Sophie Drew Eustace

NEURAL SIGNITURES OF SYLLABLE AND WORD TRANSITIONS IN SPEECH PRODUCTION

UNIVERSIDADE DO ALGARVE

Faculdade de Ciências Humanas e Sociais

2025

Sophie Drew Eustace

NEURAL SIGNITURES OF SYLLABLE AND WORD TRANSITIONS IN SPEECH PRODUCTION

Mestrado em Neurociências Cognitivas e Neuropsicologia

Trabalho efetuado sob orientação de:

Investigador Doutorado João Mendonça Correia

UNIVERSIDADE DO ALGARVE

Faculdade de Ciências Humanas e Sociais

2025

NEURAL SIGNITURES OF SYLLABLE AND WORD TRANSITIONS IN SPEECH PRODUCTION

Declaração da autoria de trabalho

Declaro ser autor deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da linguagem de referências incluída.

Assinatura:

_____

(Sophie Drew Eustace)

For my grandmother, who may not grasp all that I do,

but whose heart has always cheered me on and waits, with quiet joy,

for our next reunion.

Acknowledgements

“Alone we can do so little; together we can do so much.” - Helen Keller

This thesis marks the end of a demanding yet rewarding chapter, made possible by the generosity, patience, and encouragement of many.

To my family and friends: thank you for your unwavering support, for keeping me grounded when deadlines loomed. Your faith in me made the long nights lighter and the milestones sweeter.

My sincere gratitude to João Correia for accepting the supervision of this work and his rigour, availability, and steady guidance throughout. Your high standards, practical advice, and trust were invaluable from the first outline to the final draft.

I am grateful to the Cognitive Neuroscience group for welcoming this project, for the lively discussions that sharpened ideas, and for providing the space and resources that enabled data collection.

To my colleagues, within the programme and beyond, thank you for the shared notes, frank debates, coffee-fuelled-work sessions, and good humour when pressure peaked. Your companionship turned hard work into genuine progress.

To all who contributed, directly or behind the scenes, please accept my deepest thanks. This achievement is as much yours as it is mine.

Resumo em Português

A produção da fala assenta numa organização hierárquica, na qual unidades menores (p. ex., sílabas) se integram em unidades maiores (palavras, frases). Embora esta hierarquia esteja bem documentada na perceção, permanece pouco claro como as dinâmicas neurais diferenciam transições intra-palavra (sílaba-para-sílaba) de inter-palavra (palavra-para-palavra) durante a produção. Neste estudo, examinámos oscilações cerebrais medidas por EEG de alta densidade durante produção de fala rítmica, isolando, em tempo real, as transições em torno da terceira sílaba produzida (pk3). Vinte e um adultos falantes de português europeu participaram; dois foram excluídos por problemas técnicos, resultando em N = 19 (14 mulheres; 18-53 anos). Os participantes produziram pares de pseudopalavras com seis sílabas por ensaio, em dois arranjos equivalentes: 2+4 (palavra dissilábica seguida de palavra tetrassilábica) e 3+3 (duas palavras trissilábicas). O ritmo foi imposto por um metrónomo visual (600 ms), instruindo a emissão de uma sílaba por batida visual.

O desenho assegurou que pk3 marcava contextos distintos: uma transição entre palavras no arranjo 2+4 e dentro da palavra no 3+3. Gravámos áudio e EEG (128 canais; BioSemi ActiveTwo) e aplicámos um pipeline de pré-processamento standard (reamostragem, filtragem, rejeição de canais ruidosos, ICA com marcação automática de artefactos e interpolação). O áudio foi transformado para o envelope de amplitude e identificado o pico de cada sílaba (pk1-pk6) para alinhar EEG e produção. As análises centraram-se em épocas de −1 a +1 s em torno de pk3, minimizando contaminação de potenciais preparatórios de início de sequência e mantendo janelas simétricas para as transições pk2-pk3 e pk3-pk4.

No nível fonte, estimámos dipolos equivalentes e agrupámos componentes independentes em 20 clusters com base em mapas, espectros e localização (3–40 Hz). As ERSPs (event-related spectral perturbations) foram calculadas por componente e ensaio e depois agregadas por cluster e participante. As estatísticas entre condições recorreram a 2 000 permutações com controlo de múltiplas comparações por FDR de Benjamini–Hochberg (q = .05) nos bins tempo-frequência. A inferência confirmatória incidiu em ROIs corticais derivadas da literatura: circunvoluções frontais inferiores (IFG) bilaterais, circunvoluções temporais superiores (STG) bilaterais e áreas suplementares motoras (SMA). Reportamos, para ilustração, três clusters que exibiram efeitos corrigidos: Cluster 1 (IFG esquerdo), Cluster 3 (IFG direito) e Cluster 8 (STG esquerdo).

Os participantes acompanharam o metrónomo com elevada precisão. Ainda assim, as durações entre sílabas refletiram a estrutura lexical: o intervalo pk2–pk3 foi mais curto no 3+3 do que no 2+4 (diferença = −17 ms), $t(18) = 2.15$, p = .045; o intervalo pk3–pk4 foi mais longo

no 3+3 do que no 2+4 (diferença = +29 ms), $t(18) = -2.90$, $p = .010$. Estes desvios sugerem que o encadeamento intrapalavra é ligeiramente agilizado, enquanto a travessia de um limite lexical impõe custos adicionais de planificação temporal.

No EEG, o IFG direito (Cluster 3) exibiu uma modulação robusta pré-pk3 que se estendeu de ~ -500 a 0 ms, abrangendo ~4–30 Hz (teta/alpha até beta baixa), com maior modulação no 3+3 do que no 2+4 após correção por FDR. Este padrão é compatível com o ajustamento de estados preparatórios frontais consoante o papel hierárquico da sílaba iminente: manutenção do conjunto motor dentro de palavra (3+3) versus reconfiguração à beira de um limite lexical (2+4). No IFG esquerdo (Cluster 1) observou-se uma diferença estatística que incluí frequências em alfa (~10–14 Hz) próximo de -450 ms (i.e., antes da produção da terceira silaba), e no STG esquerdo (Cluster 8), observamos duas diferenças estatísticas significantes: antes da sílaba 3 (~9–11 Hz, -400 a -300 ms) e depois da silaba 3 (~26–30 Hz, 0–100 ms). Não emergiram efeitos corrigidos nas restantes 17 ROIs. Em conjunto, os nossos resultados apontam para uma assinatura frontal direita antecipatória sensível à hierarquia linguística, acompanhada de ajustes auditivos transitórios em torno do evento acústico.

Os resultados alinham-se com modelos onde oscilações de baixa frequência (delta/teta/alpha) suportam sequenciação e gating temporal, enquanto beta indexa conjunto preparatório e coordenação de estados motores. A preponderância do efeito no IFG direito é consistente com relatos de controlo inibitório e reset motor mediados por vias cortico-subtalâmicas (expressos na banda beta), sugerindo um mecanismo de terminação/configuração de chunks articulatórios ao aproximar-se um limite lexical. Importa notar que o principal efeito ocorre antes da articulação de pk3, reduzindo a plausibilidade de explicações baseadas em diferenças acústicas subsequentes ou artefactos mioelétricos faciais (minimizados também pela escolha de consoantes linguodentais/velares em pk3 e pelo pipeline ICA).

Metodologicamente, três aspetos reforçam a validade interna: (i) alinhamento em pk3, que equilibra contexto pré- e pós-transição e mitiga potenciais de preparação de início; (ii) clustering ao nível-fonte de componentes independentes, conferindo maior interpretabilidade anatómica face a análises por elétrodo; e (iii) controlo rigoroso de ruído e múltiplas comparações (ICLabel/ICFlag; FDR). Ao mesmo tempo, reconhecemos limitações: a amostra (N = 19) típica de estudos EEG pode limitar a deteção de efeitos mais subtis; a análise de 20 clusters, ainda que sumarizada com foco em ROIs a priori, não elimina totalmente preocupações de multiplicidade; pequenas assimetrias temporais entre condições (-17 ms vs. +29 ms) podem, em princípio, influenciar estimativas TF; e o uso de pseudopalavras sob

compasso rígido restringe a generalização para fala natural com variabilidade prosódica e semântica. Por fim, a resolução espacial do EEG de superfície é limitada para circuitos subcorticais (gânglios da base, cerebelo) implicados em temporização e manutenção de conjuntos.

Em termos de implicações, demonstramos que a estrutura hierárquica da produção modela tanto o comportamento temporal como as dinâmicas corticais pré-fala. A assinatura frontal direita anterior a pk3 sugere um mecanismo de controlo preditivo que distingue se a sílaba seguinte encerra uma palavra ou antecipa um novo início lexical. Tal padrão complementa evidência da perceção da fala, onde ritmos lentos ancoram a extração de unidades linguísticas e beta participa em previsões sensório-motoras. Como perspetivas futuras, propomos manipular a força do limite (frequência lexical, marcação prosódica), escalar a cadência (variação de ISI), e relacionar potência frontal trial-wise com tempos de transição, testando ainda acoplamento entre bandas (p. ex., teta-beta) e conetividade faseada entre IFG, STG e SMA. A combinação de EEG/MEG ao nível-fonte com perturbação do feedback auditivo poderá clarificar como previsões frontais interagem com monitorização sensorial em limites lexicais.

Em suma, quando os falantes produzem sílabas em cadência controlada, os tempos de transição e as oscilações corticais refletem a hierarquia linguística subjacente. Observa-se um padrão frontal direito robusto (teta/alpha-beta baixa) que antecede a transição-chave e diferencia contextos intra- e inter-palavra, enquanto os efeitos temporais auditivos são curtos e localizados. Estes resultados sustentam uma visão ritmo-centrada da produção, na qual redes frontais configuram estados preparatórios para a sequência articulatória e o sistema auditivo realiza ajustes transitórios em torno do evento acústico.

Palavras-chave: produção da fala; EEG; oscilações beta; giro frontal inferior; transições sílaba-palavra; ERSP.

Abstract

Speech production unfolds across different hierarchical levels, yet it remains unclear how neural dynamics differ between syllable- and word-level transitions. We recorded high-density EEG while 19 adults produced rhythmically cued pseudoword pairs comprising either two plus four syllables (2+4) or two trisyllabic words (3+3). Speech timing and EEG were time-locked to the third syllable (pk3), which corresponds to a word boundary in 2+4 and a within-word syllable transition in 3+3. Event-related spectral perturbations (ERSPs) were computed from an epoch (-1 to +1 s) around pk3 from EEG sources obtained from independent-component clusters. Behaviourally, intervals differed by condition: the pk2-pk3 interval was shorter in 3+3 condition relative to the 2+4 condition (difference = −17 ms), $t(18) = 2.15$, $p = .045$; the pk3–pk4 interval was longer in the 3+3 condition relatively to the 2+4 condition (difference = +29 ms), $t(18) = -2.90$, $p = .010$. At the cortical level, FDR-corrected ERSP differences emerged in three literature-derived ROIs: right inferior frontal gyrus (IFG; Cluster 3) showed a robust pre-pk3 modulation spanning ~4–30 Hz from approximately −500 to 0 ms, with greater power modulation for 3+3 than 2+4; left IFG (Cluster 1) exhibited a small, isolated alpha patch; and left superior temporal gyrus (Cluster 8) showed brief pre-pk3 alpha and early post-pk3 high-beta significant time-frequency clusters. No other ROI clusters yielded significant corrected effects. Our findings indicate a right-lateralised frontal preparatory signature that is sensitive to the hierarchical role of the upcoming syllable, alongside only transient auditory modulations. Together, these results support accounts in which low-frequency activity supports sequencing and beta rhythms index preparatory set during speech production, with timing and oscillatory dynamics jointly reflecting within- versus between-word transitions.

Keywords: speech production; electroencephalography; beta oscillations; inferior frontal gyrus; syllable–word transitions; event-related spectral perturbations

Acronyms and Abbreviations List

- BP - *Bereitschaftspotential*
- CMS - Common Mode Sense
- DRL - Driven Right Leg
- CNV - Contingent Negative Variation
- CV - Consonant-vowel syllable structure
- EEG - Electroencephalography
- EMG - Electromyography
- ERD - Event-related desynchronisation
- ERS - Event-related synchronisation
- ERSP(s) - Event-related spectral perturbation(s)
- FDR - False Discovery Rate
- IFG - Inferior frontal gyrus
- IPA - International Phonetic Alphabet
- ISI – Interstimulus interval
- ITI – Intertrial interval
- IC - Independent component
- ICA - Independent component analysis
- MEG - Magnetoencephalography
- MFG - Middle frontal gyrus
- ROI(s) - Region(s) of interest
- SD – Standard deviation
- SFG - Superior frontal gyrus
- SMA - Supplementary motor area
- SPL - Superior parietal lobule
- STG - Superior temporal gyrus
- STN – Subthalamic nucleus
- TF – Time-frquency
- TPJ - Temporo-parietal junction

Index

Figure Index

Table Index

Appendix Index

1. Introduction

Language is a hierarchically organised system in which smaller linguistic units are embedded within larger structures: syllables combine to form words, words combine into phrases, and phrases are assembled into sentences (Ding et al., 2016). This hierarchical structure is evident in both speech perception and production and is supported by temporally distinct patterns of neural activity (Blomert, 2011; Giraud & Poeppel, 2012). Speech production involves transforming abstract linguistic representations into precisely timed articulatory actions, engaging a distributed network of cortical and subcortical regions, including motor and premotor cortices for motor planning and execution, auditory cortex for sensory-feedback prediction and monitoring of auditory consequences, and basal ganglia and cerebellar structures for timing, sequencing, and selection of motor programmes (Christoffels et al., 2007; Herz et al., 2023; Hickok, 2012; Klostermann & Ehlen, 2020; Tourville & Guenther, 2011). Foundational neurocognitive models, such as Levelt's (1999) 'Blueprint of the Speaker', Hickok's (2012) dual-stream framework, and neuroimaging-based models of speech planning (Indefrey & Levelt, 2004), describe speech production as a multistage process encompassing conceptual preparation, lexical selection, phonological encoding, and motor execution. These accounts differ in scope and emphasis: Levelt's Blueprint outlines a stagewise architecture of information flow in speech production; Hickok's dual-stream model developed for speech perception and comprehension specifies a ventral pathway for mapping sound to meaning and a dorsal pathway that supports auditory-motor translation (Hickok & Poeppel, 2004, 2007). Hickok (2012) extends these circuits to production by emphasising internal forward models, efference-copy-based predictions of the sensory consequences of planned speech actions, within dorsal sensorimotor loops. Indefrey and Levelt (2004) provide a timing/localisation synthesis that estimates when and where these stages unfold. Together, these perspectives offer complementary constraints - process architecture (Levelt), perceptual pathways with production-relevant sensorimotor interfaces (Hickok & Poeppel, 2004, 2007; Hickok, 2012), and temporal windows (Indefrey & Levelt) - that motivate our time-frequency approach.

In speech perception, neural signals (e.g., intracranial/extracranial EEG or MEG) entrain (i.e., synchronise) with the acoustic amplitude changes of speech stimuli (Aiken & Picton, 2008; Luo & Poeppel, 2007). Interestingly, EEG oscillations relate to the hierarchical structure of speech input and are often associated with different linguistic timescales. For example, delta has been related to phrasal structure, theta to syllables, beta to word-level transitions, and gamma to phonemic detail, providing a temporal scaffold for both perception and production (Ding et al., 2016; Giraud & Poeppel, 2012; Zoefel, 2024). Delta-band activity (1-4 Hz) supports segmentation of continuous speech into larger meaningful units by tracking

the speech envelope (Aiken & Picton, 2008). Theta-band oscillations (4-8 Hz) are implicated in syllabic parsing and sublexical segmentation (Giraud & Poeppel, 2012; Luo & Poeppel, 2007). Alpha-band activity (8-13 Hz) has been linked to attentional modulation, auditory-motor coupling control, and temporal gating of linguistic information (Klostermann & Ehlen, 2020; Strauß et al., 2014), and may help regulate access to phonological and articulatory representations (Strauß et al., 2014; Klostermann & Ehlen, 2020; Zoefel, 2023; Piai et al., 2015; Giraud & Poeppel, 2012).

Furthermore, beta rhythms (15-30 Hz) have also been linked to motor planning and the temporal coordination of word-level transitions (Cao et al., 2024; Orpella et al., 2024; Piai et al., 2015). Beta activity over frontal cortices - often prominent over right inferior frontal and premotor regions - has also been associated with preparatory set/maintenance and sequential control during speech and orofacial actions (Pfurtscheller & Lopes da Silva, 1999; Piai et al., 2015; Weiss & Müller, 2012). In the cortico-basal ganglia-thalamo-cortical loop, converging evidence indicates a functional subdivision whereby low beta (~13-21 Hz) is more closely associated with the indirect basal ganglia pathway - via the striatum, external globus pallidus, subthalamic nucleus, and internal globus pallidus - whereas high beta (~21-35 Hz) reflects hyperdirect cortico-subthalamic drive, particularly coupling between SMA/premotor cortex and the STN (Cao et al., 2024; Oswal et al., 2021; Herz et al., 2023). Gamma activity (30-100 Hz) supports rapid phonemic encoding and articulatory transitions (Giraud & Poeppel, 2012). These rhythms are not only reactive but also predictive, enabling the brain to anticipate upcoming linguistic input based on contextual and rhythmic cues (Arnal & Giraud, 2012; Ding et al., 2016).

