Experimental Brain Research (2019) 237:705–721
https://doi.org/10.1007/s00221-018-5447-4

RESEARCH ARTICLE

Power and phase coherence in sensorimotor mu and temporal lobe alpha components during covert and overt syllable production

Andrew Bowers[1] · **Tim Saltuklaroglu**[2] · **David Jenson**[3] · **Ashley Harkrider**[2] · **David Thornton**[4]

Received: 16 May 2018 / Accepted: 30 November 2018 / Published online: 14 December 2018

Abstract

The sensorimotor dorsal stream is known to activate in both overt and covert speech production. However, overt production produces sensory consequences that are absent during covert production. Thus, the purpose of the current study is to investigate differences in dorsal stream activity between these two production conditions across the time course of utterances. Electroencephalography (EEG) was recorded from 68 channels while 23 participants overtly (Op) and covertly (Cp) produced orthographically cued bisyllabic targets. Sensorimotor mu and auditory alpha components (from anterior and posterior aspects of the dorsal stream) were identified using independent component analysis (ICA). Event-related spectral perturbation (ERSP) analyses identified changes in mu and alpha oscillatory power over time, while intercomponent phase coherence (IPC) measured anterior–posterior connectivity in the two conditions. Results showed greater beta (15–25 Hz) suppression during speech planning across left and right hemisphere sensorimotor and temporal ICs for Op relative to Cp. By contrast, greater intrahemispheric beta coherence was observed for Cp compared to Op during speech planning. During execution, greater beta suppression was observed along with greater low frequency (< 10 Hz) power enhancement and intrahemispheric phase coherence in Op compared to Cp. The findings implicate low frequency sensorimotor and posterior temporal phase coherence in the integration of somatosensory and acoustic feedback in overt relative to covert execution. Findings are consistent with early frontal–temporal forward models involved in planning and execution with modulations depending on whether the task goal is internal or overt syllable production.

Keywords Neural oscillations · Mu rhythm · Auditory alpha rhythm · Internal models

Introduction

Dual-stream models of language posit that a left-lateralized dorsal sensorimotor network mediates speech production and the conversion of auditory information into the motor commands required to produce auditory targets (Hickok and Poeppel 2007). In the dual-stream model, internal models of motor control initiated in premotor and motor regions internally simulate the sensory consequences of speech and subsequently compare those consequences with overt sensory outcomes in somatosensory and auditory regions (Hickok et al. 2011; Tian and Poeppel 2012; Hickok 2012; Guenther and Hickok 2015). Consistent with those models, neuroimaging studies have demonstrated that common neural substrates mediate both internal imagery of speech (i.e., covert speech production; Cp) and overt production (Op) (Price 2012). Further, despite the absence of true sensory consequences, subjects engaged in imagined speech production experience quasi-perceptual internal auditory

✉ Andrew Bowers
albowers@uark.edu

[1] Department of Communication Disorders, Epley Center for Health Professions, University of Arkansas, Epley Center 275, 606 N Razorback Road, Fayetteville, AR 72701, USA

[2] Department of Audiology and Speech-Pathology, University of Tennessee Health Science Center, 553 South Stadium Hall, Knoxville, TN, USA

[3] Department of Speech and Hearing Sciences, Elson S. Floyd College of Medicine, Washington State University, Spokane, WA, USA

[4] Department of Hearing, Speech, and Language Sciences, Gallaudet University, 800 Florida Ave, Washington, DC, USA

and somatosensory consequences of speech, suggesting that internal simulation and overt execution rely on similar mechanisms (Hickok 2012; Tian and Poeppel 2012). Despite similarities, neuroimaging studies comparing imagined to overt speech production also show differences between Cp and Op (Murphy et al. 1997; Shuster and Lemieux 2005). In particular, overt execution of speech is associated with greater activation in the premotor cortex (PMC), primary motor cortex (M1), left insula, and left posterior superior temporal gyrus (pSTG), suggesting greater activation related to processing motoric execution and associated sensory feedback (Shuster and Lemieux 2005). Despite neuroimaging evidence supporting overlap in cortical networks and differences in the magnitude of activation in Op and Cp, to our knowledge no studies have investigated how connectivity between motor and sensory regions is dynamically altered on a millisecond time-scale.

Although it is yet unclear how sensorimotor connectivity is modulated by Cp, it has been suggested that Cp follows a sequential activation pattern analogous to that in Op (Tian and Poeppel 2012). From this perspective, mental imagery in general is a top-down process that adapts the sequence of events in sensorimotor integration for overt production to the task of internally formulating and experiencing sensory goals (Tian and Poeppel 2012). In such cases, sequential estimation requires that motor plans induce somatosensory and subsequent (Tian and Poeppel 2012) or parallel internal auditory estimates related to quasiperceptual experiences (Rauschecker and Scott 2009; Hickok 2012; Guenther and Hickok 2015). It is as yet unclear whether activity associated with Op relative to Cp is lateralized to the left hemisphere or more bilaterally organized. Some studies have reported left lateralization for Op relative to Cp (Hickok et al. 2009; Murakami et al. 2015) while others report more bilateral organization for both Cp and Op, depending in part upon the analysis approach and task employed (Tremblay and Small 2011; Simmonds et al. 2014; Cogan et al. 2014). In addition to Op and Cp execution, one version of a dual-stream model of language (Hierarchical state feedback control model; HSFC) also proposes forward models prior to execution may play an important role in speech planning (Hickok 2012; Guenther and Hickok 2015). A strong prediction of that account is that sensory-phonological regions encoding the target and motor-phonological regions related to planning movements are activated in parallel (Hickok 2012). From this point of view, both Cp and Op would be expected to involve a forward model just following the selection of a speech target with modification when overt sensory feedback is anticipated (i.e., Op compared to Cp).

One potential way to test predictions from internal models is to examine millisecond temporal resolution 'neural oscillations' after the selection of speech syllable targets (i.e., speech planning) and in subsequent execution. Although relatively few studies have investigated neural oscillations in both Op and Cp (Jenson et al. 2014; Jenson et al. 2015), emerging principles derived from a dual-stream model and state feedback motor control models have been proposed (Hickok and Poeppel 2007; Giraud and Poeppel 2012; Hickok 2012). Collectively, those accounts suggest that low-frequency phase coherence and increases in power (< 10 Hz) are consistent with theoretical models of speech production proposing that speech planning and production are temporally organized along coherent oscillations converging at time constants consistent with the syllable unit (~ 200 ms) (MacNeilage 1998; Ghitza 2012; Giraud and Poeppel 2012; Hickok 2012; Gross et al. 2013; Doelling et al. 2014; Assaneo and Poeppel 2018). Therefore, one prediction from those models is that both Op and Cp should involve increased low frequency power and phase coherence between motor and sensory regions, with differences in the magnitude of coherence during overt execution (i.e., overt sensory feedback) compared to covert execution. However, to our knowledge, no studies have measured both power and phase coherence in sensorimotor and auditory sensory regions on a millisecond time-scale coinciding with speech planning and execution time-periods, respectively.

Neural oscillations in speech planning

While it is still unclear how neural oscillations are related to Op and Cp speech planning, a small body of work has investigated event-related time–frequency changes following auditory or orthographic cues but prior to imagined or overt execution (i.e., speech planning). In two complementary electrocorticographic (EcoG) studies of overt and covert planning and subsequent production, it was demonstrated that power in the gamma (40–100 Hz), alpha-mu (8–12 Hz), and beta bands (15–20 Hz) were similar across premotor, motor, somatosensory, and auditory regions, with enhanced responses in sensory (e.g., pSTG/Wernike's area) and motor regions (e.g., M1) during Op relative to Cp (Pei et al. 2011a, b). Additionally, alpha-mu suppression (i.e., power decrease) (8–13 Hz) distinguished vowels in Cp and Op conditions, suggesting that alpha-mu suppression may play a role in vowel encoding in both speaking conditions (Pei et al. 2011a). In addition to EcoG findings, a recent magnetoencephalographic (MEG) study acquired data during the 'set-up' of the speech oscillatory networks following a visual cue that indicated whether upcoming speech would be produced overtly or covertly (Gehrig et al. 2012). That study showed bilateral power enhancement maximal at low frequencies in frontal regions (e.g., superior frontal gyrus) for both conditions relative to a baseline, suggesting enhancements in the delta–theta bands are generally related to the set-up of the speech network. Further, left-lateralized beta band suppression in the medial and more lateral premotor cortex

and pSTG was greater for overt compared to covert speech planning, while alpha suppression was greater in the medial premotor cortex and middle temporal gyrus only (Gehrig et al. 2012). Taken together, those findings implicate alpha/beta suppression and low frequency enhancement in speech planning.

While a number of studies have focused on oscillatory power, fewer studies have investigated correlations between the phase of oscillations in frontal and temporal lobe regions during planning in Op and Cp. In the Gerhig et al. (2012) study, time-independent causal modeling analysis of phase coherence indicated that connectivity in the alpha and beta bands was best explained by both bottom-up and top-down connections across premotor, sensorimotor, parietal, and temporal regions. A more recent study employed a speech motor adaptation paradigm in which FAF was used to alter auditory feedback to vowel production to examine gamma and theta band coherence of the electroencephlagraph (EEG) in early and later trials in both speech planning and execution (Sengupta and Nasir 2015). A decrease in theta and gamma coherence was found at frontotemporal electrodes for late compared to early training (i.e., early trials) while theta and gamma increased at central electrodes for the same training periods. Differences in frontotemporal coherence occurred in both the planning and execution time-periods, suggesting that feedforward maps were progressively established as participants adapted to perturbed sensory input (Sengupta and Nasir 2015).

Taken together, findings in speech planning suggest that power enhancements at low frequencies (< 10 Hz) and power suppression in the alpha/beta bands are related to both Op and Cp speech planning. In addition, the evidence also implicates phase coherence across frontal motor and temporal lobe sensory regions in those same frequency bands during speech planning. However, because studies have not compared Op and Cp, it is as yet unclear from the current literature whether power and coherence in frontotemporal regions are modified by the expectation of overt auditory and somatosensory consequences in anticipation of overt execution. Such investigations are important in light of internal model proposals, suggesting that forward models in the dorsal stream are implicated in speech planning as well as execution and predict differences depending on whether speech is internally or overtly realized (Hickok 2012; Tian and Poeppel 2012).

