence for an independently generated auditory alpha that is responsive to speech perception. Parsimonious with notions of increased cognitive load and consistent with the current findings, signal degradation of speech by noise vocoding has also been shown to enhance alpha activity (Obleser and Weisz, 2012). The observed differences in early alpha ERS between the Qdis and Ndis conditions support these notions. On the other hand, in speech perception tasks, alpha ERD has been found while evaluating speech (Shahin et al., 2009). Late occurring posterior alpha ERD has been related to increased speech intelligibility (Obleser and Weisz, 2012). Both of these findings are consistent with the notion of alpha ERD during accurate performance of perceptual and memory tasks (Klimesch et al., 2006). In addition, μ-alpha is suppressed in auditory speech perception tasks (Cuellar et al., 2012; Pineda et al., 2013). Hence, the current findings of late alpha ERD suggest that following stimulus offset, the two syllables were being evaluated by participants in the decision-making process.

μ-beta

In both discrimination conditions, significant beta ERD (relative to Pasn) was found across the time course of trials, prior to, during, and after acoustic stimulation. Beta ERD spread from narrow (17–19 Hz) to wide (15–30 Hz) beta bands while gaining in strength in both discrimination conditions. However, beta ERD occurred earlier in the Qdis than the Ndis condition. Bowers et al. (2013) previously showed that early beta ERD occurred when discriminating speech but not tones. They suggested that in speech perception tasks, early beta ERD also can be explained as a function of predictive coding. That is, internal models are posited to be generated in motor regions that are delivered to higher order auditory regions (i.e., superior temporal sulcus) to help constrain analysis and functionally improve speech discrimination accuracy (Callan et al., 2010; Bowers et al., 2013). These findings are also consistent with those of Mottonen et al. (2013), in that degraded conditions do no appear necessary to induce motor activity in speech discrimination. These models are thought to be available because of the considerable experience of humans generating the movements that produce these sounds. In addition, this predictive coding may have been fine tuned within the experiment. That is, requiring participants to accurately discriminate syllables 160 times (80 per condition) may have elicited anticipatory attention to speech processing.

μ-alpha and beta in discrimination conditions

The patterns of alpha and beta μ ERS/ERD found in quiet and noisy accurate speech discrimination need to be considered in combination. While similar patterns were observed in Qdis and Ndis conditions, stronger early alpha ERS was observed in the Ndis condition, which is consistent with the requirement of discriminating in noise. That is, it is speculated that the inhibitory mechanism was stronger when background noise was present. Conversely, early beta suppression appeared to be stronger in the Qdis than the Ndis condition. Though it is likely that internal models were generated in both conditions since they both used speech and were discriminated accurately (Bowers et al., 2013), it appears that in this study, the strong alpha ERS may have dominated the μ rhythm, extended into the low beta frequencies in the Ndis condition, and perhaps negated some early beta ERD. Together, these data suggest that alpha ERS and beta ERD within the sensorimotor μ rhythms work in unison, co-operating to functionally support accurate speech discrimination. This is further evidence that examination of the μ-rhythm provides a rich, time-sensitive, and relatively unique view of SMI in speech discrimination from an oscillatory perspective.

Dorsal stream motor activity in speech perception

The source location of μ clusters and their alpha and beta ERS/ERD suggest that they provide important information regarding sensorimotor dorsal stream activity in speech perception. The current findings suggest that beta activity may provide a measure of predictive coding via internal models (Bowers et al., 2013) generated in the PMC (Skipper et al., 2007; Houde and Nagarajan, 2011; Tourville and Guenther, 2011; Rauschecker, 2012). Tamura et al. (2012) investigated μ rhythm activity in various speech tasks, including covert production and production under different types of auditory feedback. They found differential activity within the alpha band and concluded that the μ-alpha was an index of auditory monitoring for speech. In line with this notion, it is speculated that μ-alpha might index sensory feedback into the PMC. Thus, stronger alpha ERS in the Ndis condition was observed, possibly due to a stronger inhibition of auditory feedback to the PMC when speech was presented in background noise. Furthermore, in the time period following stimulus offset and prior to the button press response, it seems likely that the two syllables were held in working memory, while being compared and covertly replayed during the decision-making process. These processes may require the generation of internal speech models and the disinhibition of feedback to the PMC, which would support the current findings of alpha and beta ERD in this time period within trials.

TIME-FREQUENCY ANALYSES IN PRODUCTION CONDITIONS

The covert (Img) and overt (SylP and WorP) production conditions yielded similar general patterns of alpha/beta ERD relative to baseline across trials. However, both alpha and beta ERD were significantly stronger in the overt production conditions than the Img condition, with significant differences in ERD following the cue to speak in both conditions. Across production conditions, there appeared to be little difference between right and left μ ERD. This is consistent with others that have found movement-induced bilateral decreases in beta suppression across the sensorimotor cortex (e.g., Salmelin and Hari, 1994; Pfurtscheller et al., 1996; Stančák and Pfurtscheller, 1996; Leocani et al., 1997, 2001; Alegre et al., 2003; Rau et al., 2003; Bai et al., 2005; Doyle et al., 2005; Erbil and Ungan, 2007). No differences between SylP and WorP conditions were observed. As expected, ERSP time-frequency analysis of perilabial EMG activity showed little activity in the Img condition, confirming that participants did not articulate the target syllables. In the SylP and WorP conditions, EMG activity following the "go" cue to speak was characterized predominantly by strong ERS beginning ~300 ms in both conditions. This time lag from the "go" cue is consistent with a normal movement reaction time. Hence, μ-alpha and μ-beta ERD showed temporal alignment to lip muscle movements.