In EEG, time-frequency analysis methods such as event-related spectral perturbations (ERSPs) allow the tracking of transient power changes across frequency bands in response to experimental events, which may help characterise the neural temporal structure during speech-production paradigms (Makeig et al., 2004; Piai et al., 2015). Transitions between linguistic units, whether within a word (syllable-to-syllable) or between words (word-to-word), are hypothesised to recruit distinct oscillatory mechanisms (Cao et al., 2024; Orpella et al., 2024; Rapela, 2017), yet this has not been systematically tested during speech production. In speech-motor control, efference copy (a forward model) is an internal prediction of the sensory consequences of a planned motor command, enabling rapid comparison with incoming feedback for error detection and correction (Hickok, 2012; Tourville & Guenther, 2011). At present, a key challenge of speech production is disentangling different levels of planning from production proper and processing of its sensory consequences. Producing any utterance involves a continuously updated planning scope that spans upcoming segments and often the

next word, along with efference-copy/forward-model signals from motor to auditory cortices for sensory prediction and monitoring (Hickok, 2012; Indefrey & Levelt, 2004; Levelt, 1999). Consequently, neural activity around a speech-production boundary (e.g., a syllable or word) can reflect (i) syllabic sequencing and articulatory chunking, (ii) lexical selection and word-level planning, and (iii) sensory predictions and error monitoring, all partially overlapping in time. Finally, some studies suggest that syllables are the basic units of articulatory planning (Bürki et al., 2016; Michel Lange & Laganaro, 2014; Luo & Poeppel, 2007), whereas others suggest that speech planning can proceed at the word level, especially when lexical or prosodic cues are present (Assaneo et al., 2018; Piai et al., 2015).

To address this gap, in the present study we investigate EEG power changes during paced overt speech production to assess syllable-to-syllable versus word-to-word transitions. Participants produced rhythmically cued pseudoword pairs constructed from six consonant-vowel syllables while EEG and speech audio were recorded (see Figure 1A for the task timeline and pacing). For Condition A, word pairs comprised of a two-syllable word followed by a four-syllable word. For Condition B, word pairs comprised of a three-syllable word followed by another three-syllable word (see Figure 1B for examples of the two sequence types). This design enabled direct comparison of neural activity at syllable-to-syllable and word-to-word transitions in a time-locked manner (see Figure 1C-D for speech timing and envelope alignment). Based on prior work, we predicted (1) stronger theta/alpha modulation preceding syllable transitions and (2) beta modulation associated with planning across word boundaries. In EEG, slow preparatory potentials, such as the Bereitschaftspotential (also known as the readiness potential; Kornhuber & Deecke, 1965; Shibasaki & Hallett, 2006) and the contingent negative variation (CNV; Walter et al., 1964; Kononowicz & Penney, 2016), index motor and anticipatory preparation before self-initiated and temporally predicted actions. Accordingly, we focused on EEG changes around the production of the third syllable, since this syllable corresponds to a word-to-word transition in Condition A and a syllable-to-syllable transition in Condition B. Moreover, this syllable involved the consonants (/d/ and /k/), which are stop phonemes with reduced perioral electromyogram (EMG) changes that are known to contaminate scalp EEG signals (Stepp, 2012; Whitham et al., 2007). Together, this framework enables a controlled test of EEG signal differences between syllable- versus word-boundary transitions during speech production, including distinct oscillatory signatures.

2. Methodology

2.1. Participants

Twenty-one Portuguese-speaking adults (14 females, 5 males), aged between 18 and 53 years (Mean ± Standard Deviation (SD) = 29.4 ± 9.6), participated in the study. Two participants were excluded due to technical errors during the EEG-Audio recordings. Most participants were Portuguese, with a subset reporting Brazilian Portuguese or dual nationality (Portuguese/British or Portuguese/Dutch). All participants were right-handed, had completed secondary education or higher (bachelor's, master's, or doctoral level) and reported normal hearing and no history of psychiatric, neurological, or language-related disorders. The final participant demographics are summarised in Table 1. Participation was voluntary, and no compensation was given for taking part in the study. The study was performed in accordance with the approved guidelines and the Declaration of Helsinki and Oviedo Convention, with all participants providing written informed consent prior to testing.

Table 1

Participant demographics

Characteristic	Value
Total Participants	19
Gender	14 Female, 5 Male
Age range	18 - 53 years
Mean age (SD)	29.4 (9.6) years
Nationality	Primarily Portuguese; subset Brazilian Portuguese; small number dual nationality (Portuguese/British, Portuguese/Dutch)
Education	Secondary (12º) to doctoral level
Handedness	19 right-handed

Note. Values are counts unless otherwise indicated; age is reported as *M* (SD).

2.2. Stimuli

Stimuli consisted exclusively of sequences of two pseudowords designed to isolate neural activity related to syllable- and word-level transitions (see Figure 1B). Each sequence contained two pseudowords and conformed to one of two structures: a 2+4 condition (e.g., *náfa dacalána*), in which a disyllabic pseudoword was followed by a tetrasyllabic pseudoword, and a 3+3 condition (e.g., *náfada calána*), in which two trisyllabic pseudowords were presented. Both structures were matched for length (six syllables per trial). All syllables

followed a CV structure, composed from a set of 10 consonants (/b/, /d/, /f/, /m/, /n/, /s/, /t/, /k/, /l/, /g/) and the vowel (/a/). Phonetic symbols are presented in the International Phonetic Alphabet (IPA). Here, pk1-pk6 label the amplitude-envelope peaks aligned to syllables 1-6; pk3 is the peak aligned to the third produced syllable. We minimised bilabial closures at and before pk3 to reduce perioral EMG while maintaining clear tongue-tip versus tongue-dorsum articulations. Labial consonants were restricted to post-pk3 positions; no labials occurred at pk2-pk3 or at the pk3 onset. Critically, the third syllable was always tongue-driven (/d/, /k/); thus, any labial-related EMG would arise after pk3 and outside the pre-pk3 window (Whitham et al., 2007; Stepp, 2012). The pseudowords were phonotactically legal in Portuguese and carried stress on the first syllable of the first pseudoword and the penultimate syllable of the second, mirroring natural prosodic patterns while avoiding lexical or semantic familiarity effects (Arruda et al., 2023). Hence, the stress pattern of the stimuli was the same across conditions. Moreover, we minimised syllable position probability differences; that is, the probability of 'da' and 'ca' syllables were similar when present at the third syllable of a word and at the first syllable of a word. Finally, written stimuli were presented in lowercase text using a uniform sans-serif font (size 24), centred on the screen with a grey background using Presentation® software (Version 6; Neurobehavioral Systems, Inc., Berkeley, CA, USA; www.neurobs.com).

2.3. Experimental Procedure

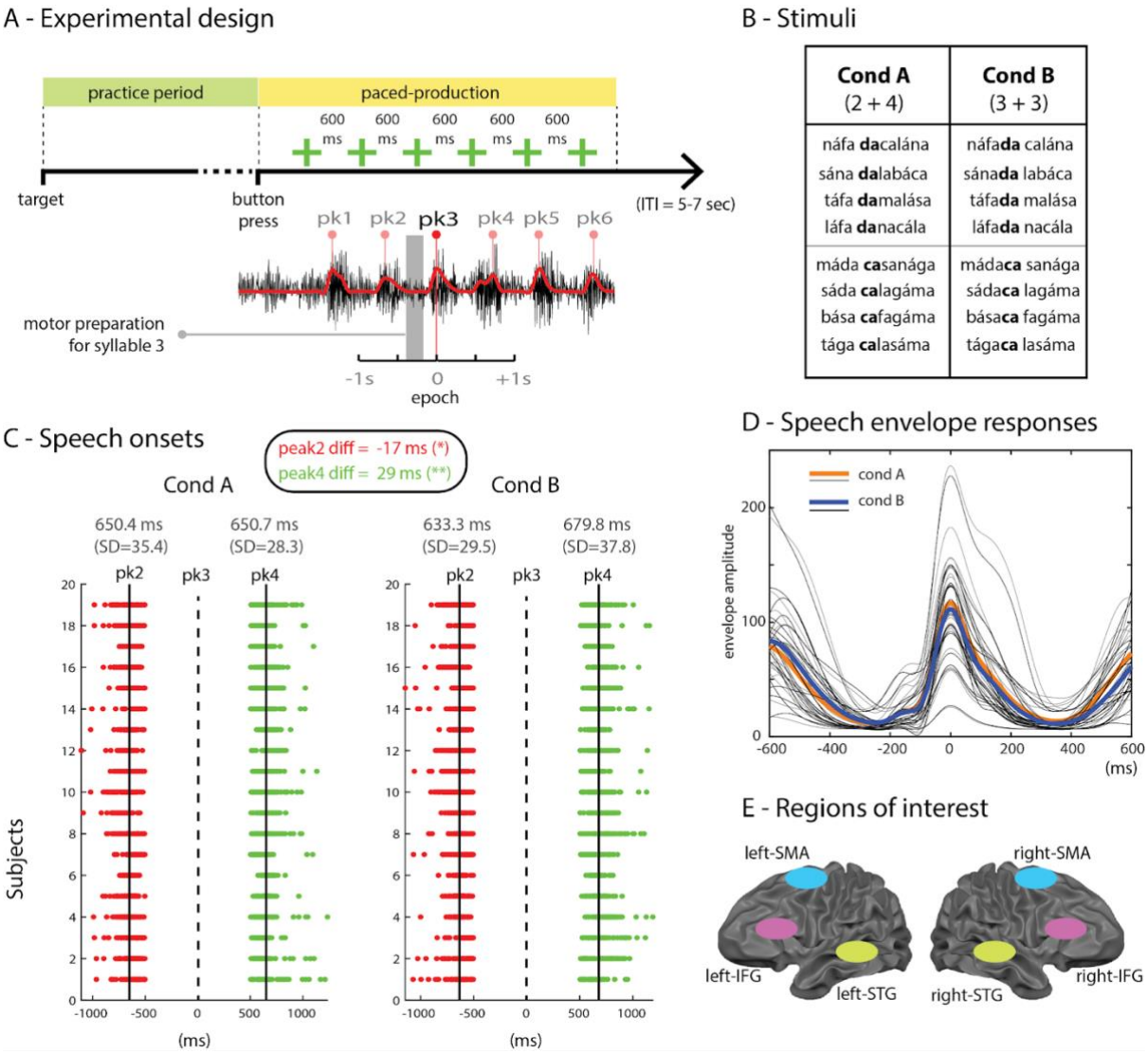
The experiment took place in an electrically shielded EEG laboratory. EEG data were recorded using a 128-channel BioSemi ActiveTwo system (BioSemi B.V., Amsterdam, the Netherlands) with electrodes positioned according to the international 10/20 system. The EEG cap was centred between the inion and the nasion anatomical fiducials on the sagittal plane, and the midpoint between the ears on the coronal plane. Conductive gel was applied to each electrode site using a blunt-tipped plastic syringe, and electrode impedance was verified to minimize offsets, in line with BioSemi system specifications. Once EEG preparation was completed, participants were seated comfortably in front of a computer screen at approximately 70 cm viewing distance.

The experiment consisted of 160 trials in total, divided into 4 blocks of 40 trials each. Each block contained five repetitions of 4 sequences from Condition A and 4 of Condition B, fully randomised. Within a block, the same sequence was never presented under Condition A and Condition B. Hence, blocks differed in how syllables were grouped, providing variation while preserving the same underlying syllabic material across blocks. The order of blocks was counterbalanced across participants, ensuring that all participants completed both conditions under comparable exposure and repetition constraints.

Following EEG cap placement and impedance checks, participants began each trial with a practice phase followed by a paced production using a visual metronome (see Figure 1A). The practice phase had the pseudoword pair displayed on the screen while participants repeatedly uttered the target sequence three times. Upon completion, participants pressed the ‘space bar’ button to signal their readiness to proceed to the paced production phase. In the paced production phase, the written sequence disappeared from the screen, and a green fixation cross appeared, flashing every 600 ms (1.67 Hz). Participants were instructed to produce one syllable per beat, aligning their speech with the visual cue rhythm. Because the lexical boundary shifts across sequence structures, the transitions flanking pk3 differ by condition. In 2+4 (Condition A), the pk2-pk3 = between-word and pk3-pk4 = within-word. In 3+3 (Condition B), pk2-pk3 = within-word and pk3-pk4 = between-word. This counterbalancing permits a within-item contrast of between- vs within-word transitions while controlling segmental content and pacing.

Figure 1

Experimental design, stimuli, speech timing, and regions of interest.



Note. A) Design. Each trial began with a practice period, followed by paced production aligned to a visual metronome (600-ms ISI). ERSP/epoch analyses were time-locked to the third-syllable peak (pk3; -1 to +1 s). ITI= 5-7 s. B) Stimuli. Two sequence types were used: Condition A = 2+4 syllables and Condition B = 3+3 syllables (six syllables total per trial). C) Speech onsets. Dots show individual syllable onsets relative to the pacing signal for each participant (red = pk2-pk3; green = pk3-pk4); vertical dashed lines mark the metronome beats. Medians are shown at the top. The within-word transition (pk2-pk3) was shorter in 3+3 than 2+4 (difference = -17 ms), $t(18) = 2.15$, $p = .045$. The between-word transition (pk3-pk4) was longer in 3+3 than 2+4 (difference = +29 ms), $t(18) = -2.90$, $p = .010$. D) Speech-envelope responses. Group speech- envelope traces time-locked to pk3 (grey = individual participants; coloured = condition means). E) Regions of interest. Bilateral inferior frontal gyrus (IFG), superior temporal gyrus (STG), and supplementary motor areas (SMA) used for hypothesis-driven visualisation. From these literature-derived, cortical ROIs (bilateral IFG, bilateral STG, bilateral SMA), only Cluster 1 (left IFG), Cluster 3 (right IFG), Cluster 8 (left STG) exhibited ERSP effects that survived

multiple-comparison control. pk = speech-envelope peak; ISI = interstimulus interval; ITI = intertrial interval. Condition A = 2+4, Condition B = 3+3. Behavioural contrasts used two-tailed paired t-tests.

2.4. Data Recording and Pre-processing

Speech responses were recorded at 16384 Hz (16-bit resolution) using a Shure SM58 microphone placed approximately 15 cm from the participants' mouth simultaneously with EEG data recording. Continuous EEG was recorded from the 128-channel BioSemi ActiveTwo system (BioSemi B.V., Amsterdam, the Netherlands), with a sampling rate of 512 Hz. Electrodes were arranged according to the international 10/20 system. The BioSemi CMS/DRL system served as the reference and ground during acquisition.

Audio data were processed to identify speech response onsets. First, speech recordings were subjected to a Hilbert transformation, and the resulting amplitude envelope was low-pass filtered at 8 Hz and resampled to 256 Hz. Within each trial, six local envelopes (one per syllable) were detected to provide syllable-level temporal markers for identifying the production of the 6 syllables composing our pseudoword pairs. Custom Matlab routines were used for audio processing, and all code is provided as appendices.

EEG preprocessing was performed in EEGLAB (Delorme & Makeig, 2004) and custom MATLAB scripts (version R2019b; MathWorks, Natick, MA). The continuous EEG recordings were first resampled to 256 Hz and high-pass filtered at 1 Hz to remove slow signal drifts. Noisy and flat-line channels were identified and removed using 'clean_rawdata' EEGLAB function, after which data were re-referenced to the average of all remaining channels (note that because we used average re-referencing, it requires prior removal of noisy channels to avoid contamination to the common average). Across datasets, a mean of 97.9 channels (*SD* = 9.6) remained following bad-channel rejection. Next, independent component analysis (ICA; INFOMAX) was performed, and components reflecting ocular, muscular, or other stereotypical artefacts were automatically identified and removed using 'ICLabel' and 'ICFlag' EEGLAB functions (Pion-Tonachini et al., 2019). Channels previously excluded were subsequently interpolated back into the dataset using the 'interp' method. Following ICA cleaning, the data were low-pass filtered at 70 Hz and a band-stop (48-52 Hz) filter was applied to further suppress line noise.

We segmented the EEG into epochs from -1 to +1 s centred on the third-syllable production event (i.e., pk3). Time-locking our EEG analysis to the third syllable minimises contamination from sequence-onset preparatory motor activity, while still providing symmetric context for preceding (pk2-pk3) and following (pk3-pk4) transitions. Although the speech-envelope peak can lag the true acoustic onset by a small amount, the same alignment procedure and fixed pacing were used in both conditions, and behavioural timing (Figure 1D)

confirmed consistent alignment across participants. This ensures that any between-condition differences in ERSPs are expected to reflect the manipulation rather than potential cross-condition misalignments.