Neuronal oscillations in speech execution

While few studies have directly compared measures power and phase coherence focusing during Op and Cp execution (Pei et al. 2011b), a number of studies have investigated those same measures focusing on the execution of speech with and without altered auditory feedback (AAF). One event-related EEG study focusing on event-related power demonstrated low frequency delta (1–4 Hz) and theta (4–7 Hz) power increases over the frontal lobes just following pitch-shifted feedback, suggesting that power at low frequencies is related to the integration of sensory consequences (Behroozmand et al. 2015). Complementary findings from the study of event-related potentials (ERPs) suggested that adaptive vocal responses were negatively correlated with parietal ERPs and positively correlated with frontal–central ERPs, consistent with the notion that pitch-shifted feedback reprograms feedforward motor commands by integrating auditory feedback (Behroozmand and Sangtian 2018). In the most recent study, Franken et al. (2018) found left hemisphere increases power in a 4–10 Hz band and a 12–16 Hz band during small (25 cents) pitch changes, suggesting that power enhancements in those bands were related to sensorimotor integration during pitch shifts as opposed to cognitive changes related to adaptation/conscious awareness. Taken as a whole, the few studies that have focused on a time-sensitive measure of power suggest that enhancements in low frequency power and suppression in alpha/beta bands are related to the integration of sensory feedback during speech execution.

In addition to studies focusing on power, a few studies have also investigated the role of interregional oscillatory phase coherence in sensorimotor integration during speech execution. Wang et al. (2014) reported localized phase coherence between prefrontal regions (including the IFG) and the pSTG prior to vowel production at frequencies < 10 Hz that was related to subsequent sensory suppression (i.e., N1 suppression), suggesting that coherence between temporal-sensory and frontal-motor regions at low frequencies in speech planning is related to convergence between forward and inverse models during overt execution (Wang et al. 2014). Evidence from an EcoG study of coherence between the dorsal premotor cortex and IFG with the pSTG also implicate low frequency theta (4–7 Hz) coherence in audiomotor integration following pitch-shifted feedback along with gamma-band coherence (Kingyon et al. 2015). Finally, while Sengupta and Nasir (2015) demonstrated robust phase coherence at theta frequencies during planning, they also found frontotemporal coherence during execution with auditory perturbation of vowels, further suggesting that low frequency coherence in planning is related to the integration of auditory feedback in execution. As such, evidence from studies of power and interregional phase coherence implicate low frequencies between frontal motor and temporal sensory regions in audiomotor integration during execution.

Independent component analysis of EEG in speech production

Consistent with previous electromagnetic studies, investigations of Op and Cp production using a blind source separation (BSS) approach known as independent component analysis (ICA) of EEG data have also demonstrated low frequency increases in power and alpha/beta suppression. The advantages of using a BSS approach, such as ICA, is that it does not require the selection of electrodes, ROIs, or averages over preselected time–frequency bins. In addtion, it is a robust means of removing movement-related artifact and accounts for the well-known volume conduction problem in the interpretation of EEG data (Makeig et al. 2004; Delorme et al. 2012). Studies conducted using this methodology have shown bilateral alpha and beta suppression in sensorimotor mu (μ) and posterior temporal lobe alpha (α) rhythms in visually cued Op and Cp consonant vowel syllable production tasks (Jenson et al. 2014, 2015). Jenson et al. (2014) demonstrated bilateral sensorimotor mu suppression during both Cp and Op, with significantly greater suppression occurring during speech execution in the Op condition. A second study examining alpha rhythms with source estimates in the posterior temporal lobe demonstrated alpha and beta suppression prior to overt execution, followed by diminished suppression in the alpha band and enhancement in theta band coinciding with overt execution (Jenson et al. 2015). During the same time-period, alpha and beta suppression peaked in the left and right sensorimotor components, suggesting that early temporal lobe alpha was associated with an internal model of speech production that was subsequently suppressed during overt execution along low frequencies. While findings in that study were interpreted as consistent with sensory suppression in the posterior temporal lobe related to early sensorimotor forward models (Houde and Nagarajan 2011), the evidence was circumstantial as it did not provide a measure of phase coherence between sensorimotor and posterior temporal lobe components.

While previous studies of Op and Cp demonstrated activation in sensorimotor mu and posterior temporal lobe components implicating interregional communication for speech production, those studies did not explore measures of phase coherence between components on a millisecond timescale. Further, the aim of those studies was to examine time–frequency activations in the time period following the cue to execute speech (i.e., the execution time period) as opposed to the speech planning period just following the orthographic cue. Given predictions from internal model proposals and previous electromagnetic evidence examining speech planning, increases in processing would be expected for Op relative to Cp in that time period in addition to the time period during execution. As such, one aim of the current study was to extend time–frequency analysis of power and phase to the speech planning time period. A second aim of the current study was to examine interregional phase coherence in Op and Cp in both planning and execution similar to previous studies investigating Op (Wang et al. 2014; Kingyon et al. 2015).

Differences in neural oscillations between Op and Cp in speech planning and execution would be expected because overt sensory feedback must be integrated with forward models in the Op condition while the Cp condition requires internal realization of syllable targets. Thus, it was hypothesized that oscillatory dynamics would show low power increases and alpha/beta suppression in sensorimotor mu and temporal lobe alpha components during planning for both Op and Cp. However, if forward models during planning are important for subsequent sensorimotor integration during execution, greater low frequency enhancement and alpha/beta suppression would be expected both during overt planning due to the anticipation and subsequent presence of overt sensory feedback in the Op condition. With regard to phase coherence between sensorimotor and temporal components, enhancements in low frequency (< 10 Hz) coherence were expected in the overt relative to covert execution time period during which overt sensory feedback must be compared with the forward model initiated during speech planning. In addition, earlier low-frequency phase coherence during covert execution was also expected due to the lack of a need to integrate delayed, overt sensory feedback during execution. In other words, because overt sensory feedback lags behind internal estimates, low frequency coherence was expected to occur earlier in the covert relative to overt condition.

Methods

Participants

Twenty-three right-handed native English speakers were recruited from the audiology and speech pathology program at the University of Tennessee Health Science Center. Participants (3 male) had a mean age of 25.16 years (range 21–46) and no history of hearing, cognitive, communicative, or attentional disorders. The Edinburgh Handedness Inventory was administered to establish handedness dominance for each subject (Oldfield 1971). The study was approved by the Institutional Review Board for the University of Tennessee Health Science Center and each participant provided written, informed consent prior to participation.

Stimuli

A timeline for the overt and covert production tasks is shown in Fig. 1. Syllable stimuli were combined to create syllable pairs such that half of the stimuli consisted of identical pairs

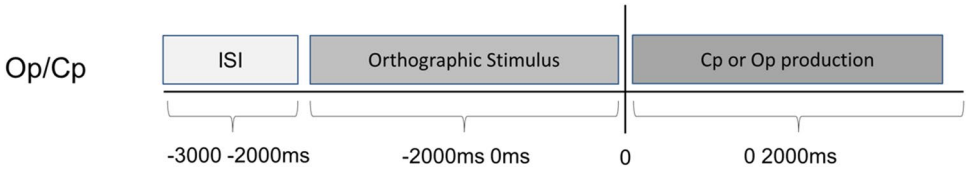

Fig. 1 Timeline of one trial in the Op and Cp conditions. A timeline of one trial from − 3000 ms prior to the orthographic cue to 2000 ms following the cue to produce speech in both the Op and Cp conditions

(e.g., /ba ba/) and half of the stimuli contained different pairs (e.g., /da ba/). Targets for the Cp and Op (i.e., speech production) conditions consisted of syllable pairings presented via graphemes (e.g., /ba da/). Visual stimuli for production were presented at the center of the visual field on Microsoft PowerPoint slides consisting of white text on a black background (Arial font, size 56) subtending a visual angle of 1.14 degrees.

Procedure

The experiment was conducted in an electrically and magnetically shielded, double-walled, sound-attenuated booth. Participants were seated in a comfortable chair with their heads and necks well supported. Stimuli were presented and button press responses were recorded by a PC computer running Compumedics NeuroScan Stim 2 software, version 4.3.3. Visual stimuli were presented on a monitor (69.5 × 39.0 cm) placed 132 cm in front of the participant's chair. Visual stimuli (syllable pairs) remained on the screen for 2000 ms, and participants were instructed to begin their production when the visual stimuli disappeared. Thus, stimulus offset (i.e., removal) was the response cue in the production conditions. The stimulus was held on the screen prior to production to reduce working memory demands in the subsequent production. In the Op speaking condition, participants were instructed to produce the syllable pairs in their normal speaking voice and at a natural rate following the cue to speak. All overt productions were complete within the 2000 ms window between the response cue and the end of the trial epoch. Each of the conditions were comprised of 2 blocks of 40 trials each. Both the order of conditions and block presentation were randomized using an algorithmic random number generator for each subject prior to the experiment.

Data acquisition

Sixty-eight electrodes were used to acquire EEG and sEMG data. EEG data were acquired using an unlinked, sintered NeuroScan Quik Cap with 64 of the channels placed on the scalp according to the extended international standard 10–10 system, along with two mastoid reference channels (M1) and (M2) (Chatrian et al. 1985). The other four channels were used for collecting various types of artifact. The electro-oculogram was recorded by means of two electrode pairs placed above and below the orbit of the left eye (VEOL, VEOU) and on the medial and lateral canthi of the left eye (HEOL, HEOR) to monitor vertical and horizontal eye movements, respectively. The electro-cardiogram was recorded from one channel placed over the carotid artery. The sEMG channel was recorded from two recessed electrodes placed above and below the lips to capture movement potentials arising from the OO complex. In accordance with previous studies, the superior electrode was placed to the right of the philtral ridges and the inferior electrode was placed just below the mentolabial sulcus in line with the superior electrode. Although the activity from single muscles cannot be measured from perioral recordings (Blair and Smith 1986; Stepp 2012), the sEMG signals from perioral structures are associated with the activity of muscles involved in oral opening/closing (Gracco 1988).