μ-alpha during production

μ-alpha ERD in speech production is again interpreted as an index of feedback to the PMC while speech is being produced. By only measuring activity in the sensorimotor μ, it is not possible to differentiate between auditory and somatosensory feedback. μ-alpha suppression is traditionally localized to the somatosensory cortex and considered to reflect somatosensory activity (Hari, 2006). However, in light of recent findings in perception (Cuellar et al., 2012; Tamura et al., 2012; Pineda et al., 2013), there is mounting evidence that it also may reflect auditory feedback. In speech production, this makes sense considering how both auditory and somatosensory integration regions provide feedback to the PMC during speech production. Furthermore, the feedback from the auditory system and somatosensory system are generally consistent during speech production such that, barring perturbation to either modality, SFC models allow for them to often be considered unitarily (Houde and Nagarajan, 2011).

μ-beta during production

During overt production (SylP and WorP), beta μ-ERD is easily explained as a consequence of motor activity. Sensorimotor beta power has been ubiquitously found to suppress to motor activity from effectors including the fingers (Gaetz et al., 2010), wrist (Alegre et al., 2003), shoulder (Stančák et al., 2000), foot (Pfurtscheller and Lopes Da Silva, 1999), and tongue (Crone et al., 1998). However, if SFC models are applied, beta ERD can be cautiously interpreted as an index of PMC activity in the generation of feedforward control to motor effectors and forward internal models (efference copies) to the feedback loop. This interpretation is supported in a recent review (Engel and Fries, 2010; Kilavik et al., 2013), suggesting the difficulty in determining a clear functional role of sensorimotor beta suppression during movement, but that it may reflect sensory and cognitive aspects (e.g., forward modeling) in addition to pure motor processes. That said, one limitation of the current interpretation is the inability of beta ERD to distinguish between motor activity in feedforward (i.e., muscle movements) and feedback (i.e., internal modeling) mechanisms.

Covert speech production

The Img (covert production) produced significantly weaker alpha/beta ERD than the overt production conditions. This condition was incorporated into the design as previous work examining motor imagery and covert speech production had shown patterns of μ suppression and sensorimotor activity similar to overt productions (Pfurtscheller and Lopes Da Silva, 1999; Neuper et al., 2006). However, there is also evidence that responses in these covert conditions have been weaker than in actual overt productions (Neuper et al., 2006). In a recent study, Holler et al. (2013) investigated μ activity to real and imagined hand movements and showed that only 11 of 18 participants produced differences in μ- alpha/beta power when imagining hand movements. Of these 11, two showed μ enhancement rather than the suppression that was shown in the real movement conditions, suggesting variable responses to covert production tasks. The results in the current Img condition showed similar variability, perhaps contraindicating future use of covert production over a large number of repeated trials. This was the only condition in the experiment that required no overt response (either button-press or speech production) and hence, it was impossible to monitor the extent of covert syllable productions that was asked of participants 80 times in this condition.

Early μ-ERD in overt production

Significantly strong μ-ERD (relative to Img) was found as speech was being produced. Weaker μ-ERD was observed in SylP and WorP conditions prior to production and even before the cue to speak. This time period coincided with the preparation of the speech network, during which similar oscillatory activity has been reported (Gehrig et al., 2012; Herman et al., 2013). μ-ERD during this time period prior to production was weaker than expected, especially in light of the findings in the speech perception tasks, and predictions from SFC models (Houde and Nagarajan, 2011). This reduced neural ERD was most likely due to the influence of EMG on overall EEG variance.

THE UTILITY OF ICA IN SPEECH PERCEPTION AND PRODUCTION

μ components were successfully identified from band-pass filtered concatenated EEG data from perception and production conditions. Though the unexplained RV of the average μ ECD was slightly higher than has been found in other studies (e.g., Bowers et al., 2013), the combination of ECD/sLORETA CSD techniques produced a reliable and valid estimate of μ sources within the standard head model that was applied to all ICA data.

In the perception conditions, time-frequency analyses revealed differential contributions from alpha and beta bands of the μ rhythm that contributed to accurate syllable discrimination. μ-alpha/beta ERD was also revealed in speech production synchronized to muscle activity. This pattern of activity had not been described previously and can be interpreted as being consistent with "normal" sensorimotor control in speech production. Future investigations involving auditory or somatosensory speech perturbations (e.g., Bauer et al., 2006; Reilly and Dougherty, 2013) might be expected to reveal differences in alpha/beta ERD in speech production. Similarly, different relative patterns of μ ERS/ERD might be observed in clinical populations with compromised sensorimotor control such as in stuttering (Max et al., 2003; Loucks and De Nil, 2006; Watkins et al., 2008; Hickok et al., 2011; Cai et al., 2014; Connally et al., 2014).