Event-related spectral perturbations (ERSPs) were computed for each trial and independent component (IC), providing a time-frequency representation of oscillatory power changes across conditions.

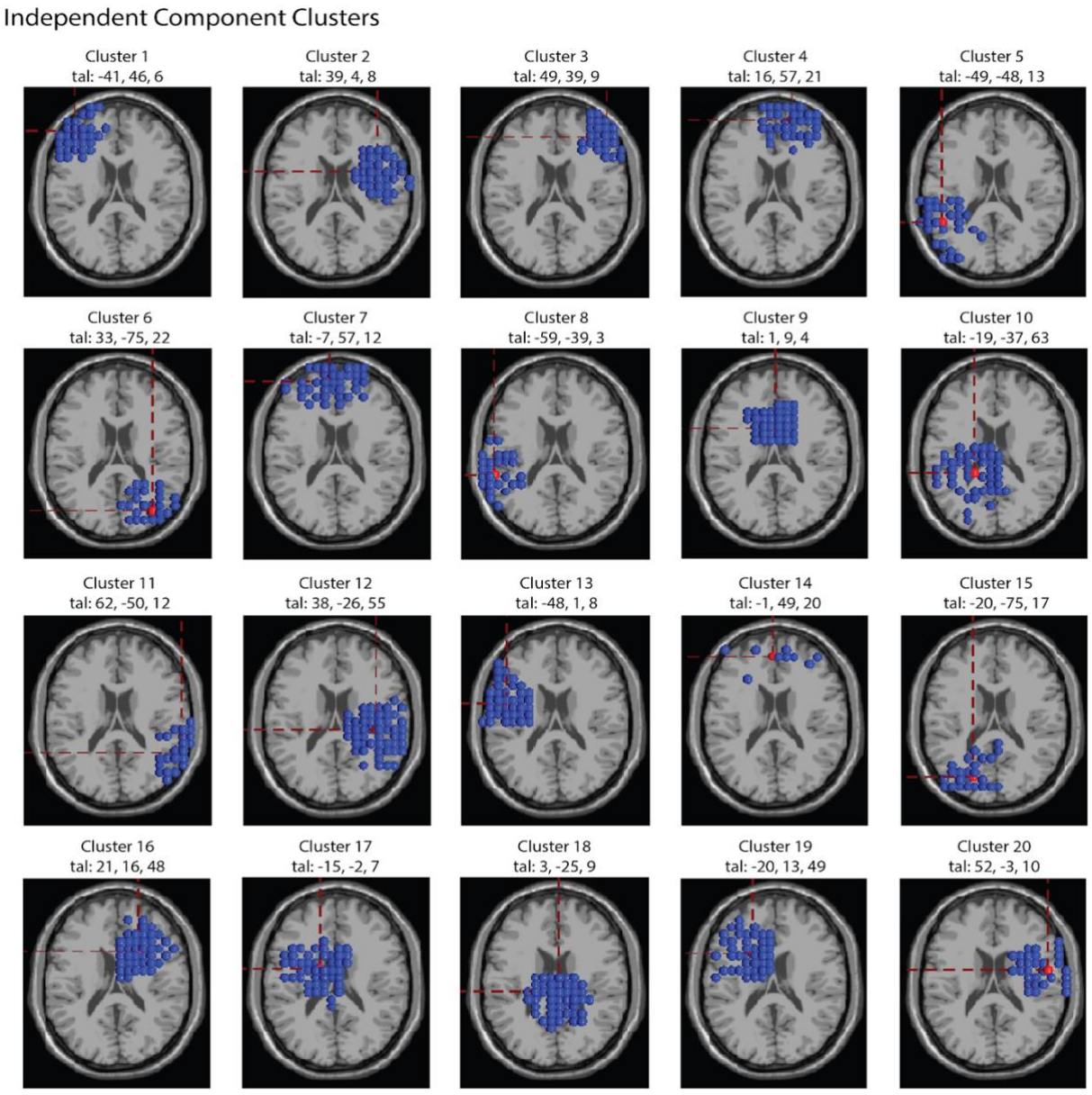
2.5. Data Analysis and Statistics

2.5.1. Independent Component Analysis and Source Clustering

After ICA, equivalent current dipoles were estimated for each independent component using a standard boundary element head model. Components with residual variance greater than 15% were excluded. Next, dipoles were computed for each IC component and participant. Finally, 20 clusters were made by matching dipole locations and spectral activity between 3 and 40 Hz (see Figure 2). With this approach, we took maximum advantage of the high-density EEG montage (i.e., 128 channels), by performing group statistics at the EEG source level instead of relying on EEG-channel signals, which are known to have low spatial specificity (Delorme & Makeig, 2004).

Figure 2

Independent Component (IC) clusters.



Note. Equivalent current dipoles (blue spheres) from all participants, after artifact IC rejection (residual variance ≤ 15%), were grouped into 20 clusters (Clusters 1-20) based on scalp maps, spectra, and dipole locations. Each panel shows an axial slice of the template brain with the cluster centroid marked by crosshairs; Talairach coordinates of the centroid are given ("tal: x, y, z"). Cluster sizes ranged from 19 to 197 ICs (M = 93.1, SD = 40.0).

2.5.2. Time-Frequency Analysis

ERSPs were averaged within each cluster (note that each cluster is composed of a combination of several ICs from participants) and assessed statistically using a permutation test (based on 2,000 condition-label permutations). For hypothesis-driven visualisation we report clusters corresponding to bilateral IFG, bilateral STG, and bilateral SMA (Figure 1E); example ERSPs are shown for L-IFG (Cluster 1), R-IFG (Cluster 3), and L-STG (Cluster 8) (Figure 3). ERSPs for all 20 clusters are provided in Appendix A (Figure A1). Multiple comparisons were controlled using Benjamini-Hochberg FDR (q = .05) across time-frequency bins (200 time-points x 68 frequencies). Behavioural contrasts were evaluated with two-tailed, paired t-tests Time-frequency contrasts were assessed with bin-wise permutation tests (2,000 label permutations) and reported after FDR correction (q < .05). We pre-specified bilateral IFG, bilateral STG, and bilateral SMA as confirmatory ROIs. Within these ROI clusters, permutation p-values were FDR-corrected across TF bins (q < .05) but not further corrected across clusters.

We refer to decreases in band-limited power as event-related desynchronisation (ERD) and increases as event-related synchronisation (ERS), reporting frequency bands and time windows relative to a pre-onset baseline (-500 to -200 ms). Consistent with our hypothesis-driven ROI plan (Figure 1E), statistical inference focused on cortical ROIs identified from the literature (bilateral IFG, bilateral STG, bilateral SMA). Unless noted otherwise, we report clusters whose ERSP contrasts survived FDR correction (q < .05); among these ROIs, significant effects were confined to Cluster 1 (left IFG), Cluster 3 (right IFG), Cluster 8 (left STG).

3. Results

3.1. Behavioural Results

Participants successfully aligned their speech production with the visual metronome (see Figure 1C-D). Because the lexical boundary differs by condition, in the 2+4 (Condition A), pk2-pk3 is a between-word syllable transition while in the 3+3 (Condition B), pk2-pk3 is a within-word syllable transition. Timing between successive syllables around pk3 differed by condition. The pk2-pk3 interval was shorter in 3+3 than in 2+4 (diff.= -17 ms, SD_diff = 34.49; paired t-test), $t(18) = 2.15$, $p = .045$, dz = 0.49, 95% CI [-33.62,-0.38] ms . The pk3-pk4 interval was longer in 3+3 than in 2+4 (diff.= +29 ms, SD_diff = 43.75), $t(18) = -2.90$, $p = .010$, dz =

0.67, 95% CI [7.91, 50.09] ms. Group speech-envelope traces time-locked to pk3 confirmed consistent alignment of the third-syllable envelope peak across participants and conditions (Figure 1D).

3.2. Event-Related Spectral Perturbations (ERSPs)

ERSPs were analysed within the three clusters selected (Clusters 1, 3, and 8; see Figure 2). EEG statistics are based on permutation-based condition tests (2,000 label permutations) with Benjamini-Hochberg FDR control across time-frequency bins (q = .05). Consistent with our analysis plan, bilateral IFG, bilateral STG, and bilateral SMA were treated as confirmatory ROIs (FDR within TF bins). In line with our regions of interest (ROI) plan, we report illustrative clusters corresponding to bilateral inferior frontal gyrus (IFG), bilateral superior temporal gyrus (STG), and supplementary motor area (SMA; Figure 1E). Of the literature-derived cortical ROIs, only Cluster 1 (left IFG), Cluster 3 (right IFG), and Cluster 8 (left STG) showed FDR-corrected ERSP differences between conditions (see Table 2); no other ROI clusters yielded corrected effects. For completeness, ERSPs for all 20 clusters are shown in Appendix A (Figure A1).

3.2.1. Frontal regions

In the right IFG (Cluster 3; tal: 49, 39, 9), a robust pre-pk3 modulation survived FDR correction from ~ -500 to 0 ms , spanning ~ 4-30 Hz (theta/alpha into low beta; Figure 3B, FDR panel). Power modulation was greater for 3+3 than 2+4, with additional weaker patches evident at the uncorrected level in the early post-pk3 window. By contrast, the left IFG (Cluster 1; tal: -41, 46, 6) showed largely similar ERSPs across conditions; only a small, isolated FDR-significant alpha patch (~ 10-14 Hz) appeared around -450 ms (Figure 3A), and broader differences did not survive correction.

3.2.2. Temporal regions

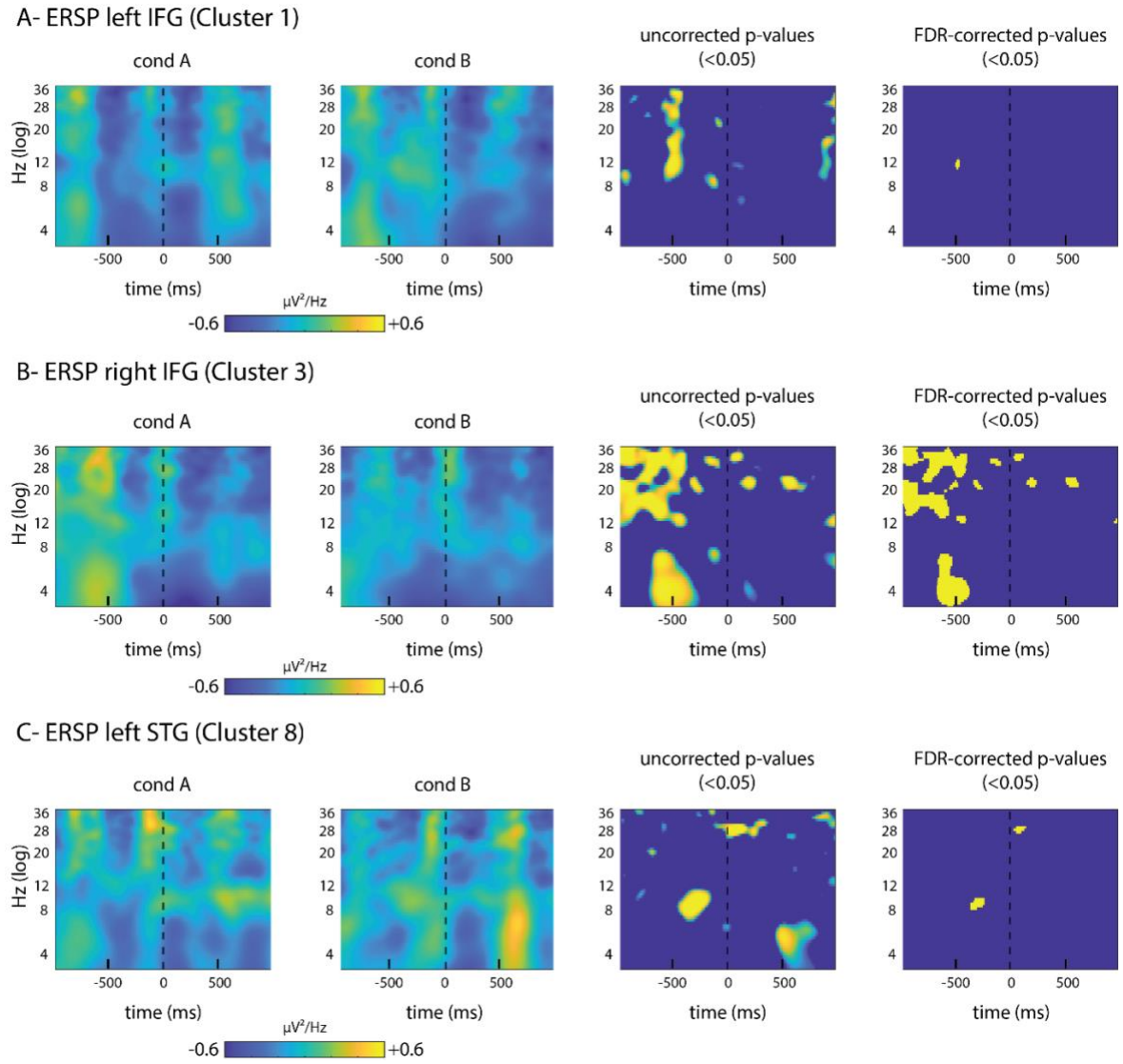
In the left STG (Cluster 8; tal: -59, -39, 3), we observed two brief FDR-significant islands (Figure 3C): a pre-pk3 alpha patch (~ 9-11 Hz, -400 to -300 ms) and a short post-pk3 high-beta patch (~ 26-30 Hz, 0 to +100 ms). Outside these windows, no sustained corrected effects were present. The right STG showed no FDR-corrected differences (see Table 2).

3.2.3. Summary across clusters

Considering all 20 clusters, the clearest FDR-corrected modulation emerged in right IFG, indicating a right-lateralised frontal preparatory signature preceding pk3. Left IFG and left STG contributed only limited, short-lived patches, and no other clusters yielded reliable corrected effects. See Table 2 for a tabular summary of the significant ERSP contrasts in the confirmatory ROIs. See Appendix A (Figure A1) for ERSPs of all 20 clusters.

Figure 3

ERSPs in frontal and temporal clusters.



Note. (A) Cluster 1- left inferior frontal gyrus (tal: -41, 46, 6). (B) Cluster 3- right inferior frontal gyrus (tal: 49, 39, 9). (C) Cluster 8- left superior temporal gyrus (tal: -59, -39, 3). For each cluster, the first two subpanels show ERSPs time-locked to pk3 (vertical dashed line; window -1 to +1 s) for Condition A and Condition B. The third and fourth subpanels display uncorrected ($p < .05$) and FDR-corrected ($p < .05$)

permutation-based condition contrasts; yellow indicates significant bins. Frequency is shown on a log scale (~ 0-36 Hz); power is baseline-normalised (-500 to -200 ms). Right IFG shows prominent pre-pk3 differences that survive FDR correction; left IFG contains a small isolated alpha patch; left STG shows two small significant islands (pre-pk3 alpha and brief post-pk3 high-beta). Condition contrasts reflect permutation tests (2,000 label permutations) displayed uncorrected ($p < .05$) and FDR-corrected ($q < .05$).

Table 2

Clusters with FDR-corrected ERSP condition differences in confirmatory ROIs (IFG, STG).

Cluster	Cluster 1	Cluster 3	Cluster 8
Approx. Location	Left inferior frontal gyrus (IFG)	Right inferior frontal gyrus (IFG)	Left superior temporal gyrus (STG)
Talairach Coord.	−41, 46, 6	49, 39, 9	−59, −39, 3
Cluster Size (ICs)	111	114	59
Frequency band(s)	~10–14 Hz (alpha)	~4–30 Hz (theta/alpha/low-beta)	(i) ~9–11 Hz (alpha); (ii) ~26–30 Hz (high-beta)
Time (ms)	~−450	−500 to 0	(i) −400 to −300; (ii) 0 to +100
Condition Effect	Small patch	B > A (greater modulation in 3+3)	Two small patches
Significance	FDR < .05	FDR < .05	FDR < .05
Peak Window	pk2-pk3	pk2-pk3	(i) pk2–pk3; (ii) pk3–pk4

Note. Values summarise time-frequency contrasts between Condition B (3+3) and Condition A (2+4) within independent-component clusters. Significance reflects Benjamin-Hochberg FDR ($q = .05$) across time-frequency bins within the cluster. Time (ms) is relative to pk3 (the third-syllable alignment point; window -1 to +1 s). Peak window labels the transition surrounding pk3: pk2-pk3 (pre-pk3) vs pk3-pk4 (post-pk3). ERSP = event-related spectral perturbation; IC = independent component; IFG = inferior frontal gyrus; STG = superior temporal gyrus; n.s. = not significant.