All electrophysiological data were recorded using Compumedics NeuroScan 4.3.3 software in tandem with the Synamps 2 system. During data acquisition, electrophysiological data were band pass filtered (0.15–100 Hz) and digitized with a 24-bit analog-to-digital converter at a sampling rate of 500 Hz. The orthographic visual stimuli for subsequent production were shown on the screen for 2000 ms prior to removal, which served as the cue to produce the stimulus. Data collection was time locked to the removal of the orthographic cue in both production conditions (Cp and Op). The impedances of all electrodes were measured at 30 Hz before, during, and following testing and were never greater than 5 kΩ. All subsequent data preprocessing was implemented in Matlab and EEGLABv12.4b, an open source Matlab signal processing toolbox designed for event-related EEG analysis (freely available at http://sccn.ucsd.edu/eeglab/).

Data processing

Data processing was accomplished using a series of steps reported in previous studies at the individual and group levels (Bowers et al. 2013, 2014, Jenson et al. 2014, 2015; Cuellar et al. 2016; Saltuklaroglu et al. 2017). In the individual level analysis, two 40 trial blocks for each condition

were appended, independent component analysis (ICA) was applied across condition for each participant, and estimates of dipole locations were computed for all components for each participant. In the group analysis, the STUDY module in EEGLAB v12 was used to perform time–frequency and spectral analysis on nonartifactual components (i.e., neural components). Subsequently, components were clustered across participants using K-mean principle component clustering. Dipole locations, event-related spectral perturbations (ERSPs) and intercomponent phase coherence (IPCs) were examined. The details of each step in both the individual and group level analyses are described in more detail in the sections below.

Independent component analysis

In the first step, unprocessed EEG/sEMG data files from the two 40 trial blocks for each condition were appended to create one dataset per condition per participant. To reduce the time required to process data files, the two 40 trial blocks were appended and then digitally downsampled to 256 Hz using a cubic spline interpolation method. Trial epochs of 5000 ms (ranging from −3000 to +2000 ms around time zero) then were extracted from the continuous EEG data resulting in 80 trials per condition (240 total trials). Following epoching, the data were then filtered from 1 to 30 Hz using an FIR (linear finite impulse response), zero-phase filter with a rolloff of 12 dB per octave. As ICA has been shown to be effective for separating non-neural signals that contribute highly repetitive artifact all epochs containing gross, non-repeating (i.e., one-time artifact) greater than 200 microvolts were removed (Makeig et al. 2004). A minimum of 50 useable trials per subject per participant across conditions was required for inclusion in the analysis. This single trial noise reduction step resulted in a mean of 65 trials contributed per block. In the second step, the unaveraged, preprocessed data were decomposed using ICA. ICA was performed using the binary version of the extended infomax algorithm (i.e., the binica algorithm). The informax algorithm spheres the data prior to ICA rotation and returns a complete ICA decomposition resulting in the same number of components as the number of original channels (Lee et al. 1999). ICA was implemented with an initial learning rate of 0.001 and the stopping weight set to 10-7. ICA decomposition yielded 68 ICs for each participant, corresponding to the number of recording electrodes. The resulting square weight matrix (68 × 68) was applied to each participant, yielding a single set of IC weights for each experimental condition expressing independence in the data. The process allows for a comparison of condition differences for the same set of component weights across experimental conditions. Thus, because the Op condition was associated with highly repetitive and identifiable sEMG artifact, it was possible to isolate a set of IC weights consistent with the sEMG signal in each condition.

IC source estimation

In the third step, equivalent current dipole models (ECD) were generated for each component using a standard template boundary element model (BEM) in the DIPFIT toolbox, an open source Matlab plugin available at http://www.sccn.ucsd.edu/eeglab/dipfit.html (Oostenveld and Oostendorp 2002). Prior to ECD estimation, an average common reference was computed. After computing the average common reference, Cartesian electrode coordinates according to the international 10-10 system were warped to the BEM. The EOG, ECG, and mastoid electrodes were omitted from dipole fitting due to lack of standard placement. Automated coarse and fine fitting to the BEM yielded a single dipole model for each of the 1564 ICs (23 × 68). Dipole localization using the DIPFIT toolbox involves a back projection of the signal to a potential source that could have generated the signal, followed by computing a forward model that accounts for the highest proportion of variance in the scalp topography. IC sources are proposed to represent Cm3 synchronized cortical patches adequately represented by a single dipole (Delorme et al. 2012). The percentage of residual variance (RV) represents the proportion of variance in the original scalp recorded signal not accounted for by a single dipole model. An RV of < 15% suggests a good fit between the dipole estimate and IC scalp topography (Makeig et al. 2004; Delorme et al. 2012).

Following dipole fitting, ICs were evaluated and classified using three criteria. First, an automated algorithm designed for rejection of IC artifactual components known as the multiple artifact rejection algorithm (MARA) was used to tag likely neural and non-neural electrophysiological components (Winkler et al. 2014). MARA uses a pre-trained linear classifier to detect non-neural component features including fit error, the absence of distinct alpha peaks typical of neural components, local skewness, scalp-topographic features, and current density norm. These features are used to estimate the probability that a given IC is artifactual in origin. Second, scalp maps and log spectra were also visually inspected for indicators of non-brain artifact including abnormal spectral slope and topographic distributions known to be associated with eye-movement and myogenic artifact. Artifacts identified as myogenic then were visually inspected for movement-related potentials consistent with the original sEMG channel. Third, ICs consistent with lip-related sEMG single trial and average sEMG potentials were computed using the erpimage function in EEGLAB and 1280 time points. The single trial and average ERPs were used to inspect trial-by-trial potential increases in both the identified sEMG ICs and original sEMG channel. The aim of the early classification

process was to identify potential sEMG sources identified by ICA.

Group level analysis

In the fourth and final step, ICs were clustered across participants and conditions using K-means principal component analysis (PCA) clustering for ICs fit with a dipole model with < 15% RV. Because the hypothesized regions are known to have distinct spectra, dipole estimates, and scalp topographies all three measures were precomputed for each component prior to PCA clustering and used in the clustering analysis (Bowers et al. 2014; Jenson et al. 2014, 2015). Log spectra were computed using a fast Fourier transform (FFT) from 1 to 30 Hz using a window length of 256 points. Scalp topographies were computed as 68 channel (*x, y*) map gradients. ECD models were precomputed in the manner described in the previous step. Components more than three standard deviations from any cluster mean were excluded from further analysis. The number of final dimensions for the PCA was set to 10. 20 possible clusters were considered based on the mean of the components with neural sources identified in a previous step. Following PCA clustering, a measure of dipole density was computed to characterize the variability of dipole location across components belonging to the cluster (Delorme et al. 2012).

For the sEMG components, a statistical toolbox known as Corrmap was used to correlate IC scalp topographies associated with increases in lip-related sEMG amplitude (Viola et al. 2009). Corrmap was designed for the quantification of relationships between similar artifact related IC scalp topographies. Corrmap uses an initial template dataset and the corresponding IC to identify similar scalp maps across IC decompositions applying an automatically determined threshold derived from an iterative process. In the first step, the inverse weights from a selected template are correlated with all ICs from all datasets entered. Here, as only one sEMG IC per participant was identified using MARA and erpimage, the number of contributing components per participant was restricted to one. In other words, because all datasets for the same subject had the same ICA decomposition, only one map from a dataset was required. The mean correlation of a cluster resulting from the initial step is computed using Fisher's *Z* transform accounting for non-normal distribution of correlation values. Following RMS normalization of the mean map, the average map is then used as a new template and the initial process described in the first step is repeated with the aim of determining the dependence of the average map on the initial template. A value close to 1 indicates that the average map is robust to the selection of the initial map template. A more detailed discussion and data-driven proof of concept for the Corrmap algorithm may be found in Viola et al. (2009). Inter-rater reliability between Corrmap IC identification and a human rater (first author) using single-trial sEMG potentials to identify the component was computed using Krippendorff's alpha (Krippendorff 2013). For each sEMG IC identified by Corrmap, single trial and average ERP images computed in a previous step were used to confirm that event-related movement potentials were consistent with the original sEMG channel. Finally, as unprocessed sEMG signals have both positive and negative polarities, sEMG signals must be full-wave rectified prior to statistical comparisons of sEMG potentials across conditions. As such, sEMG signals were rectified prior to the computation of sEMG potentials and stored in a 1/1280 matrix for later averaging and statistical comparisons between conditions.

Following PCA, neural clusters were inspected and adjusted using the following steps. First, each cluster was identified using now well-known topographic distributions, spectra, and dipole location estimates used in previous studies (Bowers et al. 2013, 2014; Jenson et al. 2014, 2015). Lateral sensorimotor clusters (i.e., the μ rhythm clusters) were identified by characteristic spectral peaks at ~ 10 and 20 Hz, dipole estimates in a region from the precentral gyrus to the posterior central sulcus/ postcentral gyrus, and topographic maps centered over the sensorimotor cortex. Temporal lobe clusters were identified by characteristic spectra with a distinct peak at ~ 10 Hz, dipole estimates clustered in the superior temporal gyrus/superior temporal sulcus, and topographic maps centered over the posterior temporal lobe. Second, to adjust for ICA over-fitting known to occur with high-density arrays, in cases where one participant submitted two components in the same cluster (e.g., two left hemisphere μ components), the component with a dipole model accounting for the highest percentage of RV in the scalp-map was selected. The rationale for this approach is that scalp topographies contributing the least mutual information in ICA decompositions are more dipolar, and consequently, have lower residual variance in accounting for scalp topographies (Delorme et al. 2012). The inclusion of only one IC from each participant also addresses concerns that subsequent statistical tests may be weighted toward a single participant contributing a higher number of ICs in a given cluster. Increases and decreases in spectral power from a 1000-ms prestimulus baseline were computed for each component across conditions as reported in previous studies (Bowers et al. 2013; Jenson et al. 2014, 2015). ERSPs in the time period following the presentation of the orthographic stimulus (−2000 ms) during speech planning to the time period following overt production (1500 ms) were examined.