In addition to the positive findings, there was also evidence of drawbacks to using EEG/ICA in production tasks. It was clear that ICA adequately separated neural from non-neural (e.g., myogenic) activity. Had this not been successfully accomplished, μ-ERD/ERS during production would likely have been overwhelmed by the EMG activity. However, it also appears that overall spectral power in μ components was reduced in the production tasks due to a greater proportion of the overall EEG variance that had to be accounted for by EMG activity. Considering motor requirements, strongest μ-ERD (especially beta) would have been expected in production conditions. However, even when at their strongest, spectral powers during production did not exceed those in perception. In addition, only weak μ-ERD was noted in the time period prior to overt production, which was expected to be stronger as the speech networks prepared to articulate. Together, these findings indicate that overall spectral power in production conditions was attenuated. As such, though interesting general patterns of μ-ERD were revealed in speech production, they should be interpreted with caution with respect to their sensitivity and without making reference to function in conditions without motor requirements.

Another limitation in the current methods was the inability to observe μ activity following speech production. Though production targets (i.e., syllables and words) were produced within the time course of trials, EMG activity (e.g., lip movement) persisted past production, such that the epoch length that did not allow for the measurement of beta rebound (i.e., ERS), which is commonly observed following termination of a movement (Kilavik et al., 2013).

CONCLUSIONS AND FUTURE DIRECTIONS

ICA successfully identified μ components in speech perception and production. Time-frequency analyses using ERSP showed real-time changes in alpha/beta power that provided indicators of PMC/sensorimotor contributions to speech-based dorsal stream activity. Localization of μ clusters and ERSP activity in perception and production are in agreement with Rauschecker's (2011) observation that, based on connections to the inferior parietal lobe and posterior auditory cortex, the PMC provides "optimal state estimation" for speech.

Sensitivity of the findings was somewhat reduced in production conditions, most likely due to concomitant myogenic activity. Further applications in speech production might consider additional filtering techniques in addition to ICA. Exquisite temporal resolution combined with economy and availability warrant further use of ICA particularly to understand speech processing in normal and clinical populations. While measuring the temporal dynamics of the μ-rhythm provide rich information about sensorimotor processing, future ICA studies may also investigate multiple components within the speech processing network in addition to measuring connectivity (i.e., coherence) between components.

REFERENCES

Adank, P. (2012). The neural bases of difficult speech comprehension and speech production: two Activation Likelihood Estimation (ALE) meta-analyses. *Brain Lang.* 122, 42–54. doi: 10.1016/j.bandl.2012.04.014

Alegre, M., Gurtubay, I. G., Labarga, A., Iriarte, J., Malanda, A., and Artieda, J. (2003). Alpha and beta oscillatory changes during stimulus-induced movement paradigms: effect of stimulus predictability. *Neuroreport* 14, 381–385. doi: 10.1097/01.wnr.0000059624.96928.c0

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

Arnal, L. H. (2012). Predicting "When" using the motor system's beta-band oscillations. *Front. Hum. Neurosci.* 6:225. doi: 10.3389/fnhum.2012.00225

Arnal, L. H., and Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* 16, 390–398. doi: 10.1016/j.tics.2012.05.003

Bai, O., Mari, Z., Vorbach, S., and Hallett, M. (2005). Asymmetric spatiotemporal patterns of event-related desynchronization preceding voluntary sequential finger movements: a high-resolution EEG study. *Clin. Neurophysiol.* 116, 1213–1221. doi: 10.1016/j.clinph.2005.01.006

Bauer, J. J., Mittal, J., Larson, C. R., and Hain, T. C. (2006). Vocal responses to unanticipated perturbations in voice loudness feedback: an automatic mechanism for stabilizing voice amplitude. *J. Acoust. Soc. Am.* 119, 2363–2371. doi: 10.1121/1.2173513

Benjamini, Y., and Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J. Educ. Behav. Stat.* 25, 60–83. doi: 10.3102/10769986025001060

Bickel, S., Dias, E. C., Epstein, M. L., and Javitt, D. C. (2012). Expectancy-related modulations of neural oscillations in continuous performance tasks. *Neuroimage* 62, 1867–1876. doi: 10.1016/j.neuroimage.2012.06.009

Bidet-Caulet, A., Barbe, P. G., Roux, S., Viswanath, H., Barthelemy, C., Bruneau, N., et al. (2012). Dynamics of anticipatory mechanisms during predictive context processing. *Eur. J. Neurosci.* 36, 2996–3004. doi: 10.1111/j.1460-9568.2012.08223.x

Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., and Ward, B. D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nat. Neurosci.* 7, 295–301. doi: 10.1038/nn1198

Blockcolsky, V., Frazer, J., and Frazer, D. (2008). *40,000 Selected Words.* San Antonio, TX: Communication Skill Builders.

Bowers, A., Saltuklaroglu, T., Harkrider, A., and Cuellar, M. (2013). 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:e72024. doi: 10.1371/journal.pone.0072024

Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *J. Acoust. Soc. Am.* 103, 3153–3161. doi: 10.1121/1.423073

Burton, M. W., and Small, S. L. (2006). Functional neuroanatomy of segmenting speech and nonspeech. *Cortex* 42, 644–651. doi: 10.1016/S0010-9452(08)70400-3

Burton, M. W., Small, S. L., and Blumstein, S. E. (2000). The role of segmentation in phonological processing: an fMRI investigation. *J. Cogn. Neurosci.* 12, 679–690. doi: 10.1162/089892900562309

Cai, S., Beal, D. S., Ghosh, S. S., Guenther, F. H., and Perkell, J. S. (2014). Impaired timing adjustments in response to time-varying auditory perturbation during connected speech production in persons who stutter. *Brain Lang.* 129, 24–29. doi: 10.1016/j.bandl.2014.01.002