4. Discussion

Participants produced syllables in tight synchrony with the 600 ms metronome, yielding broadly comparable paced speech production across conditions (Figure 1C-D). Clustering analyses identified three clusters that showed significant results: Cluster 1 (left IFG), Cluster 3 (right IFG), and Cluster 8 (left STG). The importance of these clusters was previously predicted in the literature. Across these clusters, we observed ERSP differences between conditions that survived multiple-comparison correction (FDR $q < .05$), with robust pre-pk3 modulation most prominent in right IFG and briefer patches in left IFG and left STG (Figure 3). As a cautionary note, although productions were paced to the metronome, the two conditions elicited small but statistically significant timing differences. Quantitatively, transition timing diverged by lexical structure (Figure 1C-D): the within-word interval (pk2-pk3) was significantly shorter in condition B (3+3) than condition A (2+4), whereas the between-word interval (pk3-pk4) was significantly longer in condition B than condition A. This asymmetry suggests that while within-word sequencing was slightly expedited, crossing a lexical boundary incurred additional planning time, consistent with classical accounts of boundary-related planning costs in production (Wheeldon & Lahiri, 1997).

Neural dynamics mirrored this behavioural asymmetry and were regionally selective. In right inferior frontal gyrus (IFG; Cluster 3; tal: 49, 39, 9), we observed a robust pre-pk3 modulation spanning approximately 4-30 Hz (theta/alpha into low-beta) from approximately -500 to 0 ms, that survived FDR correction, with greater power modulation for 3+3 than 2+4 (Figure 3B, FDR panel). Because pk3 indexes within-word sequencing in 3+3 and boundary-adjacent processing in 2+4, the stronger right-frontal engagement for 3+3 indicates that frontal control networks adjust preparatory states according to the hierarchical role of the upcoming syllable, maintaining within-word chunking versus reconfiguration at a boundary. Importantly, because this modulation unfolds ~ 500-0 ms before articulation of pk3, and not before the boundary-crossing interval in 3+3 (pk3-pk4), it is unlikely to reflect look-ahead planning of the next word. Rather, the pattern aligns with reports that right-frontal beta indexes preparatory set/maintenance and sequential control during speech (Pfurtscheller & Lopes da Silva, 1999; Piai et al., 2015; Weiss & Müller, 2012), suggesting stronger motor-set maintenance within a word (3+3) and greater set reconfiguration at a lexical boundary (2+4). The broadband nature of this effect (theta/alpha into low beta) is compatible with joint sequencing/gating operations often linked to lower frequencies and with motor-set or top-down maintenance associated with beta rhythms (Arnal & Giraud, 2012; Giraud & Poeppel, 2012; Pfurtscheller & Lopes da Silva, 1999; Piai et al., 2015; Weiss & Müller, 2012). In contrast, left IFG (Cluster 1; tal: -41, 46, 6) exhibited only a small, isolated alpha patch and left STG (Cluster 8; tal: -59, -39, 3) two brief

islands (pre-pk3 alpha; early post-pk3 high beta), indicating comparatively limited, event-locked sensory adjustments in this pk3-locked window (Ding et al., 2016; Luo & Poeppel, 2007). Importantly, within our literature-derived cortical ROI set (Figure 1E), significant ERSP modulations were restricted to left IFG (Cluster 1), right IFG (Cluster 3), and left STG (Cluster 8), with no FDR-corrected effects in the remaining cortical ROIs.

Building on this pattern in right IFG (Cluster 3), we note that the right inferior frontal cortex is frequently implicated in inhibitory control and rapid state re-initialisation via hyperdirect projections to the subthalamic nucleus, often expressed in the beta range (~ 21-35 Hz) and in cortico-STN coupling (Aron et al., 2014; Oswal et al., 2021; Swann et al., 2009; Wessel & Aron, 2017). Within this framework, the pre-pk3 beta modulation can be interpreted as a motor reset: when pk3 immediately precedes a lexical boundary (Condition A, 2+4), the system must terminate the current chunk and configure the next, predicting relatively stronger right-frontal beta than when pk3 occurs within a continuing word (Condition B, 3+3). This account complements the preparatory-set interpretation above and ties the Cluster 3 effect to pathway-specific control in the cortico-basal ganglia-thalamo-cortical loop.

More broadly, the present data align with accounts in which neural oscillations provide a temporal scaffold for hierarchical speech production: lower-frequency activity supports sequencing and temporal gating, whereas beta indexes preparatory set and coordination across units (Arnal & Giraud, 2012; Giraud & Poeppel, 2012; Pfurtscheller & Lopes da Silva, 1999; Piai et al., 2015; Weiss & Müller, 2012). At the same time, because corrected effects concentrated in right IFG and other clusters showed only brief patches, we refrain from inferring a global pattern (e.g. theta-dominant within-word and beta/gamma-dominant between-word responses across widespread regions). Instead, we interpret the findings as evidence for a right-lateralised frontal preparatory signature that is sensitive to whether the upcoming syllable completes a word or precedes a boundary, with only transient accompanying modulations in the auditory cortex. This interpretation is consistent with neurocognitive models in which frontal-motor systems set predictive control states for upcoming speech onsets while auditory areas provide brief, event-locked monitoring and adjustment (Hickok, 2012; Indefrey & Levelt, 2004; Levelt, 1999; Tourville & Guenther, 2011).

Several features of the design strengthen these inferences. Locking analysis to pk3 minimised contamination from initial sequence onset and enabled a balanced within- versus between-word contrast without changing segmental content or pacing; this is important when asking whether timing differences reflect hierarchical planning rather than stimulus differences (Indefrey & Levelt, 2004; Levelt, 1999). Source-resolved independent-component clustering provided a consistent cross-participant reference, improving anatomical interpretability

(Delorme et al., 2012; Makeig et al., 2004). Methodological choices also reduced common confounds: the use of tongue-driven consonants avoided bilabial closures that elevate perioral EMG above ~ 20 Hz, and ICA/ICLabel-based rejection further mitigated myogenic artefacts (Stepp, 2012; Whitham et al., 2007). The tight alignment of speech envelopes across conditions and the pre-onset locus of the principal frontal effect argue against explanations based on acoustic differences alone.

Limitations warrant caution. First, the sample size (N = 19) is typical for EEG but limits sensitivity to smaller effects, especially after multiple-comparison correction. Second, although the exploratory survey across 20 clusters was summarised descriptively, multiplicity remains a concern; accordingly, we emphasised a priori regions of interest (bilateral IFG, STG, SMA). Third, the within-word interval was shorter by ~ 17 ms in 3+3 and the between-word interval longer by ~ 29 ms in 3+3; such timing asymmetries could in principle influence ERSP estimates, though pk3-locking and identical alignment across conditions mitigate this. Future work could match trials on interval length or include trial-wise interval duration as a covariate to rule out residual confounding. Fourth, strict pacing and pseudowords enhance control but constrain generalisability to natural, prosodically rich and semantically driven speech (Assaneo & Poeppel, 2018). Fifth, because labial segments, when present, occurred after the third-syllable lock, any perioral EMG related to labial closure would arise predominantly in post-pk3 windows and therefore cannot explain the pre-onset right-IFG effect. Finally, scalp EEG has limited spatial resolution for subcortical timing circuits (basal ganglia, cerebellum) known to contribute to sequencing and set maintenance (Klostermann & Ehlen, 2020).

Altogether, the behavioural and neural data indicate that hierarchical structure shapes both timing and pre-speech cortical dynamics during production. The right-lateralised frontal signature preceding pk3 suggests predictive control mechanisms tuned to whether the upcoming syllable completes a word or indicates a new one, complementing accounts of hierarchical temporal scaffolding in speech perception and production (Ding et al., 2016; Giraud & Poeppel, 2012; Piai et al., 2015). Future work should test scaling with speech rate, manipulate boundary strength (lexical frequency, prosodic cues), relate trial-wise frontal power to transition timing, and examine cross-frequency interactions (e.g. theta-beta coupling) alongside inter-areal/cluster coupling (e.g. IFG-STG and IFG-SMA phase-synchrony/coherence) to assess whether hierarchical control is implemented via nested oscillations and network-level coordination (Arnal & Giraud, 2012). In addition, combining cross-frequency, and connectivity-based metrics with single-trial timing would test whether stronger frontal-temporal coupling predicts slower boundary crossings and faster within-word sequencing. Combining source-resolved EEG/MEG with perturbation of auditory feedback

would further clarify how frontal predictions interact with sensory monitoring at boundaries (Hickok, 2012; Tourville & Guenther, 2011).

5. Conclusion

Speakers maintained precise synchrony with the pacing signal, yet their timing systematically reflected lexical structure: within-word transitions were faster and between-word transitions slower in 3+3 than 2+4. Mirroring this behavioural asymmetry, right inferior frontal gyrus showed robust pre-onset modulation spanning theta/alpha into low-beta (approximately 4-30 Hz) that survived multiple-comparison correction, whereas left IFG and left STG exhibited only brief, isolated patches. Together, these findings indicate that hierarchical organisation in production shapes both temporal behaviour and pre-speech cortical dynamics, consistent with accounts in which frontal control networks set predictive states for articulatory sequencing and sensory cortices adjust transiently around the acoustic event. Within the a priori, cortical ROI set, significant ERSP effects were limited to left/right IFG and left STG (Clusters 1, 3, 8).

Methodologically, time-locking to the third syllable (pk3) and source-resolved ICA clustering provided a controlled test of within- versus between-word effects while minimising onset-related confounds and myogenic artefact. Nonetheless, the modest sample, the use of paced pseudowords, and the exploratory nature of cluster-wise surveys temper generalisability. Future work should manipulate boundary strength and speech rate, relate trial-wise frontal power to transition timing, test cross-frequency coupling (e.g., theta-beta), and combine EEG/MEG with feedback perturbation to delineate how predictive frontal interacts with sensory monitoring at lexical boundaries. Overall, the results support a rhythm-based account of speech production in which right-lateralised frontal activity provides the temporal scaffold for hierarchical planning.

6. References

Aiken, S. J., & Picton, T. W. (2008). Human cortical responses to the speech envelope. *Ear and Hearing, 29*(2), 139–157. https://doi.org/10.1097/AUD.0b013e31816453dc

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

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences, 18*(4), 177–185. https://doi.org/10.1016/j.tics.2013.12.003

Assaneo, M. F., & Poeppel, D. (2018). The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech–motor rhythm. *Proceedings of the National Academy of Sciences, 115*(21), E5368–E5377. https://doi.org/10.1073/pnas.1717503115

Assaneo, M. F., Ripollés, P., Orpella, J., Lin, W. M., de Diego-Balaguer, R., & Poeppel, D. (2019). Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nature Neuroscience, 22*(4), 627–632. https://doi.org/10.1038/s41593-019-0353-z

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289–300.

Blomert, L. (2011). The neural signature of orthographic–phonological binding in reading. *Trends in Cognitive Sciences, 15*(9), 372–380. https://doi.org/10.1016/j.tics.2011.06.015

Bürki, A., Sadat, J., Dubarry, A.-S., & Alario, F.-X. (2016). Sequential processing during noun phrase production. *Cognition, 146*, 90–99. https://doi.org/10.1016/j.cognition.2015.09.002

Cao, C., Delavallée, R., Wójcik, A., Chéret, P., Benazzouz, A., Modolo, J., Rizzi, F., & Denninger, R. (2024). Low-beta versus high-beta band cortico-subcortical coherence in movement inhibition and expectation. *Neurobiology of Disease, 201*, 106689. https://doi.org/10.1016/j.nbd.2024.106689

Christoffels, I. K., Formisano, E., & Schiller, N. O. (2007). Neural correlates of verbal working memory in bilinguals: An fMRI study. *Human Brain Mapping, 28*(2), 115–125. https://doi.org/10.1002/hbm.20263

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

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

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience, 19*(1), 158–164. https://doi.org/10.1038/nn.4186

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience, 15*(4), 511–517. https://doi.org/10.1038/nn.3063

Herz, D. M., Tan, H., Brittain, J.-S., Fischer, P., Cheeran, B., Green, A. L., Aziz, T. Z., Ashkan, K., Foltynie, T., Limousin, P., Zrinzo, L., Hariz, M., Bogacz, R., & Brown, P. (2023).

Dynamic modulation of subthalamic nucleus activity facilitates adaptive behavior. *PLOS Biology, 21*(6), e3002140. https://doi.org/10.1371/journal.pbio.3002140

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience, 13*(2), 135–145. https://doi.org/10.1038/nrn3158

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

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8*(5), 393–402. https://doi.org/10.1038/nrn2113

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1–2), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001

Klostermann, F., & Ehlen, F. (2020). Basal ganglia contributions to speech motor control: Evidence from neuromodulation and disease. *Parkinsonism & Related Disorders, 73*, 9–21. https://doi.org/10.1016/j.parkreldis.2020.03.018

Kononowicz, T. W., & Penney, T. B. (2016). The contingent negative variation (CNV): Timing isn't everything. *Current Opinion in Behavioral Sciences, 8*, 231–237. https://doi.org/10.1016/j.cobeha.2016.02.022

Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Archiv für die gesamte Physiologie des Menschen und der Tiere, 284*, 1–17. https://doi.org/10.1007/BF00412364

Levelt, W. J. M. (1999). Models of word production. In C. M. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 83–122). Oxford University Press.

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron, 54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences, 359*(1457), 1521–1532. https://doi.org/10.1098/rstb.2004.1515

Michel Lange, V., & Laganaro, M. (2014). *Multi-word context modulates lexical–phonological planning scope: Evidence from syllable-frequency effects*. Frontiers in Psychology, 5, 120. https://doi.org/10.3389/fpsyg.2014.00120

Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., & Carreiras, M. (2018). Delta (but not theta)–band cortical entrainment involves speech-specific processing. *European Journal of Neuroscience, 48*(7), 2642–2650. https://doi.org/10.1111/ejn.13811

Mullen, T., Kothe, C., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., Jung, T.-P., & Cauwenberghs, G. (2015). Real-time neuroimaging and cognitive monitoring using wearable dry EEG. *IEEE Transactions on Biomedical Engineering, 62*(11), 2553–2567. https://doi.org/10.1109/TBME.2015.2481482

Orpella, J., Etchell, A., de Diego-Balaguer, R., & Sebastian, R. (2024). Reactive inhibitory control precedes overt stuttering events. *Neurobiology of Language, 5*(2), 432–455. https://doi.org/10.1162/nol_a_00155

Oswal, A., Cao, C., Yeh, C.-H., Neumann, W.-J., Gratwicke, J., Akram, H., Horn, A., Li, D., Zhan, S., Zhang, C., Wang, Q., Zrinzo, L., Foltynie, T., Limousin, P., Bogacz, R., Sun, B., Husain, M., Brown, P., & Litvak, V. (2021). Neural signatures of hyperdirect

pathway activity in Parkinson's disease. *Nature Communications, 12*, 5185. https://doi.org/10.1038/s41467-021-25366-0

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology, 110*(11), 1842–1857. https://doi.org/10.1016/S1388-2457(99)00141-8

Piai, V., Roelofs, A., Rommers, J., & Maris, E. (2015). Beta oscillations reflect memory and motor aspects of spoken word production. *Human Brain Mapping, 36*(7), 2767–2780. https://doi.org/10.1002/hbm.22806

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated EEG independent component classifier, dataset, and website. *NeuroImage, 198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Rapela, J. (2017). *Entrainment of traveling waves to rhythmic motor acts* (arXiv:1606.02372). arXiv. https://doi.org/10.48550/arXiv.1606.02372

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential? *Clinical Neurophysiology, 117*(11), 2341–2356. https://doi.org/10.1016/j.clinph.2006.04.025

Stepp, C. E. (2012). Surface electromyography for speech and swallowing systems: Measurement, analysis, and interpretation. *Journal of Speech, Language, and Hearing Research, 55*(4), 1232–1246. https://doi.org/10.1044/1092-4388(2011/11-0214)

Strauß, A., Wöstmann, M., & Obleser, J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Frontiers in Human Neuroscience, 8*, 350. https://doi.org/10.3389/fnhum.2014.00350

Swann, N., Tandon, N., Canolty, R., Ellmore, T. M., McEvoy, L. K., Dreyer, S., DiSano, M., & Aron, A. R. (2009). Intracranial EEG reveals a time- and frequency-specific role for the right inferior frontal gyrus and primary motor cortex in stopping initiated responses. *Journal of Neuroscience, 29*(40), 12675–12685. https://doi.org/10.1523/JNEUROSCI.3359-09.2009

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system—An approach to cerebral imaging*. Georg Thieme.