Intercomponent phase coherence

Although two ICs are maximally independent, they also may be associated with partial synchrony over event-related

time–frequency windows (Delorme et al. 2012; Makeig et al. 2004). In other words, while ICA minimizes mutual information shared between ICs and accounts for the problem of volume conduction (Delorme et al. 2012), ICs may nonetheless show oscillatory connectivity at limited time–frequency windows not attributable to volume conduction effects (Ofori, Coombes, and Vaillancourt 2015; Chung et al. 2017). For comparison with previous studies using interelectrode or interregional phase coherence (Kingyon et al. 2015; Sengupta and Nasir 2016), in the current study, a measure of phase coherence between sensorimotor mu and temporal lobe alpha rhythms was used. The newcrssf function in EEGLAB is designed to compute measures of event-related phase coherence (ERPCOH) between two component activities using methods similar to those for computing inter-trial phase coherence (ITPC). As described in Makeig et al. (2004), calculating ITPC requires computing spectral estimates in a 2-D Cartesian coordinate frame, wherein the vectors are returned as complex vectors in 2-D phase space. The norm of each vector is represented by the magnitude of each time–frequency spectral estimate. For two component signals a and b, ERPCOH is given by:

$$\mathrm{ERPCOH}^{a,b}(f,t) = \frac{1}{n}\sum_{k=1}^{n}\frac{F_k^a(f,t)F_k^b(f,t)^*}{\left|F_k^a(f,t)F_k^b(f,t)\right|},$$

where $F_k\,(f,t)$ is the spectral estimate of trial k at frequency f and time t. // is the complex norm of the time–frequency spectral estimates for two signals a and b. The magnitude of ERPCOH varies between 1 and 0, with 1 indicating perfect synchrony and 0 indicating a complete absence of synchrony at a given time–frequency window. The normalizing factor in the ERPCOH denominator ensures that only the relative phase of the two spectral estimates at each trial is taken into account.

In the current study, custom Matlab scripts using the newcrssf function were used to compute ERPCOH between left hemisphere sensorimotor mu and temporal lobe IC clusters. For each 5000 ms epoch (i.e., −3000 to −2000 ms), Morlet wavelets set from 1 cycle at the lowest frequency (1 Hz) rising linearly by 0.5 cycles up to 30 Hz with the padratio set to 2. Coherence values are then averaged across epochs to yield a single time–frequency matrix. Because the intrinsic interconnectivity between cortex and effector was of interest, ITPC due to time-locking events was removed. In other words, as any sensory stimulus may evoke an increase in phase coherence at the same time across trials, the effect of stimulus 'phase reset' was subtracted so that only intrinsic intercomponent coherence relative to time-locking events remained. After ERPCOH was computed in the same manner across conditions and participants, each 117×200 time–frequency matrix was concatenated and stored in a single three-dimensional matrix (frequency × time × participant) in each experimental condition for further statistical analysis and averaging. To aid clarity, in the sections following, ERPCOH will be referred to simply as intercomponent phase coherence (IPC).

Statistical analysis

Analysis of condition effects was carried out using the stat_cond function in EEGLAB. Stat_cond allows for testing condition effects using parametric or non-parametric surrogate hypothesis testing methods. Because time–frequency measures are known to be non-gaussian, condition effects were initially tested using non-parametric, permutation statistics. In the current study, 2000 random permutations were computed and compared to t values for condition differences. Values were corrected for multiple comparisons across the time–frequency matrix using the false discovery rate (FDR) correction at $p < 0.05$ (Hochberg and Benjamini 1995).

Results

sEMG cluster

Individual scalp topographies for each participant and the mean scalp topography for the sEMG cluster are shown in Fig. 2a with mean sEMG potentials for both the cluster and the original perilabial channel in the Op and Cp conditions shown in Fig. 4b. In addition to neural component clusters, Corrmap identified one set of inverse component weights for each participant using a correlation threshold of $r = 0.95$. The mean correlation coefficient was 0.99 with no r value for an individual scalp-map falling below 0.98. Following the Corrmap analysis, each IC was visually inspected for agreement with the single-participant ERP analysis described in a previous step. Of the 23 participants (i.e., ICs) included in the analysis, there was only one disagreement due to a less stereotyped scalp topography. Analysis of inter-rater reliability between a human rater (first author) using visual inspection of single trial movement-related potentials and Corrmap was $\alpha = 0.89$ as assessed by Krippendorff's alpha. For further analysis, the IC consistent with the original sEMG component selected by the initial human rater was included in place of the IC selected by Corrmap. The component was the only component for that participant exhibiting consistent increases in trial-by-trial sEMG potentials associated with speech production onset.

sEMG potentials

Following clustering, sEMG potentials were computed using the original sEMG channel and sEMG components

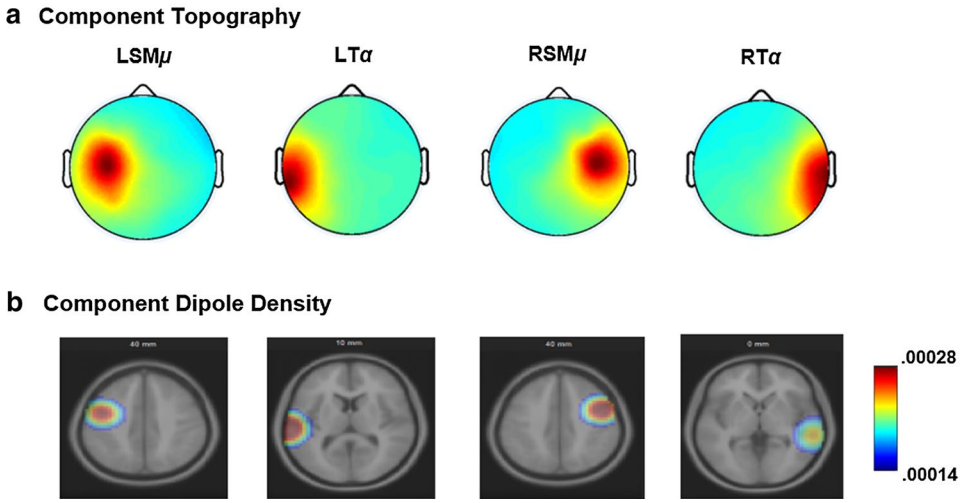

Fig. 2 sEMG cluster topographies and sEMG potentials in the Op and Cp conditions. Row (**a**) depicts the mean scalp topography and correlation coefficient for the sEMG cluster as identified by Corrmap (left) along with individual component topographies included in the cluster (right). Row (**b**) depicts movement-related potentials from the original sEMG channel (top) and sEMG component cluster identified using Corrmap as a function of Cp and Op conditions

across participants (Fig. 2b). In the channel analysis, robust increases in sEMG potentials relative to the time period prior to the cue to speak were found in the Op condition consistent with movement beginning at ~400 ms following the cue to speak (0 time). The component potentials also showed little change relative to the baseline in the Cp condition with a large increase in potentials at ~400 ms following the cue to speak in the Op condition. A comparison between conditions using permutation based statistics as described in the analysis section was performed for both the rectified channel and component sEMG signals. The paired *t* test showed statistically significant differences between conditions at $p\text{FDR} < 0.05$ (1×1280) just following the cue to speak in the Op condition. Post hoc *t* tests for the channel sEMG showed that only the speech production condition was associated with significantly higher sEMG potentials. The component paired *t* test showed significantly higher potentials occurring just following the cue to speak in the Op condition consistent with the onset of overt syllable production. Thus, both the sEMG channel and component signals were associated with increases during movement in Op and minimal movement in Cp.

Neural source IC clusters

Neural source (NS) clusters scalp topographies are shown in Fig. 3a with associated dipole densities in Fig. 3b. In accordance with our hypotheses, three component clusters were associated with the anterior bilateral network involved in the execution of speech. 17 participants submitted ICs to a cluster with mean a scalp distribution over the left sensorimotor cortex (LSMμ), a mean dipole location within the precentral gyrus [MNI (*x, y, z*) −44, 1, 39; RV 6.46%], and characteristic mu rhythm spectra with peaks at ~10 and 20 Hz reported in previous studies (Hari 2006; Jenson et al. 2014, 2015). 16 participants submitted ICs to a right sensorimotor (RSMμ) cluster with a mean scalp topography over the sensorimotor cortex, a mean dipole location in the right posterior inferior frontal gyrus (IFG) [MNI (*x, y, z*) 47, −7, 37; RV 6.18%], and mu spectra. Along with the motor clusters, two posterior temporal lobe clusters were identified. The left posterior temporal (LTα) and right posterior temporal (RTα) lobe topographic distributions and distinct spectral peaks at 10 Hz. 14 participants submitted ICs for the LTα cluster with scalp-topographic distributions centered over the left posterior temporal lobe and with a mean dipole location at [MNI (*x, y, z*) −61, −25, 3; RV 6.31%] in the pSTG. 14 participants submitted topographic distributions over the right posterior temporal lobe with a mean cluster dipole location at [MNI (*x, y, z*) 61, −26, −1; RV 5.59%] within the pSTG.

Event-related spectral perturbations (ERSPs)

ERSPs for the left sensorimotor and temporal alpha components are shown in Fig. 4a and right sensorimotor and temporal alpha components in 4b, respectively. As in previous studies and initial hypotheses, alpha and beta suppression across components relative to prestimulus baseline was found in both the Op and Cp conditions. Specifically, alpha and beta suppression occurred just following orthographic stimulus presentation and peaked in both the left and right sensorimotor components during overt production. By contrast, alpha and beta suppression peaked prior to overt production in the temporal lobe components both in the Op and Cp conditions. In addition to alpha/beta suppression, increases in power at frequencies < 10 Hz were found just following orthographic stimulus onset across conditions and during production in the Op condition. For the left (LSMμ) and right sensorimotor clusters (RSMμ) paired *t* test with an $p\text{FDR} < 0.05$ correction for multiple comparisons showed significantly