Callan, A. M., Callan, D. E., Tajima, K., and Akahane-Yamada, R. (2006). Neural processes involved with perception of non-native durational contrasts. *Neuroreport* 17, 1353–1357. doi: 10.1097/01.wnr.0000224774.66904.29

Callan, D., Callan, A., Gamez, M., Sato, M. A., and Kawato, M. (2010). Premotor cortex mediates perceptual performance. *Neuroimage* 51, 844–858. doi: 10.1016/j.neuroimage.2010.02.027

Cogan, G. B., Thesen, T., Carlson, C., Doyle, W., Devinsky, O., and Pesaran, B. (2014). Sensory-motor transformations for speech occur bilaterally. *Nature* 507, 94–98. doi: 10.1038/nature12935

Connally, E. L., Ward, D., Howell, P., and Watkins, K. E. (2014). Disrupted white matter in language and motor tracts in developmental stuttering. *Brain Lang.* 131, 25–35. doi: 10.1016/j.bandl.2013.05.013

Costa, A., Strijkers, K., Martin, C., and Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21442–21446. doi: 10.1073/pnas.0908921106

Crawcour, S., Bowers, A., Harkrider, A., and Saltuklaroglu, T. (2009). Mu wave suppression during the perception of meaningless syllables: EEG evidence of motor recruitment. *Neuropsychologia* 47, 2558–2563. doi: 10.1016/j.neuropsychologia.2009.05.001

Crone, N. E., Miglioretti, D. L., Gordon, B., Sieracki, J. M., Wilson, M. T., Uematsu, S., et al. (1998). Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. I. Alpha and beta event-related desynchronization. *Brain* 121(Pt. 12), 2271–2299. doi: 10.1093/brain/121.12.2271

Cuellar, M., Bowers, A., Harkrider, A. W., Wilson, M., and Saltuklaroglu, T. (2012). Mu suppression as an index of sensorimotor contributions to speech processing: evidence from continuous EEG signals. *Int. J. Psychophysiol.* 85, 242–248. doi: 10.1016/j.ijpsycho.2012.04.003

D'Ausilio, A., Bufalari, I., Salmas, P., and Fadiga, L. (2012). The role of the motor system in discriminating normal and degraded speech sounds. *Cortex* 48, 882–887. doi: 10.1016/j.cortex.2011.05.017

D'Ausilio, A., Pulvermuller, F., Salmas, P., Bufalari, I., Begliomini, C., and Fadiga, L. (2009). The motor somatotopy of speech perception. *Curr. Biol.* 19, 381–385. doi: 10.1016/j.cub.2009.01.017

Dell'acqua, R., Sessa, P., Peressotti, F., Mulatti, C., Navarrete, E., and Grainger, J. (2010). ERP evidence for ultra-fast semantic processing in the picture-word interference paradigm. *Front. Psychol.* 1:177. doi: 10.3389/fpsyg.2010.00177

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Delorme, A., Palmer, J., Onton, J., Oostenveld, R., and Makeig, S. (2012). Independent EEG sources are dipolar. *PLoS ONE* 7:e30135. doi: 10.1371/journal.pone.0030135

Doyle, L. M., Yarrow, K., and Brown, P. (2005). Lateralization of event-related beta desynchronization in the EEG during pre-cued reaction time tasks. *Clin. Neurophysiol.* 116, 1879–1888. doi: 10.1016/j.clinph.2005.03.017

Engel, A. K., and Fries, P. (2010). Beta-band oscillations–signaling the status quo? *Curr. Opin. Neurobiol.* 20, 156–165. doi: 10.1016/j.conb.2010.02.015

Erbil, N., and Ungan, P. (2007). Changes in the alpha and beta amplitudes of the central EEG during the onset, continuation, and offset of long-duration repetitive hand movements. *Brain Res.* 1169, 44–56. doi: 10.1016/j.brainres.2007.07.014

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *J. Phon.* 14, 3–28.

Gaetz, W., Macdonald, M., Cheyne, D., and Snead, O. C. (2010). Neuromagnetic imaging of movement-related cortical oscillations in children and adults: age predicts post-movement beta rebound. *Neuroimage* 51, 792–807. doi: 10.1016/j.neuroimage.2010.01.077

Gallese, V., Gernsbacher, M. A., Heyes, C., Hickok, G., and Davis, M. H. (2011). Mirror neuron forum. *Pespect. Psychol. Sci.* 71, 1138–1149. doi: 10.1177/1745691611413392

Ganushchak, L. Y., Christoffels, I. K., and Schiller, N. O. (2011). The use of electroencephalography in language production research: a review. *Front. Psychol.* 2:208. doi: 10.3389/fpsyg.2011.00208

Gehrig, J., Wibral, M., Arnold, C., and Kell, C. A. (2012). Setting up the speech production network: how oscillations contribute to lateralized information routing. *Front. Psychol.* 3:169. doi: 10.3389/fpsyg.2012.00169

Geranmayeh, F., Brownsett, S. L., Leech, R., Beckmann, C. F., Woodhead, Z., and Wise, R. J. (2012). The contribution of the inferior parietal cortex to spoken language production. *Brain Lang.* 121, 47–57. doi: 10.1016/j.bandl.2012.02.005