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes, 26*(7), 952–981. https://doi.org/10.1080/01690960903498424

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. *Nature, 203*(4953), 380–384. https://doi.org/10.1038/203380a0

Weiss, S., & Müller, H. M. (2012). "Too many betas do not spoil the broth": The role of beta brain oscillations in language processing. *Frontiers in Psychology, 3*, 201. https://doi.org/10.3389/fpsyg.2012.00201

Wessel, J. R., & Aron, A. R. (2017). On the globality of motor suppression: Unexpected events and their influence on behavior and cognition. *Neuron, 93*(2), 259–280. https://doi.org/10.1016/j.neuron.2016.12.013

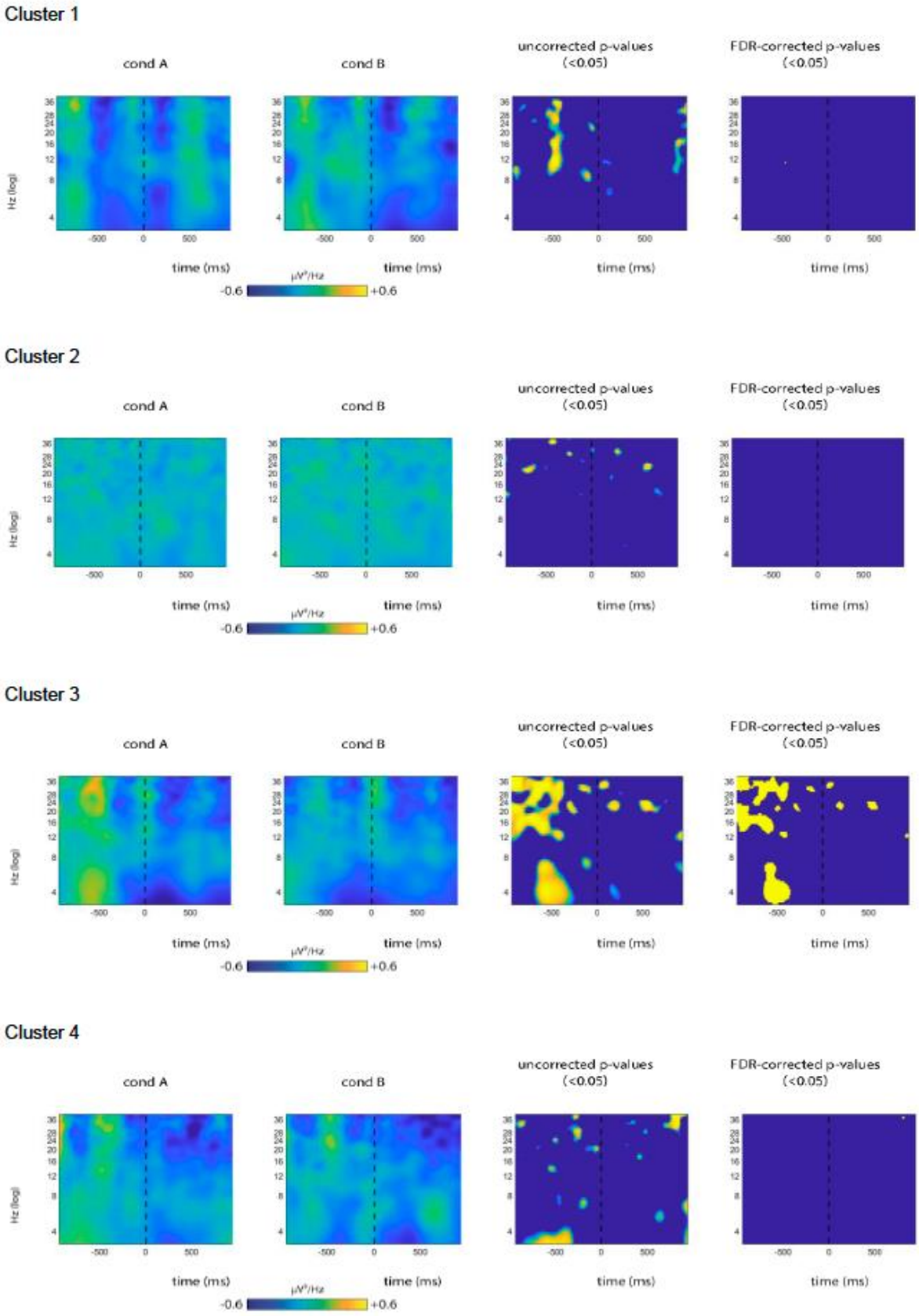
Wheeldon, K., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language, 37*(3), 356–381. https://doi.org/10.1006/jmla.1997.2517

Whitham, E. M., Pope, K. J., Fitzgibbon, S. P., Lewis, T., Clark, C. R., Loveless, S., Broberg, M., Wallace, A., DeLosAngeles, D., Lillie, P., Hardy, A., Fronsko, R., Pulbrook, A., & Willoughby, J. O. (2007). Scalp electrical recording of the EEG is contaminated by EMG. *Clinical Neurophysiology, 118*(8), 1877–1885. https://doi.org/10.1016/j.clinph.2007.04.019

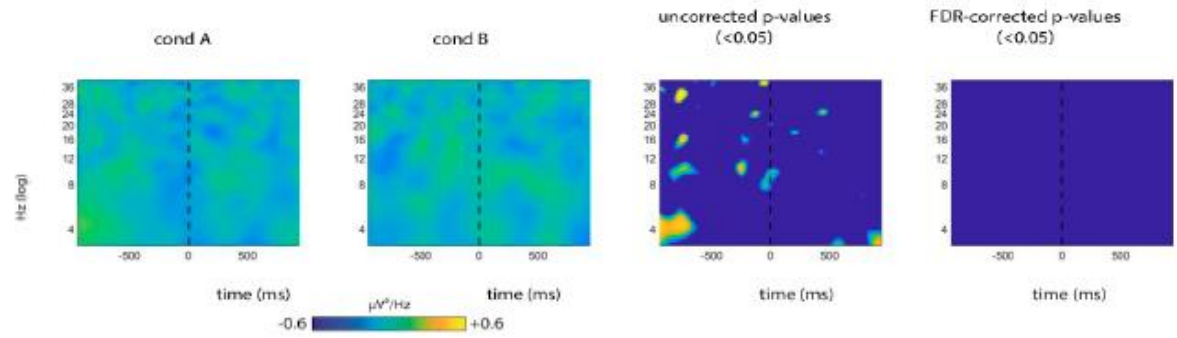
Zoefel, B., & Kösem, A. (2024). Neural tracking of continuous acoustics: Properties, speech-specificity and open questions. *European Journal of Neuroscience, 59*(3), 394–414. https://doi.org/10.1111/ejn.16221

7. Appendix

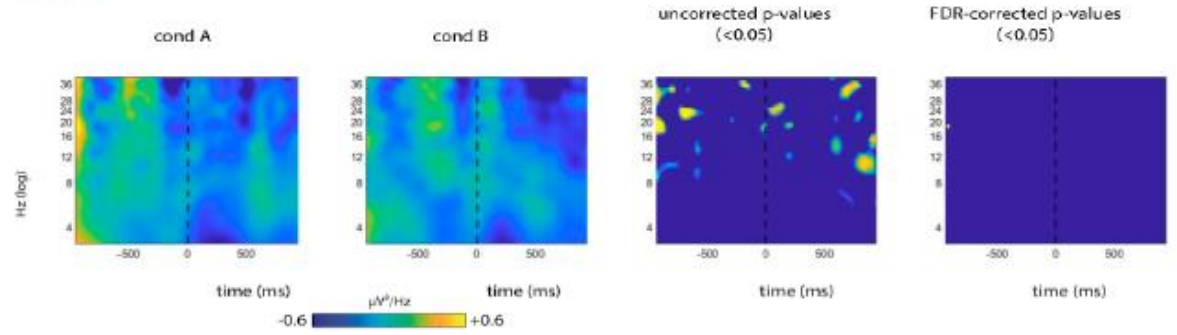
7.1. Appendix A. Figure A1. ERSPs of Clusters 1-20.