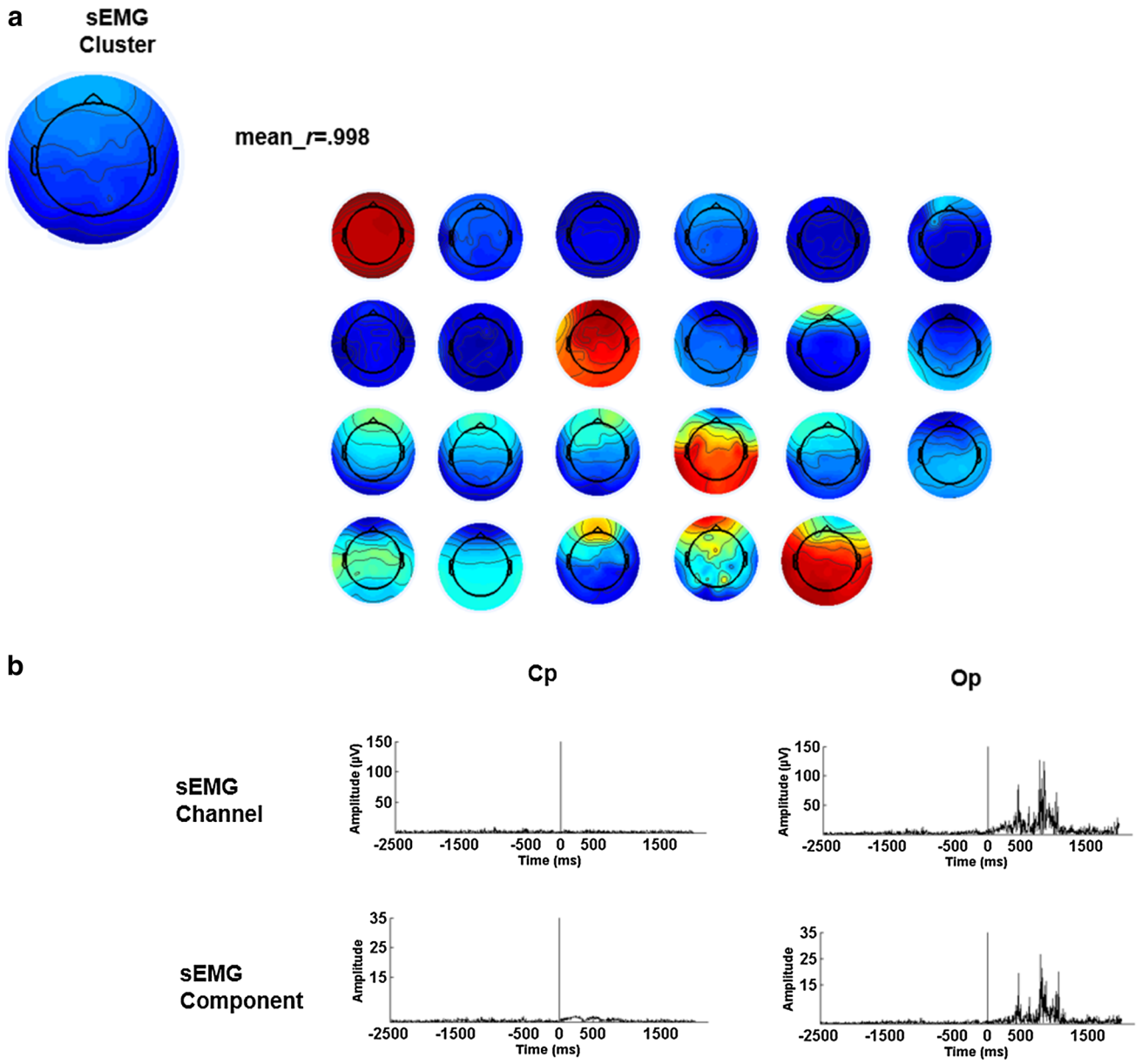

Fig. 3 Mean topographies and dipole density left and right lateralized component clusters. Row (**a**) depicts mean scalp topographies for the left and right sensorimotor mu (μ) and temporal lobe alpha (α) component clusters. Row (**b**) depicts a measure of dipole density for left and right sensorimotor mu (μ) and temporal lobe alpha (α) components

higher beta suppression in the time period prior to speech production in the Op condition with significantly higher beta and alpha suppression during production. In addition, enhancements in the theta range (4–7 Hz) following the orthographic stimulus along with delta–theta enhancements during production were found for both the left and right component clusters. A paired t test with an $p\text{FDR} < 0.05$ correction for the left and right component clusters, respectively, showed significantly higher beta suppression prior to the cue to speak (0 time) and just prior to movement termination (~ 1000 ms). Relative increases in the theta range during the speech planning period were also found along with significant delta–theta increases during speech production. Overall, the results showed that relative to covert production, the overt production was associated with higher alpha/beta suppression across components along with significant enhancements in the theta range during the planning period and delta–theta enhancements during execution.

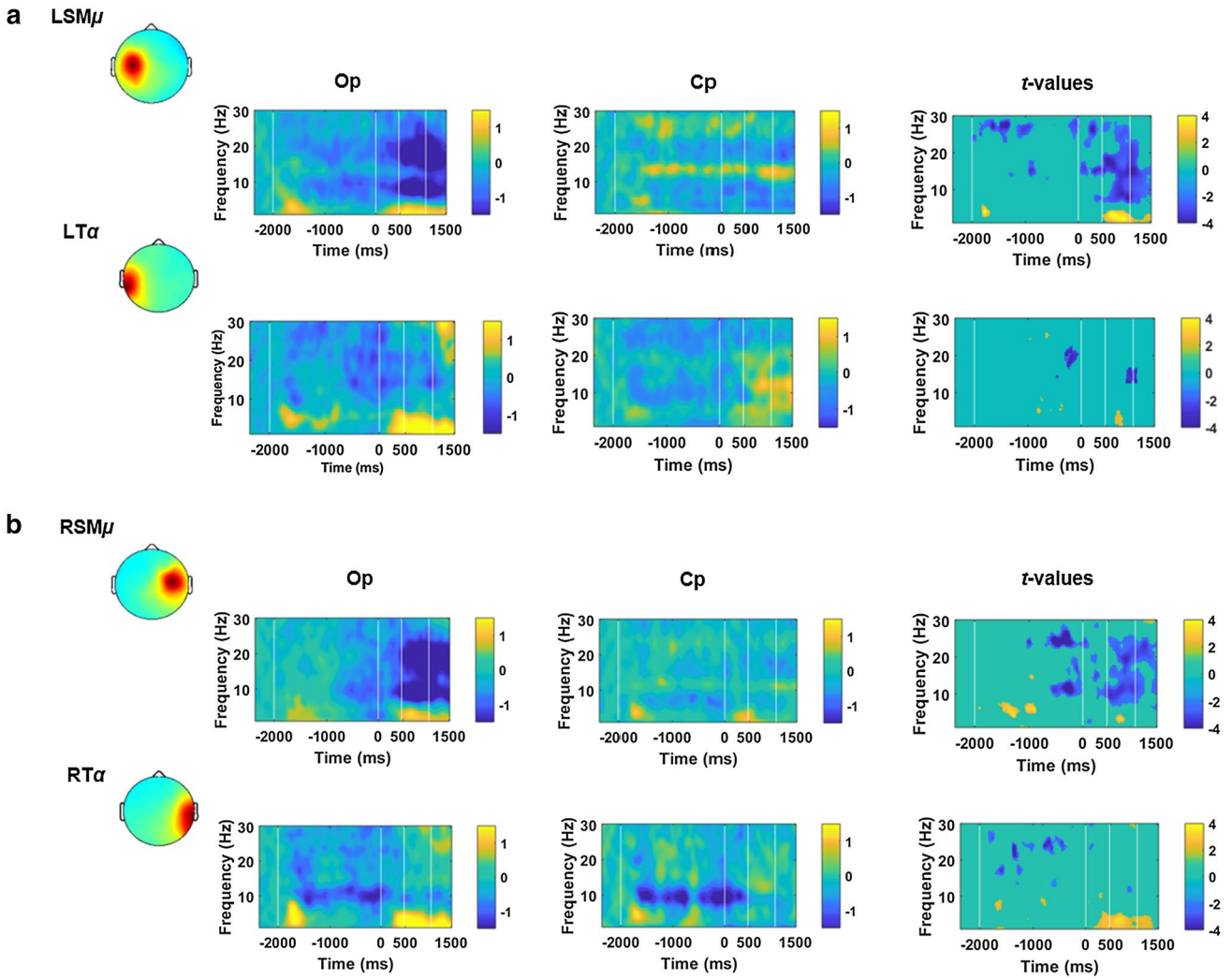

Fig. 4 ERSPs in the Op and Cp conditions across sensorimotor and temporal lobe components. Row (**a**) depicts ERSPs in the 1–30 Hz range for the left hemisphere sensorimotor mu (μ) (top) and temporal lobe alpha (α) component clusters (bottom) in the Op and Cp conditions followed by t values for condition differences at pFDR < 0.05. Row (**b**) depicts ERSPs in the 1–30Hz range for the right hemisphere sensorimotor mu (μ) (top) and temporal lobe alpha (α) components (bottom) followed by t values for condition differences at pFDR < 0.05. Blue t values represent significant decreases in power for a comparison of Op > Cp while yellow–white values represent increases in power for the comparison Op > Cp. White vertical lines from left to right represent orthographic stimulus onset (− 2000 ms), the cue to execute (0 ms), the mean onset of movement (~ 480 ms) and mean offset of movement (~ 1100 ms) in the Op condition

Intercomponent phase coherence (IPC)

Intrahemsipheric IPCs for the left and right sensorimotor and temporal lobe components are shown in Fig. 5a, b, respectively. Left hemisphere IPCs showed increased coherence at frequencies < 10 Hz during speech planning with a peak occurring during execution in the Op condition while a similar peak was evident in the Cp condition just following the cue to speak (0 time) with diminishing values thereafter. Intermittent peaks in the beta band were also evident in the Cp condition. Paired t *test*s with a pFDR < 0.05 correction showed significant increases in phase coherence in the Op condition at < 10 Hz during speech execution. Significant decreases in < 10 Hz coherence were found for the Op relative to Cp condition in the speech planning period just following the cue to speak (0–500 ms), along with intermittent increases in the beta band prior to and following speech execution. In the right hemisphere, phase coherence appeared to show diminished values during execution in the Op condition relative to the Cp condition. As in the left hemisphere, in the Cp condition increased coherence at frequencies < 10 Hz and intermittent beta coherence were evident throughout the trial. A paired t test with a pFDR < 0.05 correction showed significantly lower coherence in the Op relative to Cp condition in the < 10 Hz and beta ranges during speech planning and following the cue to speak. In