Ghitza, O., Giraud, A. L., and Poeppel, D. (2012). Neuronal oscillations and speech perception: critical-band temporal envelopes are the essence. *Front. Hum. Neurosci.* 6:340. doi: 10.3389/fnhum.2012.00340

Giraud, A. L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15, 511–517. doi: 10.1038/nn.3063

Gladwin, T. E., Lindsen, J. P., and de Jong, R. (2006). Pre-stimulus EEG effects related to response speed, task switching and upcoming response hand. *Biol. Psychol.* 72, 15–34. doi: 10.1016/j.biopsycho.2005.05.005

Golfinopoulos, E., Tourville, J. A., Bohland, J. W., Ghosh, S. S., Nieto-Castanon, A., and Guenther, F. H. (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *Neuroimage* 55, 1324–1338. doi: 10.1016/j.neuroimage.2010.12.065

Grabski, K., Tremblay, P., Gracco, V. L., Girin, L., and Sato, M. (2013). A mediating role of the auditory dorsal pathway in selective adaptation to speech: a state-dependent transcranial magnetic stimulation study. *Brain Res.* 1515, 55–65. doi: 10.1016/j.brainres.2013.03.024

Graimann, B., and Pfurtscheller, G. (2006). Quantification and visualization of event-related changes in oscillatory brain activity in the time—frequency domain. *Prog. Brain Res.* 159, 79–97. doi: 10.1016/S0079-6123(06)59006-5

Gramann, K., Gwin, J. T., Bigdely-Shamlo, N., Ferris, D. P., and Makeig, S. (2010). Visual evoked responses during standing and walking. *Front. Hum. Neurosci.* 4:202. doi: 10.3389/fnhum.2010.00202

Grin-Yatsenko, V. A., Baas, I., Ponomarev, V. A., and Kropotov, J. D. (2010). Independent component approach to the analysis of EEG recordings at early stages of depressive disorders. *Clin. Neurophysiol.* 121, 281–289. doi: 10.1016/j.clinph.2009.11.015

Guenther, F. H., and Vladusich, T. (2012). A neural theory of speech acquisition and production. *J. Neurolinguist.* 25, 408–422. doi: 10.1016/j.jneuroling.2009.08.006

Gwin, J. T., and Ferris, D. (2011). High-density EEG and independent component analysis mixture models distinguish knee contractions from ankle contractions. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2011, 4195–4198. doi: 10.1109/IEMBS.2011.6091041

Gwin, J. T., Gramann, K., Makeig, S., and Ferris, D. P. (2010). Removal of movement artifact from high-density EEG recorded during walking and running. *J. Neurophysiol.* 103, 3526–3534. doi: 10.1152/jn.00105.2010

Haegens, S., Osipova, D., Oostenveld, R., and Jensen, O. (2010). Somatosensory working memory performance in humans depends on both engagement and

disengagement of regions in a distributed network. *Hum. Brain Mapp.* 31, 26–35. doi: 10.1002/hbm.20842

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

Herman, A. B., Houde, J. F., Vinogradov, S., and 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. doi: 10.1523/JNEUROSCI.1472-12.2013

Hickok, G. (2009). The functional neuroanatomy of language. *Phys. Life Rev.* 6, 121–143. doi: 10.1016/j.plrev.2009.06.001

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nat. Rev. Neurosci.* 13, 135–145. doi: 10.1038/nrn3158

Hickok, G. (2014). The architecture of speech production and the role of the phoneme in speech processing. *Lang. Cogn. Process.* 29, 2–20. doi: 10.1080/01690965.2013.834370

Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69, 407–422. doi: 10.1016/j.neuron.2011.01.019

Hickok, G., and Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99. doi: 10.1016/j.cognition.2003.10.011

Hirschfeld, G., Jansma, B., Bolte, J., and Zwitserlood, P. (2008). Interference and facilitation in overt speech production investigated with event-related potentials. *Neuroreport* 19, 1227–1230. doi: 10.1097/WNR.0b013e328309ecd1

Holler, Y., Bergmann, J., Kronbichler, M., Crone, J. S., Schmid, E. V., Thomschewski, A., et al. (2013). Real movement vs. motor imagery in healthy subjects. *Int. J. Psychophysiol.* 87, 35–41. doi: 10.1016/j.ijpsycho.2012.10.015

Houde, J. F., and Nagarajan, S. S. (2011). Speech production as state feedback control. *Front. Hum. Neurosci.* 5:82. doi: 10.3389/fnhum.2011.00082

Jasper, H. H. (1958). The ten twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.* 10, 371–375.

Jensen, O., Kaiser, J., and Lachaux, J. P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci.* 30, 317–324. doi: 10.1016/j.tins.2007.05.001

Kilavik, B. E., Zaepffel, M., Brovelli, A., Mackay, W. A., and Riehle, A. (2013). The ups and downs of beta oscillations in sensorimotor cortex. *Exp. Neurol.* 245, 15–26. doi: 10.1016/j.expneurol.2012.09.014

Klimesch, W. (2012). alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn. Sci.* 16, 606–617. doi: 10.1016/j.tics.2012.10.007

Klimesch, W., Doppelmayr, M., and Hanslmayr, S. (2006). Upper alpha ERD and absolute power: their meaning for memory performance. *Prog. Brain Res.* 159, 151–165. doi: 10.1016/S0079-6123(06)59010-7