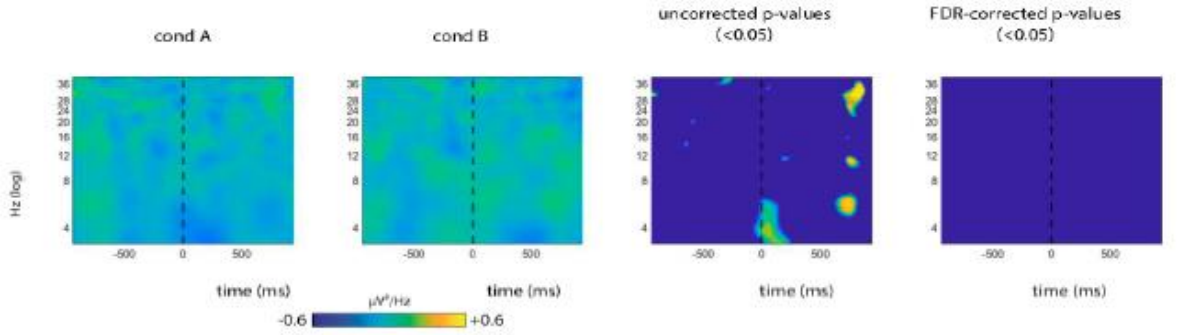
Cluster 5



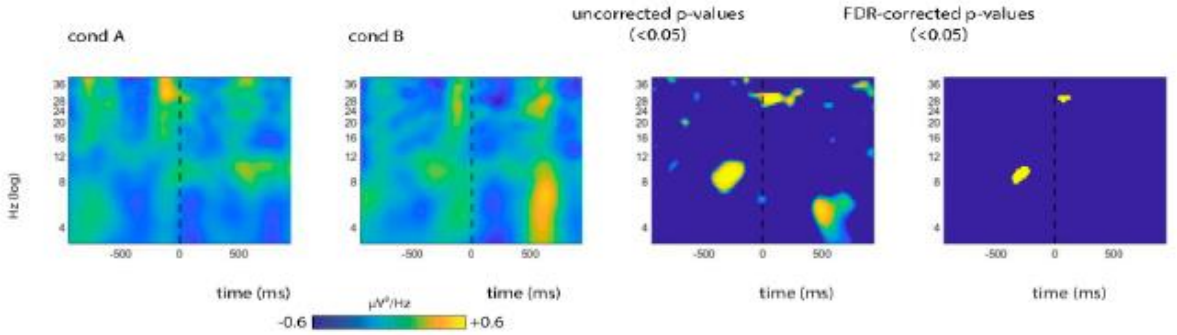
Cluster 6



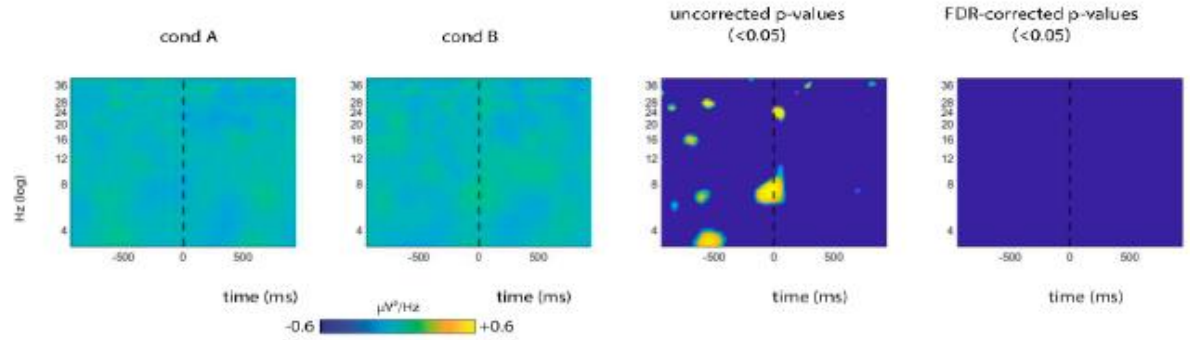
Cluster 7



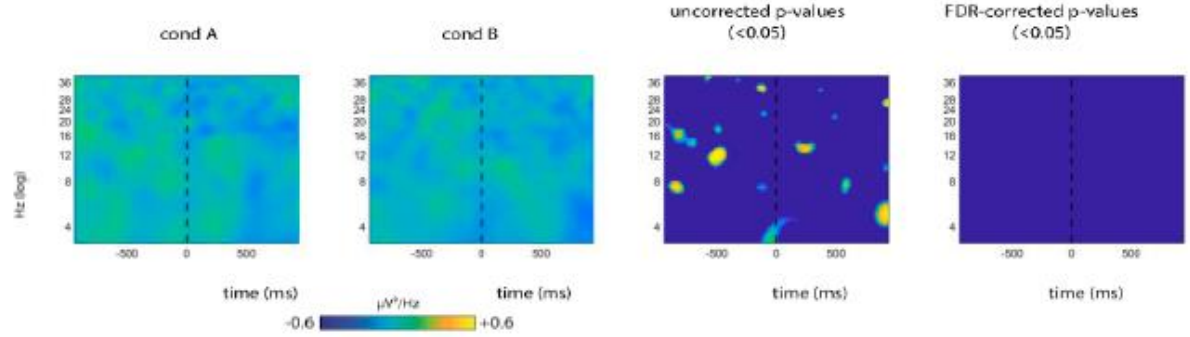
Cluster 8



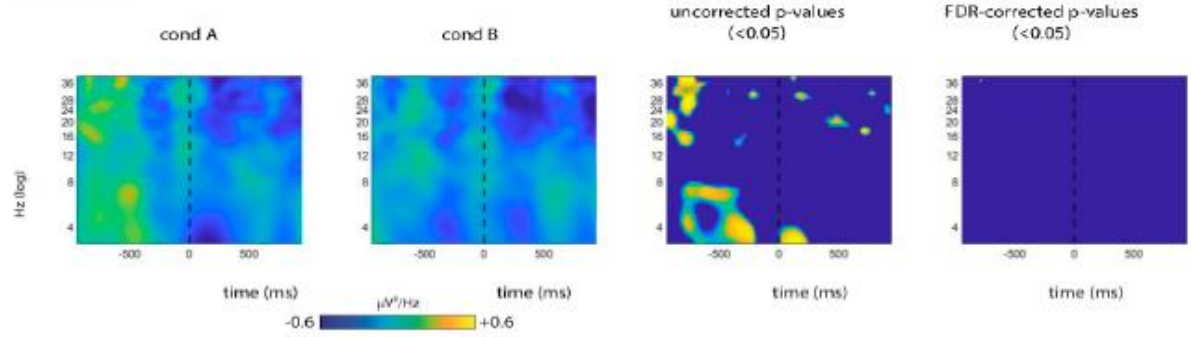
Cluster 9



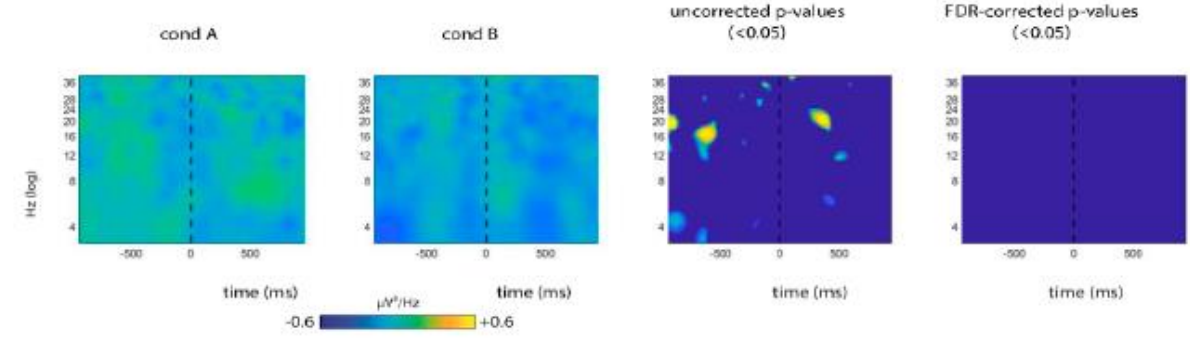
Cluster 10



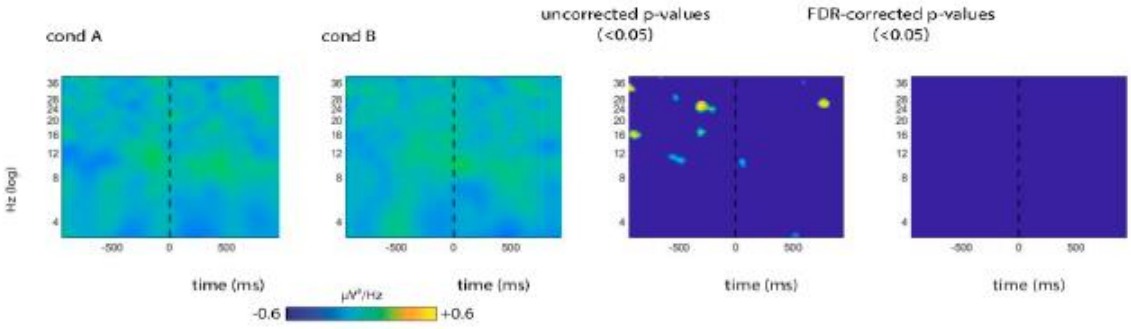
Cluster 11



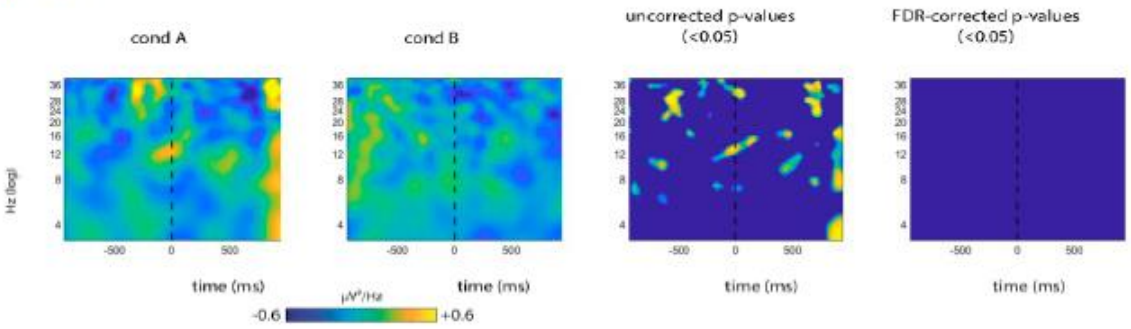
Cluster 12



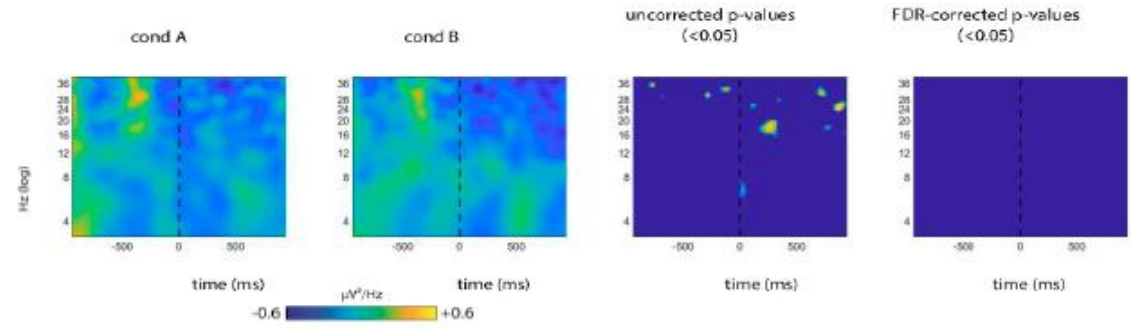
Cluster 13



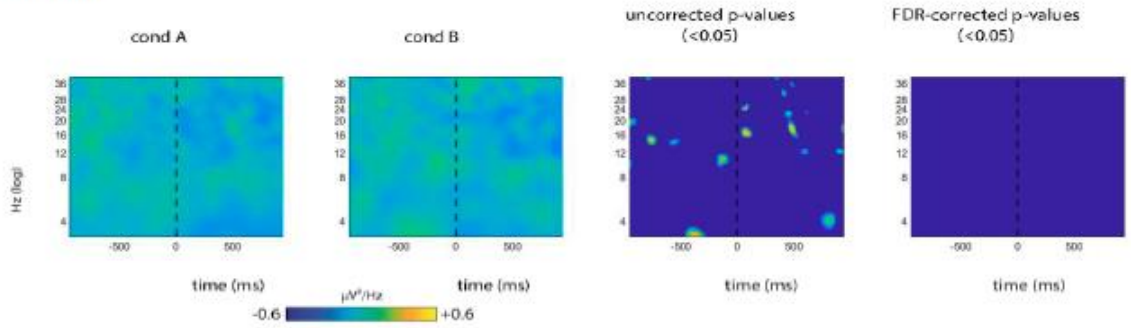
Cluster 14



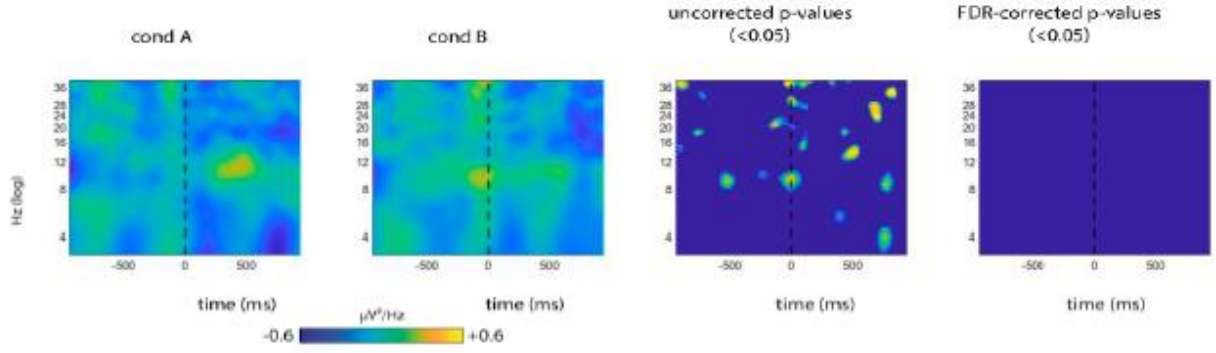
Cluster 15



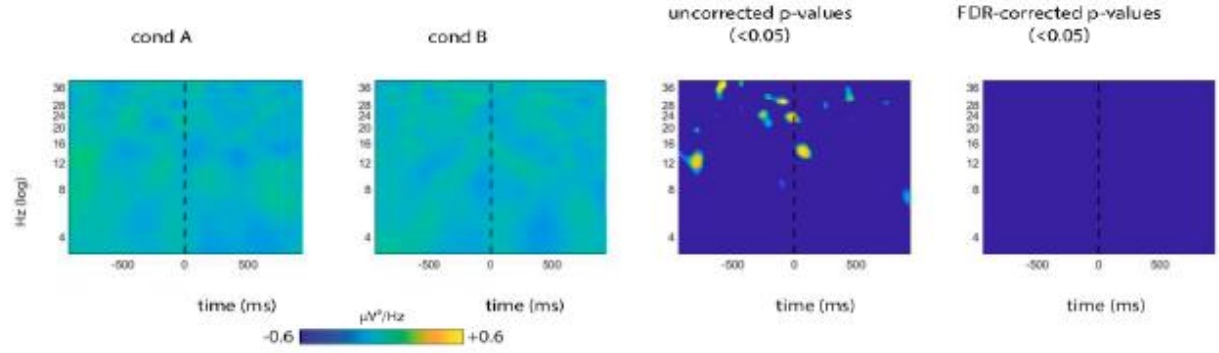
Cluster 16



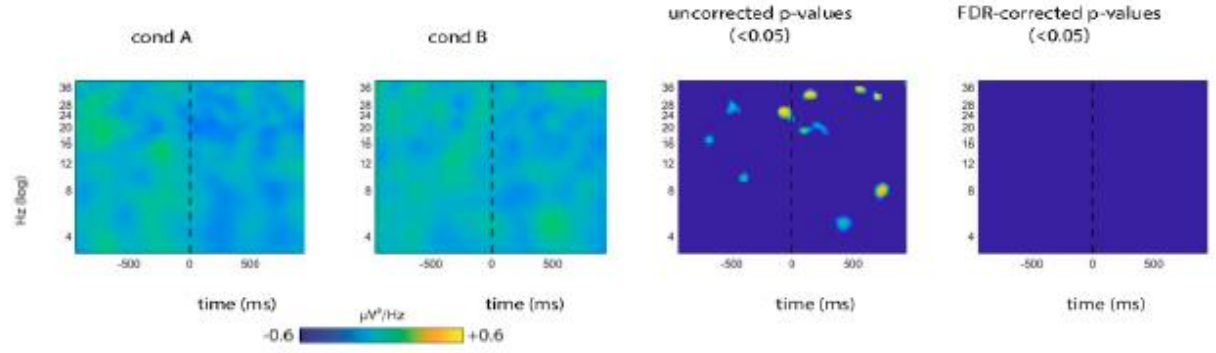
Cluster 17



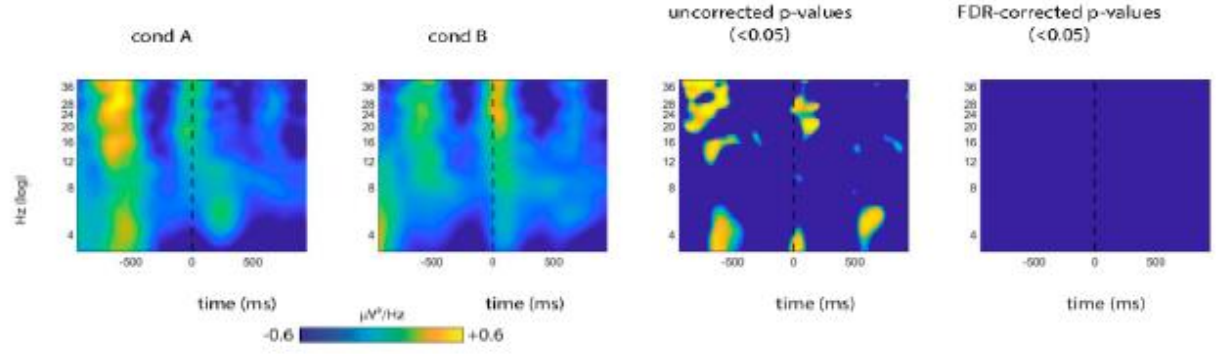
Cluster 18



Cluster 19



Cluster 20



7.2. Appendix B. preprocess_single_subject_v2

```
function preprocess_single_subject_v2(subjectID, dataPath)


%     subjectID = 3;
    dataPath = '../DATA';
    outputPath = '../DATA/PROCESSED/';
    target_SR = 256;
    env_lpf = 8;
    env_hpf = 0.1;
    EpochSize = [-1 20];
    EpochBaseline = [-1000 0];

    fprintf(['Preprocessing subject ' int2str(subjectID) '\n']);

    if ~exist([outputPath 'Subject_' sprintf('%02d',subjectID) '_EEG_continuous_ready.set']) || ~exist([outputPath 'Subject_'
sprintf('%02d',subjectID) '_AUDIO_continuous_ready.set'])

       fprintf(['Stage 1 \n']);

       %% 1-load audio, compute envelope, downsample, synchronize with EEG, epoch data
    thisAUDIO = pop_biosig([dataPath '/AUDIO/Subject_' sprintf('%02d',subjectID) '_audio.bdf']);
    thisAUDIO = eeg_checkset( thisAUDIO );


    x = double(thisAUDIO.data(end,:));
    x = x-mean(x);
    Yenv = abs(hilbert(x));

    [bl, al] = butter(4,env_lpf/(thisAUDIO.srate/2),'low');
    Yenv_filt = filter(bl,al,Yenv(:,end:-1:1),[],2);
    Yenv_filt = filter(bl,al,Yenv_filt(:,end:-1:1),[],2);

    thisAUDIO.data = [];
    thisAUDIO.data(1,:) = x;
    thisAUDIO.data(2,:) = Yenv;
    thisAUDIO.data(3,:) = Yenv_filt;
    thisAUDIO.nbchan = 3;
    thisAUDIO.chanlocs = [];
    thisAUDIO.chanlocs(1).labels = 'audio';
    thisAUDIO.chanlocs(2).labels = 'env';
    thisAUDIO.chanlocs(3).labels = 'env_filt';


    thisAUDIO = pop_resample(thisAUDIO,target_SR);

    %correct the event type column
    for ii=1:length(thisAUDIO.event)
       thisAUDIO.event(ii).type = int2str(thisAUDIO.event(ii).type);
    end

    %remove events until event 250
    event250 = 0;
    while ~event250
       if isequal(thisAUDIO.event(1).type , '250')
          event250 = 1;
       else
          event250 = 0;
          thisAUDIO.event(1) = [];
       end
    end

    % reference events to event 250
    refPointlatencyAUDIO = thisAUDIO.event(1).latency;
    thisAUDIO = pop_select(thisAUDIO, 'nopoint', [1 refPointlatencyAUDIO]);


    %% 2-load eeg
    origEEG = pop_biosig([dataPath '/EEG/Subject_' sprintf('%02d',subjectID) '_eeg.bdf']);
    origEEG = eeg_checkset( origEEG );

    %load channel locations
```

```
    origEEG = pop_chanedit(origEEG, 'load',
'D:\Dropbox\DROPBOX_JOAO\FARO\TESE_DAIANA\LIBS\eeglab2021.1\eeglab2021.1\sample_locs/BioSemi128.sfp');

    %resample EEG
    origEEG = pop_resample(origEEG,target_SR);

     % high-pass filter the data at 1 Hz
    origEEG = pop_eegfiltnew( origEEG, 1, []);

    %get orig_chanlocs
    orig_chanlocs = origEEG.chanlocs;
    thisEEG = origEEG;

    thisEEG = clean_rawdata(thisEEG, 5, [0.25 0.75], 0.85, 4, 5, []);
    removedChannels = setdiff({origEEG.chanlocs.labels},{thisEEG.chanlocs.labels});


    % Remove channels
    thisEEG = pop_select(origEEG, 'nochannel', removedChannels);

    % re-reference EEG to average
    thisEEG = pop_reref( thisEEG, []);
    thisEEG = eeg_checkset( thisEEG );


    % run and clean ICA
    thisEEG = pop_runica(thisEEG, 'icatype', 'picard', 'pca',thisEEG.nbchan - 1);
    thisEEG = eeg_checkset( thisEEG );
    thisEEG = pop_iclabel(thisEEG,'default');
    thisEEG = pop_icflag(thisEEG, [NaN NaN;0.8 1;0.8 1;NaN NaN;0.8 1;0.8 1;NaN NaN]);
    thisEEG = pop_saveset(thisEEG, 'filename', ['Subject_' sprintf('%02d',subjectID) '_EEG_beforeICA.set'], 'filepath',
outputPath);
    thisEEG = pop_subcomp( thisEEG, find(thisEEG.reject.gcompreject), 0);
    thisEEG = eeg_checkset( thisEEG );

    %interpolate channels
    thisEEG = pop_interp(thisEEG, orig_chanlocs, 'spherical');



    % filter the data
    thisEEG = pop_eegfiltnew( thisEEG, [], 70);
    thisEEG = pop_eegfiltnew(thisEEG, 48, 52, [], 1);



    % correct the event type column
    for ii=1:length(thisEEG.event)
       if ~isempty(thisEEG.event(ii).edftype)
          thisEEG.event(ii).type = int2str(thisEEG.event(ii).edftype);
       end
    end

    % remove events until event 250
    event250 = 0;
    while ~event250
       if isequal(thisEEG.event(1).type , '250')
         event250 = 1;
       else
          event250 = 0;
          thisEEG.event(1) = [];
       end
    end

    % reference events to event 250
    refPointlatencyEEG = thisEEG.event(1).latency;
    thisEEG = pop_select(thisEEG, 'nopoint', [1 refPointlatencyEEG]);


    % remove event 100 related to button presses
    events2remove = [];
    for ii=1:length(thisEEG.event)
       if isequal(thisEEG.event(ii).type, '100')
          events2remove = [events2remove ii];
       end
```

```
    end
    thisEEG.event(events2remove) = [];

    % change event type codes
    blockCode = [2 1 1 2 2 1 1 2]; trials_per_block = 40;
    blockIdx = 1;
    countTrials = 1;

    for ii=1:length(thisEEG.event)
        if isequal(thisEEG.event(ii).type, '61')
            if ~rem(countTrials,trials_per_block)
                blockIdx = blockIdx + 1;
            end
            wordEvent = str2num(thisEEG.event(ii-1).type);
            wordEvent = ((blockCode(blockIdx)-1)*1000)+(wordEvent*10);
            thisEEG.event(ii-1).type = int2str(wordEvent);
            thisEEG.event(ii).type = int2str(wordEvent + 1);
            thisEEG.event(ii+1).type = int2str(wordEvent + 2);
            thisEEG.event(ii+2).type = int2str(wordEvent + 3);
            thisEEG.event(ii+3).type = int2str(wordEvent + 4);
            thisEEG.event(ii+4).type = int2str(wordEvent + 5);
            thisEEG.event(ii+5).type = int2str(wordEvent + 6);

            countTrials=countTrials+1;
        end
    end


    %transfer the EEG events to the AUDIO file
    thisAUDIO.event = thisEEG.event;

    thisEEG = pop_saveset(thisEEG, 'filename', ['Subject_' sprintf('%02d',subjectID) '_EEG_continuous_ready.set'], 'filepath',
outputPath);
    thisAUDIO = pop_saveset(thisAUDIO, 'filename', ['Subject_' sprintf('%02d',subjectID) '_AUDIO_continuous_ready.set'],
'filepath', outputPath);

    else

    fprintf(['Stage 2 \n']);
    thisEEG = pop_loadset('filename', ['Subject_' sprintf('%02d',subjectID)  '_EEG_continuous_ready.set'], 'filepath', outputPath);
    thisAUDIO = pop_loadset('filename', ['Subject_' sprintf('%02d',subjectID)  '_AUDIO_continuous_ready.set'], 'filepath',
outputPath);


    %% 4- epoch data A
    thisAUDIO_epochedA = pop_epoch( thisAUDIO, {'1010' '1020' '1030' '1040' '2050' '2060' '2070' '2080'} , EpochSize);
    thisEEG_epochedA = pop_epoch( thisEEG, {'1010' '1020' '1030' '1040' '2050' '2060' '2070' '2080'} , EpochSize);
    thisEEG_epochedA = pop_rmbase(thisEEG_epochedA, EpochBaseline);
    all_events2rem = [];
    for ii=1:length(thisAUDIO_epochedA.epoch)
        tmp_events2rem = thisAUDIO_epochedA.epoch(ii).event(7:end);
        all_events2rem = [all_events2rem tmp_events2rem];
    end
    thisAUDIO_epochedA.event(all_events2rem) = [];
    thisEEG_epochedA.event(all_events2rem) = [];

    thisAUDIO_epochedB = pop_epoch( thisAUDIO, {'2010' '2020' '2030' '2040' '1050' '1060' '1070' '1080'} , EpochSize);
    thisEEG_epochedB = pop_epoch( thisEEG, {'2010' '2020' '2030' '2040' '1050' '1060' '1070' '1080'} , EpochSize);
    thisEEG_epochedB = pop_rmbase(thisEEG_epochedB, EpochBaseline);
    all_events2rem = [];
    for ii=1:length(thisAUDIO_epochedB.epoch)
        tmp_events2rem = thisAUDIO_epochedB.epoch(ii).event(7:end);
        all_events2rem = [all_events2rem tmp_events2rem];
    end
    thisAUDIO_epochedB.event(all_events2rem) = [];
    thisEEG_epochedB.event(all_events2rem) = [];

    clear thisEEG thisAUDIO

%     thisAUDIO_epochedA_baseline = pop_select(thisAUDIO_epochedA, 'time', [EpochSize(1) 0]);
%     thisEEG_epochedA_baseline = pop_select(thisEEG_epochedA, 'time', [EpochSize(1) 0]);
% %     [powbase1, freqs, times] = timefreq( thisEEG_epochedA.data(1:64,:,:), thisEEG_epochedA.srate, 'cycles', [3 0.5], ...
%         'nfreqs', 38, 'freqs', [3 40], 'tlimits', [thisEEG_epochedA.xmin 0]);
%     [powbase2, freqs, times] = timefreq( thisEEG_epochedA.data(65:128,:,:), thisEEG_epochedA.srate, 'cycles', [3 0.5], ...
%         'nfreqs', 38, 'freqs', [3 40], 'tlimits', [thisEEG_epochedA.xmin 0]);
```

```
%     powbase = cat(1,powbase1,powbase2);
%     clear powbase1 powbase2
%     powbase = abs(powbase).^2;
%     thisEEG_epochedA_powbase = squeeze(mean(powbase,3));
%
%     EEGbaselineA = squeeze(mean(thisEEG_epochedA_baseline.data,2));

    [thisAUDIO_epochedA_tmp indices] = pop_epoch( thisAUDIO_epochedA, {'1011' '1021' '1031' '1041' '2051' '2061' '2071'
'2081'} , [-1 5]);
    thisEEG_epochedA_tmp = pop_epoch( thisEEG_epochedA, {'1011' '1021' '1031' '1041' '2051' '2061' '2071' '2081'} , [-1 5]);


%     thisAUDIO_epochedA_baseline = pop_select(thisAUDIO_epochedA_baseline, 'trial', indices);
%     thisEEG_epochedA_baseline = pop_select(thisEEG_epochedA_baseline, 'trial', indices);
%
%     thisAUDIO_epochedA_baseline = pop_saveset(thisAUDIO_epochedA_baseline, 'filename', ['BASELINE/Subject_'
sprintf('%02d',subjectID) '_AUDIO_condA_baseline.set'], 'filepath', outputPath);
%     thisEEG_epochedA_baseline = pop_saveset(thisEEG_epochedA_baseline, 'filename', ['BASELINE/Subject_'
sprintf('%02d',subjectID) '_EEG_condA_baseline.set'], 'filepath', outputPath);




    % 6-find two peaks in speech envelope (trial-by-trial) and create a list of events with it.
    for epIdx=1:thisAUDIO_epochedA_tmp.trials
        [pks locs] =
findpeaks(thisAUDIO_epochedA_tmp.data(3,:,epIdx),"NPeaks",6,"SortStr","descend",'MinPeakDistance',0.5*thisAUDIO_epoch
edA_tmp.srate);
        [tmp pks_order] = sort(locs);

        for pkIdx=1:6
            peak_loc = locs(pks_order(pkIdx));
            thisEEG_epochedA_tmp.event(end+1).type = ['peak' int2str(pkIdx)];
            thisEEG_epochedA_tmp.event(end).latency = peak_loc+((epIdx-1)*thisEEG_epochedA_tmp.pnts);
            thisEEG_epochedA_tmp.event(end).epoch = epIdx;
        end
    end
    thisAUDIO_epochedA_tmp.event = thisEEG_epochedA_tmp.event;

    peak_epoch_interval_peak2 = [-1.5 1.5];
    peak_epoch_interval_peak3 = [-1.5 1.5];
    peak_epoch_interval_peak4 = [-1.5 1.5];


    % epoch condition A in relation to peak2 and peak3
    thisAUDIO_epochedA_peak2 = pop_epoch( thisAUDIO_epochedA_tmp, {'peak2'} , peak_epoch_interval_peak2);
    [thisEEG_epochedA_peak2 ind_peak2] = pop_epoch( thisEEG_epochedA_tmp, {'peak2'} , peak_epoch_interval_peak2);
    thisAUDIO_epochedA_peak3 = pop_epoch( thisAUDIO_epochedA_tmp, {'peak3'} , peak_epoch_interval_peak3);
    [thisEEG_epochedA_peak3 ind_peak3] = pop_epoch( thisEEG_epochedA_tmp, {'peak3'} , peak_epoch_interval_peak3);
    thisAUDIO_epochedA_peak4 = pop_epoch( thisAUDIO_epochedA_tmp, {'peak4'} , peak_epoch_interval_peak4);
    [thisEEG_epochedA_peak4 ind_peak4] = pop_epoch( thisEEG_epochedA_tmp, {'peak4'} , peak_epoch_interval_peak4);


    clear thisAUDIO_epochedA_tmp thisEEG_epochedA_tmp

    %compute ERSP for each peak and normalize by the powbase previously
%     %computed
%     [ersp_A_peak2_1, freqs, ersp_times] = timefreq( thisEEG_epochedA_peak2.data(1:64,:,:),