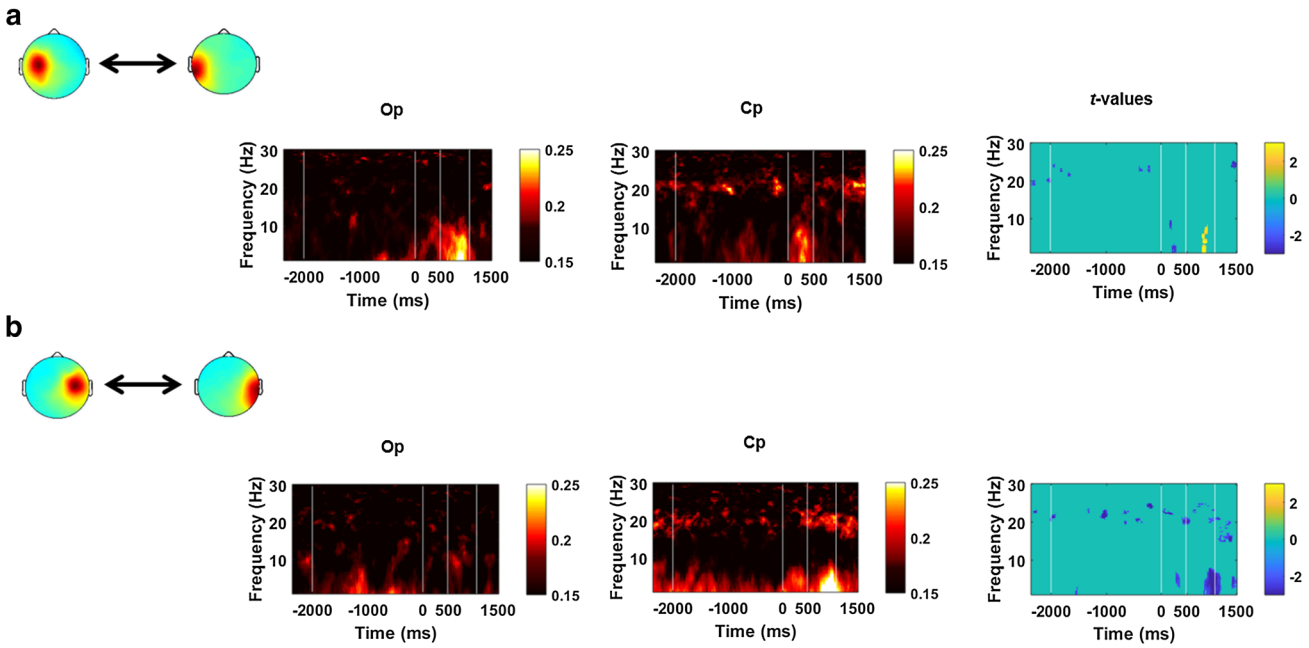

Fig. 5 Intercomponent coherence for the left and right sensorimotor and alpha components. Row (**a**) depicts interhemispheric phase coherence values (red–yellow) for the left hemisphere in the Op and Cp conditions followed by *t* values for condition differences at *p*FDR < 0.05. Row (**b**) depicts interhemispheric phase coherence values (red–yellow) for the left hemisphere in the Op and Cp conditions followed by *t* values for condition differences at *p*FDR < 0.05. Blue *t* values represent significant decreases in power for a comparison of Op > Cp while yellow–white values represent increases in power for the comparison Op > Cp. White vertical lines from left to right represent orthographic stimulus onset (− 2000 ms), the cue to execute (0 ms), the mean onset of movement (~ 480 ms) and mean offset of movement (~ 1100 ms) in the Op condition

summary, left hemisphere IPC at frequencies < 10 Hz was maximal during speech execution in the Op condition with a significantly earlier peak occurring in the Cp condition. By contrast, the right IPC was maximal at frequencies < 10 Hz in the Cp condition during the time period (500–1000 ms) associated with overt execution in the Op condition. In addition, significantly higher intermittent beta peaks were observed in both the left and right hemispheres in the Cp condition.

Discussion

The primary aim of the current study was to use ICA of EEG data to investigate differences in the time course of sensorimotor mu and temporal lobe alpha ERSPs and IPCs between overt relative and covert speech production with minimal influence from movement-related artifact. Consistent with previous studies using the same EEG data analysis approach and task, myographic artifact was readily identifiable and associated with stereotyped scalp topographies and movement-related potentials (Jenson et al. 2015). Using a broader time range to encompass speech planning, there were also several novel findings regarding ERSPs relative to our previous studies (Jenson et al. 2014, 2015). First, greater suppression in the beta band was observed across components in the speech planning period for the Op relative to Cp condition. Second, greater enhancements were observed in theta band during the speech planning period just following stimulus presentation and in the delta–theta range during overt speech execution relative to Cp. Further, in the temporal lobe alpha components, low frequency enhancement during execution also encompassed the lower alpha band (8–10 Hz) as compared to the sensorimotor components in which power enhancements during execution were more restricted to the delta–theta range (1–5 Hz).

Finally, analysis of IPCs also showed significant condition differences in the < 10 Hz and beta (15–25 Hz) frequency ranges. In the Op condition, left hemisphere low frequency coherence was greater in a short time window during speech execution, while a significantly earlier peak in the same frequency range was observed in the Cp condition. Right hemisphere contrasts showed enhanced low frequency coherence in the Cp relative to Op condition primarily in the time period following the cue to speak. One unexpected finding was that IPC was higher in the beta band across left and right components in the Cp relative to Op condition throughout the trial, suggesting task-related enhancements of the beta band. Overall, Op and Cp were associated with similar time–frequency distributions in both phase and power

with some task-related modulations. Taken together, the current findings implicate increases in both low frequency power enhancement and phase coherence along with higher frequency beta/alpha suppression in speech sensorimotor integration for Op relative to Cp conditions. Those findings are consistent with the prediction that forward models are activated in parallel and maintained during execution with significant differences in power and the timing of interregional coherence to accommodate overt sensory feedback. In the sections following, findings are discussed in the context of the relatively small body of work investigating neural oscillations in speech planning and execution, respectively, and predictions derived from dynamic internal models of motor control.

Power and phase coherence in speech planning

Speech planning in the current study was defined as the time between the orthographic cue and a cue to execute covert or overt speech. Internal models of speech motor preparation have proposed that early forward models of upcoming sensory targets are initiated in sensory and motor regions prior to speech execution, though few high time-resolution studies have investigated cued speech planning in both Op and Cp (Hickock 2012; Tian and Poeppel 2012). Studies that have investigated neural oscillations following an auditory or orthographic cue have implicated suppression in the alpha and beta bands prior to both overt and covert production in cortical premotor, sensorimotor, auditory, and audiomotor regions (Pei et al. 2011a, b; Gehrig et al. 2012; Cogan et al. 2014). In particular, Gehrig et al. (2012) focused on analysis of MEG oscillations in a set-up period following a cue, indicating that subsequent stimuli would be produced either overtly or covertly. That study reported increases in beta band suppression over the left lateral sensorimotor cortex and pSTG regions (planum temporale/supramarginal gyrus) prior to overt relative to covert production, whereas greater alpha suppression was localized to the more inferior temporal lobe (middle temporal gyrus), medial premotor, and parietal regions. Further, Gehrig et al. (2012) reported significantly larger delta–theta band power across frontal regions in speech set up that was not significantly greater for overt relative to covert conditions, suggesting similar prefrontal processing in both conditions.

Findings in the current study of greater beta band suppression during the speech planning period are consistent with greater suppression in sensorimotor/premotor and pSTG regions as reported in Gehrig et al. (2012). However, findings in the current study also suggest that differences in beta band suppression during speech planning may also be extended to the right hemisphere sensorimotor and pSTG regions, consistent with more recent accounts and findings supporting sensorimotor integration for speech as a bilateral process (Cogan et al. 2014; Pei et al. 2011b; Tremblay and Small 2011; Simmonds et al. 2014). Further, in contrast with the Gehrig et al. (2012) study, significantly larger theta band (4–7 Hz) power enhancements were found across components immediately following stimulus presentation in the Op relative to Cp condition. As such, findings in the current study implicate theta band increases in speech planning prior to the cue to execute speech over sensorimotor and posterior temporal regions in addition to enhanced beta suppression.

One likely reason for more bilateral beta suppression and theta enhancement in the current study as compared to Gehrig et al. (2012) is a difference in the paradigm used to investigate speech planning as opposed to speech 'set up.' In the current study, the syllables were known to participants prior to the cue to execute either covert or overt speech and as such participants could begin planning their response prior to the cue to execute. By contrast, in the Gehrig et al. (2012) study, participants were given a cue indicating that the upcoming sentence level stimuli would subsequently be produced either covertly or overtly but were not able to plan to produce speech targets. Those findings imply that general cognitive mechanisms are related to preparation to produce both overt and covert speech in the absence of known speech targets. However, under conditions in which the upcoming syllable targets are known prior to production (i.e., speech planning), power enhancement in the theta band and suppression in the beta band may be related to preparation for subsequent overt or covert execution. Such a distinction may be important because delta–theta frequencies are implicated in sensorimotor integration under perturbed sensory feedback (Behroozmand et al. 2015), show interregional or interelectrode phase coherence related to sensory suppression (Wang et al. 2014; Kingyon et al. 2015; Sengupta and Nasir 2015; Behroozmand and Sangtian 2018), and are implicated more generally in syllable parsing in both perception and production (Doelling et al. 2014; Asseneo and Poppel 2018). As such, delta–theta enhancement would be expected to be important both for speech planning as part of a forward model as well as during execution when sensory feedback is integrated with the initial forward estimates (Hickok 2012).

Similar to differences in power between the current study and Gehrig et al. (2012) study, task differences may account for discrepancies in phase coherence. In the current study, increased beta band IPCs in both the left and right hemispheres were observed for Cp relative to Op in the planning period, while increases in low frequency coherence were associated with processing just prior and during overt execution in the Op condition. Increased beta band coherence in the Cp conditions was somewhat surprising given previous reports of increased beta band phase coherence in speech set-up for Op relative to Cp (Gehrig et al. 2012). As in the current study and in contrast to Gehrig et al. (2012) study,

Wang et al. (2014) reported only low frequency (< 10 Hz) coherence during the speech planning period that was subsequently related to overt sensory feedback (i.e., N1 suppression), suggesting that interregional coherence may be altered for speech planning as compared to speech 'set-up.' Such an interpretation is reasonable as low frequencies are thought to be related to the integration of forward models and overt sensory feedback. In such cases, interregional communication would be expected to shift toward frequencies required for sensorimotor integration. As such, it appears that when speech targets are known prior to execution, coherence in the planning period is enhanced at low frequencies < 10 Hz for Op compared to Cp, while phase at both low and upper beta frequencies is enhanced for covert speech.