Lau, T. M., Gwin, J. T., and Ferris, D. P. (2014). Walking reduces sensorimotor network connectivity compared to standing. *J. Neuroeng. Rehabil.* 11, 14. doi: 10.1186/1743-0003-11-14

Lee, T. W., Girolami, M., and Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput.* 11, 417–441. doi: 10.1162/089976699300016719

Leiberg, S., Lutzenberger, W., and Kaiser, J. (2006). Effects of memory load on cortical oscillatory activity during auditory pattern working memory. *Brain Res.* 1120, 131–140. doi: 10.1016/j.brainres.2006.08.066

Leocani, L., Toro, C., Manganotti, P., Zhuang, P., and Hallett, M. (1997). Event-related coherence and event-related desynchronization/synchronization in the 10 Hz and 20 Hz EEG during self-paced movements. *Electroencephalogr. Clin. Neurophysiol.* 104, 199–206. doi: 10.1016/S0168-5597(96)96051-7

Leocani, L., Toro, C., Zhuang, P., Gerloff, C., and Hallett, M. (2001). Event-related desynchronization in reaction time paradigms: a comparison with event-related potentials and corticospinal excitability. *Clin. Neurophysiol.* 112, 923–930. doi: 10.1016/S1388-2457(01)00530-2

Locasto, P. C., Krebs-Noble, D., Gullapalli, R. P., and Burton, M. W. (2004). An fMRI investigation of speech and tone segmentation. *J. Cogn. Neurosci.* 16, 1612–1624. doi: 10.1162/0898929042568433

Loucks, T. M., and De Nil, L. F. (2006). Anomalous sensorimotor integration in adults who stutter: a tendon vibration study. *Neurosci. Lett.* 402, 195–200. doi: 10.1016/j.neulet.2006.04.002

Makeig, S., Debener, S., Onton, J., and Delorme, A. (2004). Mining event-related brain dynamics. *Trends Cogn. Sci.* 8, 204–210. doi: 10.1016/j.tics.2004.03.008

Max, L., Caruso, A. J., and Gracco, V. L. (2003). Kinematic analyses of speech, orofacial nonspeech, and finger movements in stuttering and nonstuttering adults. *J. Speech Lang. Hear. Res.* 46, 215–232. doi: 10.1044/1092-4388(2003/017)

McMenamin, B. W., Shackman, A. J., Greischar, L. L., and Davidson, R. J. (2011). Electromyogenic artifacts and electroencephalographic inferences revisited. *Neuroimage* 54, 4–9. doi: 10.1016/j.neuroimage.2010.07.057

McMenamin, B. W., Shackman, A. J., Maxwell, J. S., Bachhuber, D. R., Koppenhaver, A. M., Greischar, L. L., et al. (2010). Validation of ICA-based myogenic artifact correction for scalp and source-localized EEG. *Neuroimage* 49, 2416–2432. doi: 10.1016/j.neuroimage.2009.10.010

McMenamin, B. W., Shackman, A. J., Maxwell, J. S., Greischar, L. L., and Davidson, R. J. (2009). Validation of regression-based myogenic correction techniques for scalp and source-localized EEG. *Psychophysiology* 46, 578–592. doi: 10.1111/j.1469-8986.2009.00787.x

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., and Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Curr. Biol.* 17, 1692–1696. doi: 10.1016/j.cub.2007.08.064

Mottonen, R., and Watkins, K. E. (2012). Using TMS to study the role of the articulatory motor system in speech perception. *Aphasiology* 26, 1103–1118. doi: 10.1080/02687038.2011.619515

Mottonen, R., Dutton, R., and Watkins, K. E. (2013). Auditory-motor processing of speech sounds. *Cereb. Cortex* 23, 1190–1197. doi: 10.1093/cercor/bhs110

Moulin-Frier, C., and Arbib, M. A. (2013). Recognizing speech in a novel accent: the motor theory of speech perception reframed. *Biol. Cybern.* 107, 421–447. doi: 10.1007/s00422-013-0557-3

Murakami, T., Ugawa, Y., and Ziemann, U. (2013). Utility of TMS to understand the neurobiology of speech. *Front. Psychol.* 4:446. doi: 10.3389/fpsyg.2013.00446

Muthukumaraswamy, S. D., and Johnson, B. W. (2004). Primary motor cortex activation during action observation revealed by wavelet analysis of the EEG. *Clin. Neurophysiol.* 115, 1760–1766. doi: 10.1016/j.clinph.2004.03.004

Neuper, C., Wortz, M., and Pfurtscheller, G. (2006). ERD/ERS patterns reflecting sensorimotor activation and deactivation. *Prog. Brain Res.* 159, 211–222. doi: 10.1016/s0079-6123(06)59014-4

Nystrom, P. (2008). The infant mirror neuron system studied with high density EEG. *Soc. Neurosci.* 3, 334–347. doi: 10.1080/17470910701563665

Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S., and Pineda, J. A. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Brain Res. Cogn. Brain Res.* 24, 190–198. doi: 10.1016/j.cogbrainres.2005.01.014

Obleser, J., and Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb. Cortex* 22, 2466–2477. doi: 10.1093/cercor/bhr325

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Onton, J., Westerfield, M., Townsend, J., and Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neurosci. Biobehav. Rev.* 30, 808–822. doi: 10.1016/j.neubiorev.2006.06.007

Oostenveld, R., and Oostendorp, T. F. (2002). Validating the boundary element method for forward and inverse EEG computations in the presence of a hole in the skull. *Hum. Brain Mapp.* 17, 179–192. doi: 10.1002/hbm.10061

Osnes, B., Hugdahl, K., and Specht, K. (2011). Effective connectivity analysis demonstrates involvement of premotor cortex during speech perception. *Neuroimage* 54, 2437–2445. doi: 10.1016/j.neuroimage.2010.09.078

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find. Exp. Clin. Pharmacol.* 24(Suppl. D), 5–12.