thisEEG_epochedA_peak2.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     [ersp_A_peak2_2, freqs, ersp_times] = timefreq( thisEEG_epochedA_peak2.data(65:128,:,:),
thisEEG_epochedA_peak2.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     ersp_A_peak2 = cat(1,ersp_A_peak2_1,ersp_A_peak2_2);
%     clear ersp_A_peak2_1 ersp_A_peak2_2
%     ersp_A_peak2 = abs(ersp_A_peak2).^2;
%     ersp_A_peak2 = permute(ersp_A_peak2,[1 2 4 3]);
%     ersp_A_peak2_psc = (ersp_A_peak2-
thisEEG_epochedA_powbase(:,:,ind_peak2))./thisEEG_epochedA_powbase(:,:,ind_peak2);
%     ersp_A_peak2_psc = permute(ersp_A_peak2_psc,[1 2 4 3]);
%     ersp_A_peak2_psc = mean(ersp_A_peak2_psc,4);
%     data_times = thisEEG_epochedA_peak2.times;
%     data = mean(thisEEG_epochedA_peak2.data,3);
%     ersp = ersp_A_peak2_psc; save([outputPath '/peak2_condA/Subject_' sprintf('%02d',subjectID) '.mat'], 'ersp', 'data_times',
'ersp_times', 'freqs', 'data');
```

```
%     clear ersp_A_peak2 ersp ersp_A_peak2_psc data
%
%     [ersp_A_peak3_1, freqs, ersp_times] = timefreq( thisEEG_epochedA_peak3.data(1:64,:,:),
thisEEG_epochedA_peak3.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     [ersp_A_peak3_2, freqs, ersp_times] = timefreq( thisEEG_epochedA_peak3.data(65:128,:,:),
thisEEG_epochedA_peak3.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     ersp_A_peak3 = cat(1,ersp_A_peak3_1,ersp_A_peak3_2);
%     clear ersp_A_peak3_1 ersp_A_peak3_2
%     ersp_A_peak3 = abs(ersp_A_peak3).^2;
%     ersp_A_peak3 = permute(ersp_A_peak3,[1 2 4 3]);
%     ersp_A_peak3_psc = (ersp_A_peak3-
thisEEG_epochedA_powbase(:,:,ind_peak3))./thisEEG_epochedA_powbase(:,:,ind_peak3);
%     ersp_A_peak3_psc = permute(ersp_A_peak3_psc,[1 2 4 3]);
%     ersp_A_peak3_psc = mean(ersp_A_peak3_psc,4);
%     data_times = thisEEG_epochedA_peak3.times;
%     data = mean(thisEEG_epochedA_peak3.data,3); ersp = ersp_A_peak3_psc; save([outputPath '/peak3_condA/Subject_'
sprintf('%02d',subjectID) '.mat'], 'ersp', 'data_times', 'ersp_times', 'freqs', 'data');
%     clear ersp_A_peak3 ersp ersp_A_peak3_psc data
%
%     [ersp_A_peak4_1, freqs, ersp_times] = timefreq( thisEEG_epochedA_peak4.data(1:64,:,:),
thisEEG_epochedA_peak4.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     [ersp_A_peak4_2, freqs, ersp_times] = timefreq( thisEEG_epochedA_peak4.data(65:128,:,:),
thisEEG_epochedA_peak4.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     ersp_A_peak4 = cat(1,ersp_A_peak4_1,ersp_A_peak4_2);
%     clear ersp_A_peak4_1 ersp_A_peak4_2
%     ersp_A_peak4 = abs(ersp_A_peak4).^2;
%     ersp_A_peak4 = permute(ersp_A_peak4,[1 2 4 3]);
%     ersp_A_peak4_psc = (ersp_A_peak4-
thisEEG_epochedA_powbase(:,:,ind_peak4))./thisEEG_epochedA_powbase(:,:,ind_peak4);
%     ersp_A_peak4_psc = permute(ersp_A_peak4_psc,[1 2 4 3]);
%     ersp_A_peak4_psc = mean(ersp_A_peak4_psc,4);
%     data_times = thisEEG_epochedA_peak4.times;
%     data = mean(thisEEG_epochedA_peak4.data,3); ersp = ersp_A_peak4_psc; save([outputPath '/peak4_condA/Subject_'
sprintf('%02d',subjectID) '.mat'], 'ersp', 'data_times', 'ersp_times', 'freqs', 'data');
%     clear ersp_A_peak4 ersp ersp_A_peak4_psc data


    % save datasets
    thisEEG_epochedA_peak2 = pop_saveset(thisEEG_epochedA_peak2, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_EEG_condA_peak2.set'], 'filepath', outputPath);
    thisAUDIO_epochedA_peak2 = pop_saveset(thisAUDIO_epochedA_peak2, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_AUDIO_condA_peak2.set'], 'filepath', outputPath);
    thisEEG_epochedA_peak3 = pop_saveset(thisEEG_epochedA_peak3, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_EEG_condA_peak3.set'], 'filepath', outputPath);
    thisAUDIO_epochedA_peak3 = pop_saveset(thisAUDIO_epochedA_peak3, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_AUDIO_condA_peak3.set'], 'filepath', outputPath);
    thisEEG_epochedA_peak4 = pop_saveset(thisEEG_epochedA_peak4, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_EEG_condA_peak4.set'], 'filepath', outputPath);
    thisAUDIO_epochedA_peak4 = pop_saveset(thisAUDIO_epochedA_peak4, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_AUDIO_condA_peak4.set'], 'filepath', outputPath);



    %% 4- epoch data B
%     thisAUDIO_epochedB_baseline = pop_select(thisAUDIO_epochedB, 'time', [EpochSize(1) 0]);
%     thisEEG_epochedB_baseline = pop_select(thisEEG_epochedB, 'time', [EpochSize(1) 0]);
%     [powbase1, freqs, times] = timefreq( thisEEG_epochedB.data(1:64,:,:), thisEEG_epochedB.srate, 'cycles', [3 0.5], ...
%         'nfreqs', 38, 'freqs', [3 40], 'tlimits', [thisEEG_epochedB.xmin 0]);
%     [powbase2, freqs, times] = timefreq( thisEEG_epochedB.data(65:128,:,:), thisEEG_epochedB.srate, 'cycles', [3 0.5], ...
%         'nfreqs', 38, 'freqs', [3 40], 'tlimits', [thisEEG_epochedB.xmin 0]);
%     powbase = cat(1,powbase1,powbase2);
%     clear powbase1 powbase2
%     powbase = abs(powbase).^2;
%     thisEEG_epochedB_powbase = squeeze(mean(powbase,3));

    [thisAUDIO_epochedB_tmp indices] = pop_epoch( thisAUDIO_epochedB, {'2011' '2021' '2031' '2041' '1051' '1061' '1071'
'1081'} , [-1 5]);
    thisEEG_epochedB_tmp = pop_epoch( thisEEG_epochedB, {'2011' '2021' '2031' '2041' '1051' '1061' '1071' '1081'} , [-1 5]);


%     thisAUDIO_epochedB_baseline = pop_select(thisAUDIO_epochedB_baseline, 'trial', indices);
```

```
%     thisEEG_epochedB_baseline = pop_select(thisEEG_epochedB_baseline, 'trial', indices);
%
%     thisAUDIO_epochedB_baseline = pop_saveset(thisAUDIO_epochedB_tmp, 'filename', ['BASELINE/Subject_'
sprintf('%02d',subjectID) '_AUDIO_condB_baseline.set'], 'filepath', outputPath);
%     thisEEG_epochedB_baseline = pop_saveset(thisEEG_epochedB_tmp, 'filename', ['BASELINE/Subject_'
sprintf('%02d',subjectID) '_EEG_condB_baseline.set'], 'filepath', outputPath);


    % 6-find two peaks in speech envelope (trial-by-trial) and create a list of events with it.
    for epIdx=1:thisAUDIO_epochedB_tmp.trials
        [pks locs] =
findpeaks(thisAUDIO_epochedB_tmp.data(3,:,epIdx),"NPeaks",6,"SortStr","descend",'MinPeakDistance',0.5*thisAUDIO_epoch
edB_tmp.srate);
        [tmp pks_order] = sort(locs);

        for pkIdx=1:6
            peak_loc = locs(pks_order(pkIdx));
            thisEEG_epochedB_tmp.event(end+1).type = ['peak' int2str(pkIdx)];
            thisEEG_epochedB_tmp.event(end).latency = peak_loc+((epIdx-1)*thisEEG_epochedB_tmp.pnts);
            thisEEG_epochedB_tmp.event(end).epoch = epIdx;
        end
    end
    thisAUDIO_epochedB_tmp.event = thisEEG_epochedB_tmp.event;


%     epoch condition B in relation to peak2 and peak3
    thisAUDIO_epochedB_peak2 = pop_epoch( thisAUDIO_epochedB_tmp, {'peak2'} , peak_epoch_interval_peak2);
    [thisEEG_epochedB_peak2 ind_peak2] = pop_epoch( thisEEG_epochedB_tmp, {'peak2'} , peak_epoch_interval_peak2);
    thisAUDIO_epochedB_peak3 = pop_epoch( thisAUDIO_epochedB_tmp, {'peak3'} , peak_epoch_interval_peak3);
    [thisEEG_epochedB_peak3 ind_peak3] = pop_epoch( thisEEG_epochedB_tmp, {'peak3'} , peak_epoch_interval_peak3);
    thisAUDIO_epochedB_peak4 = pop_epoch( thisAUDIO_epochedB_tmp, {'peak4'} , peak_epoch_interval_peak4);
    [thisEEG_epochedB_peak4 ind_peak4] = pop_epoch( thisEEG_epochedB_tmp, {'peak4'} , peak_epoch_interval_peak4);


    clear thisAUDIO thisEEG thisAUDIO_epochedB_tmp thisEEG_epochedB_tmp

    %compute ERSP for each peak and normalize by the powbase previously
    %computed
%     [ersp_B_peak2_1, freqs, ersp_times] = timefreq( thisEEG_epochedB_peak2.data(1:64,:,:),
thisEEG_epochedB_peak2.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     [ersp_B_peak2_2, freqs, ersp_times] = timefreq( thisEEG_epochedB_peak2.data(65:128,:,:),
thisEEG_epochedB_peak2.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     ersp_B_peak2 = cat(1,ersp_B_peak2_1,ersp_B_peak2_2);
%     clear ersp_B_peak2_1 ersp_B_peak2_2
%     ersp_B_peak2 = abs(ersp_B_peak2).^2;
%     ersp_B_peak2 = permute(ersp_B_peak2,[1 2 4 3]);
%     ersp_B_peak2_psc = (ersp_B_peak2-
thisEEG_epochedB_powbase(:,:,ind_peak2))./thisEEG_epochedB_powbase(:,:,ind_peak2);
%     ersp_B_peak2_psc = permute(ersp_B_peak2_psc,[1 2 4 3]);
%     ersp_B_peak2_psc = mean(ersp_B_peak2_psc,4);
%     data_times = thisEEG_epochedB_peak2.times;
%     data = mean(thisEEG_epochedB_peak2.data,3); ersp = ersp_B_peak2_psc; save([outputPath '/peak2_condB/Subject_'
sprintf('%02d',subjectID) '.mat'], 'ersp', 'data_times', 'ersp_times', 'freqs', 'data');
%     clear ersp_B_peak2 ersp ersp_B_peak2_psc
%
%     [ersp_B_peak3_1, freqs, ersp_times] = timefreq( thisEEG_epochedB_peak3.data(1:64,:,:),
thisEEG_epochedB_peak2.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     [ersp_B_peak3_2, freqs, ersp_times] = timefreq( thisEEG_epochedB_peak3.data(65:128,:,:),
thisEEG_epochedB_peak2.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     ersp_B_peak3 = cat(1,ersp_B_peak3_1,ersp_B_peak3_2);
%     clear ersp_B_peak3_1 ersp_B_peak3_2
%     ersp_B_peak3 = abs(ersp_B_peak3).^2;
%     ersp_B_peak3 = permute(ersp_B_peak3,[1 2 4 3]);
%     ersp_B_peak3_psc = (ersp_B_peak3-
thisEEG_epochedB_powbase(:,:,ind_peak3))./thisEEG_epochedB_powbase(:,:,ind_peak3);
%     ersp_B_peak3_psc = permute(ersp_B_peak3_psc,[1 2 4 3]);
%     ersp_B_peak3_psc = mean(ersp_B_peak3_psc,4);
%     data_times = thisEEG_epochedB_peak3.times;
%     data = mean(thisEEG_epochedB_peak3.data,3); ersp = ersp_B_peak3_psc; save([outputPath '/peak3_condB/Subject_'
sprintf('%02d',subjectID) '.mat'], 'ersp', 'data_times', 'ersp_times', 'freqs', 'data');
%     clear ersp_B_peak3 ersp ersp_B_peak3_psc
```

```
%
%     [ersp_B_peak4_1, freqs, ersp_times] = timefreq( thisEEG_epochedB_peak4.data(1:64,:,:),
thisEEG_epochedB_peak2.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     [ersp_B_peak4_2, freqs, ersp_times] = timefreq( thisEEG_epochedB_peak4.data(65:128,:,:),
thisEEG_epochedB_peak2.srate, 'cycles', [3 0.5], ...
%     'nfreqs', 38, 'freqs', [3 40], 'tlimits', [-1 1]);
%     ersp_B_peak4 = cat(1,ersp_B_peak4_1,ersp_B_peak4_2);
%     clear ersp_B_peak4_1 ersp_B_peak4_2
%     ersp_B_peak4 = abs(ersp_B_peak4).^2;
%     ersp_B_peak4 = permute(ersp_B_peak4,[1 2 4 3]);
%     ersp_B_peak4_psc = (ersp_B_peak4-
thisEEG_epochedB_powbase(:,:,ind_peak4))./thisEEG_epochedB_powbase(:,:,ind_peak4);
%     ersp_B_peak4_psc = permute(ersp_B_peak4_psc,[1 2 4 3]);
%     ersp_B_peak4_psc = mean(ersp_B_peak4_psc,4);
%     data_times = thisEEG_epochedB_peak4.times;
%     data = mean(thisEEG_epochedB_peak4.data,3); ersp = ersp_B_peak4_psc; save([outputPath '/peak4_condB/Subject_'
sprintf('%02d',subjectID) '.mat'], 'ersp', 'data_times', 'ersp_times', 'freqs', 'data');
%     clear ersp_B_peak4 ersp ersp_B_peak4_psc

    % save datasets
    thisEEG_epochedB_peak2 = pop_saveset(thisEEG_epochedB_peak2, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_EEG_condB_peak2.set'], 'filepath', outputPath);
    thisAUDIO_epochedB_peak2 = pop_saveset(thisAUDIO_epochedB_peak2, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_AUDIO_condB_peak2.set'], 'filepath', outputPath);
    thisEEG_epochedB_peak3 = pop_saveset(thisEEG_epochedB_peak3, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_EEG_condB_peak3.set'], 'filepath', outputPath);
    thisAUDIO_epochedB_peak3 = pop_saveset(thisAUDIO_epochedB_peak3, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_AUDIO_condB_peak3.set'], 'filepath', outputPath);
    thisEEG_epochedB_peak4 = pop_saveset(thisEEG_epochedB_peak4, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_EEG_condB_peak4.set'], 'filepath', outputPath);
    thisAUDIO_epochedB_peak4 = pop_saveset(thisAUDIO_epochedB_peak4, 'filename', ['Subject_' sprintf('%02d',subjectID)
'_AUDIO_condB_peak4.set'], 'filepath', outputPath);


    end
```