Power and phase coherence in speech execution

Findings regarding power during execution highlight substantial similarities along with some differences between the Op and Cp conditions. In both conditions, alpha and beta power were maximal prior to execution in the temporal lobe components and during execution in the sensorimotor components. In addition, low frequency power enhancements reaching into the lower alpha band (1–10 Hz) were observed in both conditions for temporal lobe components. The current data are consistent with inhibition of temporal lobe components at low frequencies proposed to be critical for sensory analysis and suppression. At the same time, 1–4 Hz sensorimotor and < 10 Hz temporal lobe alpha power increases during execution were also significantly higher in the Op compared to Cp conditions, suggesting that low frequency power increases are related to the presence of sensory feedback. Because fMRI studies have shown an inverse relationship between alpha/beta suppression and blood oxygen level dependent (BOLD) measures, those findings are broadly consistent with previous neuroimaging findings showing greater BOLD activation in sensorimotor and posterior temporal lobe regions during Op compared to Cp (Shuster and Lemieux 2005; Tremblay and Small 2011; Simmonds et al. 2014). Findings of low frequency power increases are also in accord with EEG/EcOG evidence demonstrating delta–theta enhancement during sensory perturbation (Kingyon et al. 2015). As such, findings in the current study are consistent with a role for low frequency power increases in both planning and execution, albeit with significantly larger responses for Op relative to Cp conditions.

While few studies have investigated phase coherence in speech production, those that have suggested that frontotemporal coherence at < 10 Hz is associated with sensory feedback in overt production. Low frequency coherence is correlated with subsequent sensory suppression (Wang et al. 2014), the onset of perturbed auditory feedback (Kingyon et al. 2015), and decreased coherence between frontal and temporal electrodes occurs with practice compensating for sensory perturbation (Sengupta and Nasir 2015). The current finding of a left hemisphere low frequency peak coincident with movement onset and offset is a novel finding as previous studies have not measured coherence in relation to the onset of movement-related potentials. That finding suggests that sensorimotor and temporal lobe regions coordinate low frequency phase to process sensory input in the left hemisphere near movement onset and diminish following offset. By contrast, in the Cp conditions low-frequency coherence peaked earlier in the left hemisphere and was associated with significantly greater coherence in the right hemisphere, suggesting that intrahemispheric connectivity is dynamically altered in time along low frequencies to accommodate the internal realization of sensory targets as opposed to the integration of overt sensory feedback. In addition, as discussed in the speech planning section, pervasive increases in beta coherence were observed in Cp following the cue to execute, implicating as yet unclear task-related factors differentiating Cp from Op. To our knowledge, this is the first study to investigate differences in intrahemispheric phase coherence between Op and Cp conditions, and as such, more research is needed to determine what functional role alterations in phase coherence perform in Cp conditions during internal execution.

Dynamic oscillations in internal models of motor control

Because Cp is considered to be an adaptation of internal models of motor control for realizing internally generated speech, the task goal is an important consideration (Tian and Poeppel 2012). One way to interpret time-sensitive changes in power and interregional coherence is to discuss them in the context of larger scale cognitive-sensorimotor interactions necessary for achieving task goals (Seigel et al. 2012; Fries 2005). According to a 'dynamic' framework, the cognitive processes at work in any task (including internal models of motor control) are altered by task-specific requirements associated with different time–frequency distributions within the same cortical networks (i.e., multiple realizability). For example, tasks requiring a sensorimotor response may be associated with large-scale interregional beta coherence (Buchman and Miller 2007; Gross et al. 2004; Saalmann et al. 2007; Hipp et al. 2011) or a shift toward gamma (~ 30–60 Hz) coherence depending on task requirements (Siegel et al. 2008; Gregoriou et al. 2009). In addition, interregional phase coherence and local band power are often disassociated, suggesting that each process may index different underlying neural computations and associations with task requirements (Buchman and Miller 2007; Pesaran et al. 2008; Siegel et al. 2008; Hipp et al. 2011). As such, neural oscillations related to 'mechanistic' internal model

accounts of speech motor control may also be considered within broader dynamic-mechanistic models predicting task-specific alterations in neural oscillations (Siegel et al. 2012).

Broadly consistent with a dynamic model, findings in the current study show both similarities and differences between ERSP and IPC measures related to Op and Cp tasks. While ERSPs appear to index primarily differences in the magnitude of response between Op and Cp conditions, IPCs are more closely associated with differences in interregional timing and coordination related to overt sensory feedback (i.e., overt execution) or the lack of overt sensory feedback (i.e., covert execution). One finding common to analyses of both ERSP and IPC measures was an increase at < 10 Hz during overt execution relative to covert execution, further suggesting a role in processing sensory feedback both in local power and larger scale phase coherence between regions. Such coupling at low frequencies may be related to intrinsic oscillatory properties of the sensorimotor cortex, allowing for sensory processing at preferential speech rates consistent with the syllable unit (MacNeilage 1998; Ghitza 2012; Giraud and Poeppel 2012; Assaneo and Poeppel 2018). Further, the findings are broadly consistent with internal model proposals implicating a frontotemporal dorsal sensorimotor stream in both speech planning and execution (Hickok 2012) and with theories of speech production predicting preferential processing consistent with the syllable unit (McNeilage et al. 1998; Ghitza 2012; Giraud and Poeppel 2012; Assaneo and Poeppel 2018).

Although the cause of increased beta coherence in the Cp condition is unclear, one plausible explanation is that Cp is associated with increased communication between motor and sensory regions along a band closely related to both motor processing and task-related goals. According to a relatively recent state feedback (SFC) motor control model, Cp requires forward models relying on simulations of somatosensory and auditory targets during speech planning even without anticipation of subsequent input from overt feedback (Hickok et al. 2011, Hickok 2012). Because the beta band is closely related to both motor processing and attentional control (Siegel et al. 2012), we speculate that the results may reflect differences in internal cognitive states related to Cp not required in the Op condition. Such an explanation is consistent with pervasive beta increases even prior to the orthographic stimulus and throughout the trial. Alternatively, it is plausible that increased beta coherence in the Cp condition reflects cooperation between sensorimotor and temporal lobe components in the inhibition of overt execution during Cp. Similar explanations have been proposed to account for inhibition of motoric execution related to the beta band in movement more generally (Engel and Fries 2010). However, given that this is, to our knowledge, the only study to compute coherence between sensorimotor and posterior temporal lobe regions during Cp, more research is needed to determine what role sensorimotor and temporal lobe beta coherence may play in internal speech production.

Limitations and conclusions

One limitation of the current study is related to the analysis method used to process EEG data. Because ICA is an unsupervised and entirely data driven approach, coherence and power could only be computed for participants submitting ICs consistent with spectra, scalp topographies, and dipole sources in the sensorimotor mu and posterior temporal lobe. As such, interpretation of the results is limited to the participants submitting dipolar ICs consistent with inclusion criteria and future studies should investigate power and phase in a larger sample. A second limitation is that the measure of phase coherence used in this study does not indicate the direction of coherence. For that reason, the analysis does not suggest that sensorimotor regions caused increases in phase coherence with posterior temporal regions or vice versa. One possible avenue for future study using the ICA approach is to compute measures of directed phase coherence or directed transfer function that would indicate causal influences similar to previous studies exploring nonspeech motor control (Chung et al. 2017). Finally, the current study measured only low frequencies < 30 Hz in Cp and Op speech planning and production. Data and models have also demonstrated low (30–60 Hz) and high gamma (60–120 Hz) coherence/power increases in speech planning and execution (Pei et al. 2011a; Kingyon et al. 2015; Giraud and Poeppel 2012; Behroozmand et al. 2016; Chang et al. 2013; Kort et al. 2016). For that reason, future studies should also study gamma-band coherence and power in relation to speech planning and execution.

Findings in the current study demonstrate that the power and phase of sensorimotor mu and posterior temporal lobe alpha rhythms index sensorimotor processes involved in preparation for and execution of Op and Cp. Our findings suggest that beta rhythm suppression and theta rhythm enhancement are greater during planning for Op relative to Cp. During overt execution, < 10 Hz power and phase are enhanced when compared to Cp, consistent with a role for both power and phasic coordination of < 10 Hz rhythms in the integration of sensory feedback during movement. In future studies, the current signal processing approach could be used to further investigate the functional role of low-frequency rhythms in feedback integration using sensory perturbation paradigms (e.g., AAF paradigms) (Kittilstved et al. 2018). Interestingly, decreases in frontotemporal coherence have been reported in adults who stutter prior to dysfluent relative to fluent nonword repetition trials (Sengupta and Nasir 2016; Sengupta et al. 2017). As such, future studies may have implications for sensorimotor integration deficits

in children and adults who stutter (Bowers et al. 2018; Jenson et al. 2018) or people with acquired neurogenic language disorders related to sensorimotor integration (e.g., conduction aphasia) (Buchsbaum et al. 2011). Due to its relatively low cost and noninvasive nature, the EEG recording and signal processing approach used in the current study has high potential for translational use in both the evaluation and treatment of clinical populations.