Perrier, P., Ostry, D. J., and Laboissiere, R. (1996). The equilibrium point hypothesis and its application to speech motor control. *J. Speech Hear. Res.* 39, 365–378.

Perry, A., and Bentin, S. (2010). Does focusing on hand-grasping intentions modulate electroencephalogram mu and alpha suppressions? *Neuroreport* 21, 1050–1054. doi: 10.1097/WNR.0b013e32833fcb71

Pfurtscheller, G., and Lopes Da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110, 1842–1857. doi: 10.1016/S1388-2457(99)00141-8

Pfurtscheller, G., Stancak, A. Jr., and Neuper, C. (1996). Post-movement beta synchronization. A correlate of an idling motor area? *Electroencephalogr. Clin. Neurophysiol.* 98, 281–293. doi: 10.1016/0013-4694(95)00258-8

Pickering, M. J., and Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends Cogn. Sci.* 11, 105–110. doi: 10.1016/j.tics.2006.12.002

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

Pineda, J. A., Grichanik, M., Williams, V., Trieu, M., Chang, H., and Keysers, C. (2013). EEG sensorimotor correlates of translating sounds into actions. *Front. Neurosci.* 7:203. doi: 10.3389/fnins.2013.00203

Purcell, D. W., and Munhall, K. G. (2006). Compensation following real-time manipulation of formants in isolated vowels. *J. Acoust. Soc. Am.* 119, 2288–2297. doi: 10.1121/1.2173514

Rau, C., Plewnia, C., Hummel, F., and Gerloff, C. (2003). Event-related desynchronization and excitability of the ipsilateral motor cortex during simple self-paced finger movements. *Clin. Neurophysiol.* 114, 1819–1826. doi: 10.1016/S1388-2457(03)00174-3

Rauschecker, J. P. (2011). An expanded role for the dorsal pathway in sensorimotor control and integration. *Hear. Res.* 27, 16–25. doi: 10.1016/j.heares.2010.09.001

Rauschecker, J. P. (2012). Ventral and dorsal streams in the evolution of speech and language. *Front. Evol. Neurosci.* 4:7. doi: 10.3389/fnevo.2012.00007

Reilly, K. J., and Dougherty, K. E. (2013). The role of vowel perceptual cues in compensatory responses to perturbations of speech auditory feedback. *J. Acoust. Soc. Am.* 134, 1314–1323. doi: 10.1121/1.4812763

Rogalsky, C., Love, T., Driscoll, D., Anderson, S. W., and Hickok, G. (2011). Are mirror neurons the basis of speech perception? Evidence from five cases with damage to the purported human mirror system. *Neurocase* 17, 178–187. doi: 10.1080/13554794.2010.509318

Salmelin, R., and Hari, R. (1994). Spatiotemporal characteristics of sensorimotor neuromagnetic rhythms related to thumb movement. *Neuroscience* 60, 537–550. doi: 10.1016/0306-4522(94)90263-1

Sams, M., Mottonen, R., and Sihvonen, T. (2005). Seeing and hearing others and oneself talk. *Brain Res. Cogn. Brain Res.* 23, 429–435. doi: 10.1016/j.cogbrainres.2004.11.006

Sato, M., Tremblay, P., and Gracco, V. L. (2009). A mediating role of the premotor cortex in phoneme segmentation. *Brain Lang.* 111, 1–7. doi: 10.1016/j.bandl.2009.03.002

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., and Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Curr. Opin. Neurobiol.* 20, 172–176. doi: 10.1016/j.conb.2010.02.010

Scott, S. K., and Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends Neurosci.* 26, 100–107. doi: 10.1016/S0166-2236(02)00037-1

Scott, S. K., McGettigan, C., and Eisner, F. (2009). A little more conversation, a little less action—candidate roles for the motor cortex in speech perception. *Nat. Rev. Neurosci.* 10, 295–302. doi: 10.1038/nrn2603

Shackman, A. J., McMenamin, B. W., Slagter, H. A., Maxwell, J. S., Greischar, L. L., and Davidson, R. J. (2009). Electromyogenic artifacts and electroencephalographic inferences. *Brain Topogr.* 22, 7–12. doi: 10.1007/s10548-009-0079-4

Shahin, A. J., Picton, T. W., and Miller, L. M. (2009). Brain oscillations during semantic evaluation of speech. *Brain Cogn.* 70, 259–266. doi: 10.1016/j.bandc.2009.02.008

Shou, G., Ding, L., and Dasari, D. (2012). Probing neural activations from continuous EEG in a real-world task: time-frequency independent component analysis. *J. Neurosci. Methods* 209, 22–34. doi: 10.1016/j.jneumeth.2012.05.022

Simmonds, A. J., Wise, R. J., Collins, C., Redjep, O., Sharp, D. J., Iverson, P., et al. (2014). Parallel systems in the control of speech. *Hum. Brain Mapp.* 35, 1930–1943. doi: 10.1002/hbm.22303