7.3. Appendix C. makeClusters_andPlotResults2

```
% Lock randomness
rng(0);
% rng('shuffle');



[STUDY ALLEEG] = pop_loadstudy('filename', 'SOPHIE_PEAK3.study', 'filepath',
'D:\Dropbox\DROPBOX_JOAO\FARO\SOPHIE\DATA\PROCESSED\PEAK3');


[STUDY ALLEEG] = std_preclust(STUDY, ALLEEG, 1,{'spec','npca',3,'weight',1,'freqrange',[3 40] },{'dipoles','weight',1});
% [STUDY ALLEEG] = std_preclust(STUDY, ALLEEG, 1,{'dipoles','weight',1});
% [STUDY ALLEEG] = std_preclust(STUDY, ALLEEG, 1,{'scalp', 'weight',1},{'spec','npca',3,'weight',1,'freqrange',[3 40]
},{'dipoles','weight',1});


[STUDY, ALLEEG] = pop_clust(STUDY, ALLEEG, 'algorithm','kmeans','clus_num', 20);


STUDY =  std_dipplot(STUDY, ALLEEG, 'verbose', 'off', 'figure', 'off',  'mode', 'together');
saveas(gcf,'../FIGURES/ClustersDipoles.png');


[STUDY, ALLEEG] = pop_savestudy( STUDY, ALLEEG, 'filename', 'study_peak3.study', 'filepath',
'D:\Dropbox\DROPBOX_JOAO\FARO\SOPHIE\DATA\PROCESSED\PEAK3');


clustersPEAK3 = STUDY.cluster;
```

```
changrpPEAK3    = STUDY.changrp;
etcPEAK3        = STUDY.etc;



STUDY.etc.erspparams.subbaseline = 'on';
STUDY.etc.erspparams.commonbaseline = 'on';


%%


for cc=[2:21]
%
   [STUDY erspdata ersptimes erspfreqs pgroup pcond pinter] = std_erspplot(STUDY,ALLEEG, 'clusters', cc, 'freqrange', [3 40],
...             % specify your frequency band (e.g., alpha band)
   'statistics', 'perm', 'naccu', 2000,...             % 'param' or 'perm'
   'alpha', 0.05, ...                  % significance threshold
   'groupstats', 'off', ...
   'condstats', 'on',...
   'mcorrect', 'fdr', 'recompute', 'on');                % cluster-based correct);


% [STUDY erspdata ersptimes erspfreqs pgroup pcond pinter] = std_erspplot(STUDY,ALLEEG, 'clusters', cc, 'freqrange', [3 40],
...             % specify your frequency band (e.g., alpha band)
%    'statistics', 'param', ...             % 'param' or 'perm'
%    'alpha', 0.05, ...                  % significance threshold
%    'groupstats', 'off', ...
%    'condstats', 'on',...
%    'mcorrect', 'fdr', 'recompute', 'on');


   figure('Position', [100, 100, 1400, 200])

   subplot(141);
   [T,F] = meshgrid(ersptimes, erspfreqs);
   pcolor(T, F, mean(erspdata{1,1},3));
   shading flat;      % removes grid lines
   axis tight;       % trim whitespace
   caxis([-0.6 0.6]);
   xline(0, '--k', 'LineWidth', 1.5);   % black dashed line at time=0
   set(gca,'yscale','log');
   yticks([4 8 12 16 20 24 28 36]);
   yticklabels(string([4 8 12 16 20 24 28 36]));

   subplot(142);
   [T,F] = meshgrid(ersptimes, erspfreqs);
   pcolor(T, F, mean(erspdata{2,1},3));
   shading flat;      % removes grid lines
   axis tight;       % trim whitespace
   caxis([-0.6 0.6]);
   xline(0, '--k', 'LineWidth', 1.5);   % black dashed line at time=0
   set(gca,'yscale','log');
   yticks([4 8 12 16 20 24 28 36]);
   yticklabels(string([4 8 12 16 20 24 28 36]));

   subplot(144);
   [T,F] = meshgrid(ersptimes, erspfreqs);
   pcolor(T, F, pcond{1,1});
   shading flat;      % removes grid lines
   axis tight;       % trim whitespace
   caxis([0 1]);
   xline(0, '--k', 'LineWidth', 1.5);   % black dashed line at time=0
   set(gca,'yscale','log');
   yticks([4 8 12 16 20 24 28 36]);
   yticklabels(string([4 8 12 16 20 24 28 36]));


   subplot(143);
   tmpA = permute(erspdata{1,1},[3 1 2]);
   tmpB = permute(erspdata{2,1},[3 1 2]);
   [h p] = ttest(tmpA,tmpB);
   [T,F] = meshgrid(ersptimes, erspfreqs);
   pcolor(T, F, squeeze(1-p));
```

```
    shading flat;      % removes grid lines
    axis tight;        % trim whitespace
    caxis([0.95 1]);
    xline(0, '--k', 'LineWidth', 1.5);   % black dashed line at time=0
    set(gca,'yscale','log');
    yticks([4 8 12 16 20 24 28 36]);
    yticklabels(string([4 8 12 16 20 24 28 36]));


    saveas(gcf,['../FIGURES/Cluster_' int2str(cc) 'ERSP_PEAK3_stat.png']);

end









%% plot peak times across subjects


sz = 10;


AVG_PEAK2_latency_A = [];
AVG_PEAK4_latency_A = [];
AVG_PEAK2_latency_B = [];
AVG_PEAK4_latency_B = [];


figure('color', 'white')


subplot(1,2,1)


tmp=1;


for datasetIdx=1:2:38


PEAK2_latency = [];
PEAK4_latency = [];




for evIdx=1:length(ALLEEG(datasetIdx).event)

    if strcmp(ALLEEG(datasetIdx).event(evIdx).type , 'peak3')
        if evIdx>2
            if strcmp(ALLEEG(datasetIdx).event(evIdx-2).type , 'peak2')
                PEAK2_latency = [PEAK2_latency ALLEEG(datasetIdx).event(evIdx).latency-ALLEEG(datasetIdx).event(evIdx-
2).latency];
            end
        end

        if evIdx+2<=length(ALLEEG(datasetIdx).event)
            if strcmp(ALLEEG(datasetIdx).event(evIdx+2).type , 'peak4')
                PEAK4_latency = [PEAK4_latency ALLEEG(datasetIdx).event(evIdx+2).latency-
ALLEEG(datasetIdx).event(evIdx).latency];
            end
        end
    end

end


PEAK2_latency = PEAK2_latency/256*1000;
PEAK4_latency = PEAK4_latency/256*1000;
```

```
scatter(-PEAK2_latency,tmp*ones(1,length(PEAK2_latency)), sz, 'filled','MarkerFaceColor', [1 0 0]); hold on;
scatter(PEAK4_latency,tmp*ones(1,length(PEAK4_latency)), sz, 'filled','MarkerFaceColor', [0 1 0]); hold on;


scatter(-mean(PEAK2_latency), tmp, sz, 'filled','MarkerFaceColor', [0 0 0]); hold on;
scatter(mean(PEAK4_latency), tmp, sz, 'filled','MarkerFaceColor', [0 0 0]); hold on;



AVG_PEAK2_latency_A = [AVG_PEAK2_latency_A mean(PEAK2_latency)];
AVG_PEAK4_latency_A = [AVG_PEAK4_latency_A mean(PEAK4_latency)];


tmp=tmp+1;
end


plot([-mean(AVG_PEAK2_latency_A) -mean(AVG_PEAK2_latency_A)], [0 20], 'linewidth', 1, 'color', [0.3 0.3 0.3]); hold on;
plot([mean(AVG_PEAK4_latency_A) mean(AVG_PEAK4_latency_A)], [0 20], 'linewidth', 1, 'color', [0.3 0.3 0.3]); hold off;




subplot(1,2,2)


tmp=1;


for datasetIdx=2:2:38


PEAK2_latency = [];
PEAK4_latency = [];



for evIdx=1:length(ALLEEG(datasetIdx).event)

   if strcmp(ALLEEG(datasetIdx).event(evIdx).type , 'peak3')
      if evIdx>2
         if strcmp(ALLEEG(datasetIdx).event(evIdx-2).type , 'peak2')
            PEAK2_latency = [PEAK2_latency ALLEEG(datasetIdx).event(evIdx).latency-ALLEEG(datasetIdx).event(evIdx-
2).latency];
         end
      end

      if evIdx+2<=length(ALLEEG(datasetIdx).event)
         if strcmp(ALLEEG(datasetIdx).event(evIdx+2).type , 'peak4')
            PEAK4_latency = [PEAK4_latency ALLEEG(datasetIdx).event(evIdx+2).latency-
ALLEEG(datasetIdx).event(evIdx).latency];
         end
      end
   end

end


PEAK2_latency = PEAK2_latency/256*1000;
PEAK4_latency = PEAK4_latency/256*1000;


scatter(-PEAK2_latency,tmp*ones(1,length(PEAK2_latency)), sz, 'filled','MarkerFaceColor', [1 0 0]); hold on;
scatter(PEAK4_latency,tmp*ones(1,length(PEAK4_latency)), sz, 'filled','MarkerFaceColor', [0 1 0]); hold on;


scatter(-mean(PEAK2_latency), tmp, sz, 'filled','MarkerFaceColor', [0 0 0]); hold on;
scatter(mean(PEAK4_latency), tmp, sz, 'filled','MarkerFaceColor', [0 0 0]); hold on;
```

```
AVG_PEAK2_latency_B = [AVG_PEAK2_latency_B mean(PEAK2_latency)];
AVG_PEAK4_latency_B = [AVG_PEAK4_latency_B mean(PEAK4_latency)];


tmp=tmp+1;


end


plot([-mean(AVG_PEAK2_latency_B) -mean(AVG_PEAK2_latency_B)], [0 20], 'linewidth', 1, 'color', [0.3 0.3 0.3]); hold on;
plot([mean(AVG_PEAK4_latency_B) mean(AVG_PEAK4_latency_B)], [0 20], 'linewidth', 1, 'color', [0.3 0.3 0.3]); hold off;



[h_latency_PEAK2 p_latency_PEAK2 ci_latency_PEAK2 stats_latency_PEAK2] =
ttest(AVG_PEAK2_latency_A,AVG_PEAK2_latency_B);
[h_latency_PEAK4 p_latency_PEAK4 ci_latency_PEAK4 stats_latency_PEAK4] =
ttest(AVG_PEAK4_latency_A,AVG_PEAK4_latency_B);


saveas(gcf,'../FIGURES/SyllableTimes.eps');
```

7.4. Appendix D. generateIC_dipoles

```
subs = [4 5 6 7 9 10 11 12 13 14 15 16 18 19 20 21 22 23 24];


peak = 3;
cond = 'B';


dataPath = ['../DATA/PROCESSED/PEAK' int2str(peak) '/'];


for ii=1:length(subs)

    subjectID = subs(ii);

    EEG = pop_loadset('filename', ['Subject_' sprintf('%02d',subjectID)  '_EEG_cond' cond '_peak' int2str(peak) '.set'], 'filepath',
dataPath);

    EEG = pop_dipfit_settings(EEG, ...

'hdmfile','D:\Dropbox\DROPBOX_JOAO\FARO\LIBS\eeglab_current\eeglab2024.2\plugins\dipfit5.5\dipfit\standard_BEM/standa
rd_vol.mat', ...
    'coordformat','MNI', ...

'mrifile','D:\Dropbox\DROPBOX_JOAO\FARO\LIBS\eeglab_current\eeglab2024.2\plugins\dipfit5.5\dipfit\standard_BEMstandard
_mri.mat', ...

'chanfile','D:\Dropbox\DROPBOX_JOAO\FARO\LIBS\eeglab_current\eeglab2024.2\plugins\dipfit5.5\dipfit\standard_BEMstanda
rd_1005.elc', ...
    'coord_transform',[0 0 0 0 0 0 1 1 1], ...
    'chansel', 1:EEG.nbchan);


%     pop_dipfit_batch(EEG);

    EEG = pop_dipfit_gridsearch(EEG, 1:size(EEG.icaweights, 1), linspace(-85,85,24), linspace(-85,85,24), linspace(0,85,12),
40);

    pop_saveset(EEG, 'filename',['Subject_' sprintf('%02d',subjectID)  '_EEG_cond' cond '_peak' int2str(peak) '.set'], 'filepath',
dataPath);
end
```