References

Assaneo MF, Poeppel D (2018) The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm. Sci Adv 4:eaao3842. https://doi.org/10.1126/sciadv.aao3842

Behroozmand R, Sangtian S (2018) Neural bases of sensorimotor adaptation in the vocal motor system. Exp Brain Res 236:1881–1895

Behroozmand R, Ibrahim N, Korzyukov O et al (2015) Functional role of delta and theta band oscillations for auditory feedback processing during vocal pitch motor control. Front Neurosci 9:1–13. https://doi.org/10.3389/fnins.2015.00109

Behroozmand R, Oya H, Nourski K, Kawassaki H, Larson C, Brugge J, Howard M, Greenlee J (2016) Neural correlates of vocal production and motor control in human heschl's gyrus. J Neuroscience 36:2302–2315

Blair C, Smith A (1986) EMG recording in human lip muscles: can single muscles be isolated? J Speech Hear Res 29:256–266

Bowers A, Saltuklaroglu T, Harkrider A, 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

Bowers AL, Saltuklaroglu T, Harkrider A et al (2014) Dynamic modulation of shared sensory and motor cortical rhythms mediates speech and non-speech discrimination performance. Front Psychol 5:366. https://doi.org/10.3389/fpsyg.2014.00366

Bowers A, Bowers L, Hudock D, Ramsdell-Hudock R (2018) Phonological working memory in developmental stuttering: potential insights from the neurobiology of language and cognition. J Flu Dis 58:94–117

Buchsbaum B, Baldo J, Okada K, Berman K (2011) Conduction aphasia, sensory-motor integration, and phonological short-term memory–an aggregate analysis of lesion and fMRI data. Brain Lang 119:119–128

Buchman TJ, Miller EK (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science 315:1860–1861

Campos Viola F, Thorne J, Edmonds B et al (2009) Semi-automatic identification of independent components representing EEG artifact. Clin Neurophysiol 120:868–877. https://doi.org/10.1016/j.clinph.2009.01.015

Chang EF et al (2013) Human cortical sensorimotor network underlying feedback control of vocal pitch. Proc Natl Acad Sci USA 110:2653–2658

Chatrian GE, Lettich E, Nelson PL (1985) Ten percent electrode system for topographic studies of spontaneous and evoked EEG activity. Am J EEG Technol 25:83–92. https://doi.org/10.1080/00029238.1985.11080163

Chung JW, Ofori E, Misra G et al (2017) Beta-band activity and connectivity in sensorimotor and parietal cortex are important for accurate motor performance. Neuroimage 144:164–173. https://doi.org/10.1016/j.neuroimage.2016.10.008

Cogan G, Thesen T, Carlson C, Doyle W, Devinsky O, Pesaran B (2014) Sensory-motor transformations for speech occur bilaterally. Nature 507:94–97

Cuellar M, Harkrider AW, Jenson D et al (2016) Time–frequency analysis of the EEG mu rhythm as a measure of sensorimotor integration in the later stages of swallowing. Clin Neurophysiol 127:2625–2635. https://doi.org/10.1016/j.clinph.2016.04.027

Delorme A, Palmer J, Onton J et al (2012) Independent EEG sources are dipolar. PLoS One. https://doi.org/10.1371/journal.pone.0030135

Doelling K, Arnal L, Ghitza O, Poeppel D (2014) Speech comprehension by facilitating perceptual parsing. Neuroimage. https://doi.org/10.1016/j.neuroimage.2013.06.035.Acoustic

Engel A, Fries P (2010) Beta band oscillations: signaling the status quo. Curr Opin Neurobiol 50:156–165

Franken MK, Eisner F, Acheson DJ, McQueen JM, Hagoort P, Schoffelen JM (2018) Self-monitoring in the cerebral cortex: neural responses to small pitch shifts in auditory feedback during speech production. Neuroimage 179:326–336

Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn Sci 9:474–479

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

Ghitza O (2012) On the role of theta-driven syllabic parsing in decoding speech: intelligibility of speech with a manipulated modulation spectrum. Front Psychol 3:238. https://doi.org/10.3389/fpsyg.2012.00238

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

Gracco VL (1988) Timing factors in the coordination of speech movements. J Neurosci 8:4628–4639

Gregoriou GG, Gotts SJ, Zhou H, Desimone R (2009) High-Frequency, long-range coupling between prefrontal and visual cortex during attention. Science 324:1207–1210. https://doi.org/10.1126/science.1171402

Gross J, Schmitz F, Schnitzler I et al (2004) Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. Proc Nat Acad Sci 101:13050–13055

Gross J, Hoogenboom N, Thut G et al (2013) Speech rhythms and multiplexed oscillatory sensory coding in the human brain. PLoS Biol 11:e1001752. https://doi.org/10.1371/journal.pbio.1001752

Guenther F, Hickok G (2015) Role of the auditory system in speech production. Handb Clin Neurol 129:161–175

Hari R (2006) Action–perception connection and the cortical mu rhythm. Prog Brain Res 159:253–260

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

Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci 15:393–402

Hickok G, Buchsbaum B, Humphries H, Muftuler (2009) Auditory-motor interaction revealed by fmri: speech, music, and working memory in area Spt. J Cog Neuroci 15:673–682

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

Hipp JF, Engel AK, Siegel M (2011) Oscillatory synchronization in large-scale cortical networks predicts perception. Neuron 69:387–396. https://doi.org/10.1016/j.neuron.2010.12.027

Hochberg Y, Benjamini Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57:289–300. https://doi.org/10.2307/2346101

Houde JF, Nagarajan SS (2011) Speech production as state feedback control. Front Hum Neurosci 5:82. https://doi.org/10.3389/fnhum.2011.00082

Jenson D, Bowers AL, Harkrider AW et al (2014a) Temporal dynamics of sensorimotor integration in speech perception and production: independent component analysis of EEG data. Front Psychol 5:1–17. https://doi.org/10.3389/fpsyg.2014.00656

Jenson D, Harkrider AW, Thornton D et al (2015a) Auditory cortical deactivation during speech production and following speech perception: an EEG investigation of the temporal dynamics of the auditory alpha rhythm. Front Hum Neurosci. https://doi.org/10.3389/fnhum.2015.00534

Jenson D, Reilly KJ, Harkrider AW, Thornton D, Saltuklaroglu T (2018) Trait related sensorimotor deficits in people who stutter: an EEG investigation of μ rhythm dynamics during spontaneous fluency. Neuroimage Clin 19:690–702

Kittilstved T, Reilly KJ, Harkrider AW, Casenhiser D, Thornton D, Jenson DE, Hedinger T, Bowers AL, Saltuklaroglu T (2018) The effects of fluency enhancing conditions on sensorimotor control of speech in typically fluent speakers: an EEG mu rhythm study. Front Hum Neurosci 12:126

Kingyon J, Behroozmand R, Kelley R, Oya H, Kawasaki H, Narayanan N, Greenlee J (2015) High-gamma band fronto-temporal coherence as a measure of functional connectivity in speech motor control. Neuroscience 305:15–25

Kort NS, Cuesta P, Houde JF, Nagarajan SS (2016) Bihemispheric network dynamics coordinating vocal feedback control. Hum Brain Mapp 37(4):1474–1485

Krippendorff K (2013) Content analysis: an introduction to its methodology. Sage, Thousand Oaks

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

MacNeilage P (1998) The frame/content theory of evolution of speech production. Behav Brain Sci 21:499–546

Makeig S, Debener S, Onton J, Delorme A (2004) Mining event-related brain dynamics. Trends Cognit Sci 8:204–210

Murakami T, Kell C, Restle J, Ugawa Y, Ziemann U (2015) Left dorsal speech stream components and their contribution to phonological processing. J Neurosci 35:1411–1422

Murphy K, Corfield DR, Guz A et al (1997) Cerebral areas associated with motor control of speech in humans. J Appl Physiol 83:1438–1447

Oldfield R (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113

Oostenveld TF, Oostendorp (2002) Validating the boundary element method for forward and inverse EEG computations in the presence of a hole in the skull. Hum Brain Map 17:179–192

Pei X, Barbour DL, Leuthardt EC, Schalk G (2011a) Decoding vowels and consonants in spoken and imagined words using electrocorticographic signals in humans. J Neural Eng 8:46028. https://doi.org/10.1088/1741-2560/8/4/046028

Pei X, Leuthardt EC, Gaona CM et al (2011b) Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. Neuroimage 54:2960–2972. https://doi.org/10.1016/j.neuroimage.2010.10.029

Pesaran B, Nelson MJ, Andersen RA (2008) Free choice activates a decision circuit between frontal and parietal cortex. Nature 453:406–409. https://doi.org/10.1038/nature06849

Price CJ (2012) A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. Neuroimage 62:816–847

Rauschecker J, Scott S (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nature Neurosci 12:718

Saalmann Y, Pigarev I, Vidyasagar T (2007) Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. Science 316:1612–1614

Saltuklaroglu T, Harkrider A, Thornton D, Jenson D (2017) EEG Mu (μ) rhythm spectra and oscillatory activity differentiate stuttering from non-stuttering adults

Sengupta R, Nasir SM (2015) Redistribution of neural phase coherence reflects establishment of feedforward map in speech motor adaptation. J Neurophysiol 113:2471–2479. https://doi.org/10.1152/jn.00731.2014

Sengupta R, Nasir SM (2016) Anomaly in neural phase coherence accompanies reduced sensorimotor integration in adults who stutter. Neuropsychologia 93:242–250

Sengupta R, Shah S, Loucks T, Pelczarski K, Yaruss S, Gore K, Nasir S (2017) Cortical dynamics of disfluency in adults who stutter. Physiol Rep 5:e13194. https://doi.org/10.1481/phy2.13194

Shuster LI, Lemieux SK (2005) An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. Brain Lang 93:20–31. https://doi.org/10.1016/j.bandl.2004.07.007

Siegel M, Donner TH, Oostenveld R et al (2008) Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. Neuron 60:709–719. https://doi.org/10.1016/j.neuron.2008.09.010

Siegel M, Donner T, Engel AK (2012) Spectral fingerprints of large-scale neuronal interactions. Nat Rev Neuro 13:121–131

Simmonds A, Leech R, Colliins C, Redjep O, Wise RJ (2014) Sensorymotor integration during speech production localizes to both the left and right plana temporale. J Neuroci 34:12963–12972

Stepp C (2012) Surface electromyography for speech and swallowing systems: measurement, analysis, and interpretation. J Speech Lang Hear Res 55:1232–1247. https://doi.org/10.1044/1092-4388(2011/11-0214)a

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

Tremblay P, Small S (2011) On the context dependent nature of the contribution of the ventral premotor cortex to speech perception. Neuroimage 50:1561–1571

Viola F, Thorn J, Edmonds B et al (2009) Semi-automatic identification of independent components representing EEG artifact. Clin Neurophys 120:868–877

Wang J, Mathalon DH, Roach BJ et al (2014) Action planning and predictive coding when speaking. Neuroimage 91:91–98. https://doi.org/10.1016/j.neuroimage.2014.01.003

Winkler I, Brandl S, Horn F et al (2014) Robust artifactual independent component classification for BCI practitioners. J Neural Eng 11:35013. https://doi.org/10.1088/1741-2560/11/3/035013