Skipper, J. I., Van Wassenhove, V., Nusbaum, H. C., and Small, S. L. (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cereb. Cortex* 17, 2387–2399. doi: 10.1093/cercor/bhl147

Sohoglu, E., Peelle, J. E., Carlyon, R. P., and Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *J. Neurosci.* 32, 8443–8453. doi: 10.1523/jneurosci.5069-11.2012

Specht, K. (2013). Mapping a lateralization gradient within the ventral stream for auditory speech perception. *Front. Hum. Neurosci.* 7:629. doi: 10.3389/fnhum.2013.00629

Specht, K. (2014). Neuronal basis of speech comprehension. *Hear. Res.* 307, 121–135. doi: 10.1016/j.heares.2013.09.011

Stančák, A. Jr., and Pfurtscheller, G. (1996). Event-related desynchronisation of central beta-rhythms during brisk and slow self-paced finger movements of dominant and nondominant hand. *Brain Res. Cogn. Brain Res.* 4, 171–183. doi: 10.1016/S0926-6410(96)00031-6

Stančák, A. Jr., Feige, B., Lucking, C. H., and Kristeva-Feige, R. (2000). Oscillatory cortical activity and movement-related potentials in proximal and distal movements. *Clin. Neurophysiol.* 111, 636–650. doi: 10.1016/S1388-2457(99)00310-7

Stevens, K. N., and Halle, M. (1967). *Remarks on Analysis by Synthesis and Distinctive Features.* Cambridge, MA: MIT press.

Strijkers, K., Costa, A., and Thierry, G. (2010). Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. *Cereb. Cortex* 20, 912–928. doi: 10.1093/cercor/bhp153

Stuart, A., Kalinowski, J., Rastatter, M. P., and Lynch, K. (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *J. Acoust. Soc. Am.* 111, 2237–2241. doi: 10.1121/1.1466868

Szenkovits, G., Peelle, J. E., Norris, D., and Davis, M. H. (2012). Individual differences in premotor and motor recruitment during speech perception. *Neuropsychologia* 50, 1380–1392. doi: 10.1016/j.neuropsychologia.2012.02.023

Tamura, T., Gunji, A., Takeichi, H., Shigemasu, H., Inagaki, M., Kaga, M., et al. (2012). Audio-vocal monitoring system revealed by mu-rhythm activity. *Front. Psychol.* 3:225. doi: 10.3389/fpsyg.2012.00225

Tian, X., and Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Front. Psychol.* 1:166. doi: 10.3389/fpsyg.2010.00166

Tian, X., and Poeppel, D. (2012). Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation. *Front. Hum. Neurosci.* 6:314. doi: 10.3389/fnhum.2012.00314

Tourville, J. A., and Guenther, F. H. (2011). The DIVA model: a neural theory of speech acquisition and production. *Lang. Cogn. Process.* 26, 952–981. doi: 10.1080/01690960903498424

Towle, V. L., Bolanos, J., Suarez, D., Tan, K., Grzeszczuk, R., Levin, D. N., et al. (1993). The spatial location of EEG electrodes: locating the best-fitting sphere relative to cortical anatomy. *Electroencephalogr. Clin. Neurophysiol.* 86, 1–6. doi: 10.1016/0013-4694(93)90061-Y

Tran, Y., Craig, A., Boord, P., and Craig, D. (2004). Using independent component analysis to remove artifact from electroencephalographic measured during stuttered speech. *Med. Biol. Eng. Comput.* 42, 627–633. doi: 10.1007/BF02347544

Venezia, J. H., Saberi, K., Chubb, C., and Hickok, G. (2012). Response bias modulates the speech motor system during syllable discrimination. *Front. Psychol.* 3:157. doi: 10.3389/fpsyg.2012.00157

Ventura, M. I., Nagarajan, S. S., and Houde, J. F. (2009). Speech target modulates speaking induced suppression in auditory cortex. *BMC Neurosci.* 10:58. doi: 10.1186/1471-2202-10-58

Watkins, K. E., Smith, S. M., Davis, S., and Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain* 131, 50–59. doi: 10.1093/brain/awm241

Weisz, N., Hartmann, T., Muller, N., Lorenz, I., and Obleser, J. (2011). Alpha rhythms in audition: cognitive and clinical perspectives. *Front. Psychol.* 2:73. doi: 10.3389/fpsyg.2011.00073

Wilsch, A., Henry, M. J., Herrmann, B., Maess, B., and Obleser, J. (2014). Alpha oscillatory dynamics index temporal expectation benefits in working memory. *Cereb. Cortex.* doi: 10.1093/cercor/bhu004. [Epub ahead of print].

Wilson, S. M., Saygin, A. P., Sereno, M. I., and Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nat. Neurosci.* 7, 701–702. doi: 10.1038/nn1263

Zaepffel, M., Trachel, R., Kilavik, B. E., and Brochier, T. (2013). Modulations of EEG beta power during planning and execution of grasping movements. *PLoS ONE* 8:e60060. doi: 10.1371/journal.pone. 0060060

Zion Golumbic, E. M., Poeppel, D., and Schroeder, C. E. (2012). Temporal context in speech processing and attentional stream selection: a behavioral and neural perspective. *Brain Lang.* 122, 151–161. doi: 10.1016/j.bandl.2011. 12.010

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.