

TREAT: a toolbox for trial selection and alignment of unlabeled neural data

S.H. Geukes¹, A. Bekius¹, F. Fabris^{1,2}, Z.V. Freudenburg¹, N.F. Ramsey^{1,2}, M.P. Branco^{1,2}

¹ Department of Neurology and Neurosurgery, University Medical Center Utrecht Brain Center, Utrecht University, Utrecht, The Netherlands

² Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

E-mail: m.pedrosobranco@umcutrecht.nl

ABSTRACT: Brain–computer interfaces (BCIs) enable paralyzed individuals to control external devices using brain signals. For training decoders, trial-to-trial variability in response time may obscure meaningful task-related neural activity. We introduce the TRial sElection and AlignmenT (TREAT) toolbox to provide an automated neural event onset detection method that supports both trial selection and event alignment, based on the slope of the neural response (neural slope marker, NSM). We validated TREAT using high-density electrocorticography recordings from three able-bodied individuals performing a speech task. We extracted high-frequency and alpha band power and compared the NSM to the voice onset times. For two of the three participants, the neural slope markers preceded voice onset by 0.4 seconds. Trial selection improved the precision of neural onset estimates across trials as the spread of NSM timing compared to voice onset time decreased. TREAT offers event-detection in the absence of behavioral markers and may thereby improve interpretation of neural activity and training of BCI decoding models.

INTRODUCTION

Brain–computer interfaces (BCIs) enable individuals with paralysis to control external devices for communication or environmental control, without requiring any muscle output [1]. Variability in task responses due to missed trials and variable reaction times can obscure meaningful task-related neural activity, particularly when signals are averaged across trials [2]. Time-locking neural activity to behavioral markers (obtained for example with a data glove or microphone) typically mitigates this variability [3], [4], [5]. However, behavioral markers are unavailable for individuals who cannot move or speak; individuals who may particularly benefit from a BCI. On top of that, the absence of movement-related sensory feedback in this population can itself introduce variability in task responses, which may be further increased by disease-related factors such as reduced attention, as observed for example in amyotrophic lateral sclerosis (ALS) [6]. These limitations underscore the need for signal-curation methods for trial selection/rejection and response alignment that do not depend on behavioral markers but rely solely on neural data. Others previously presented

trial selection based on both the trial’s contribution to classification accuracy and templates of independent components reflecting expected motor-related power changes in the neural signal [7]. Here we propose to accomplish trial selection by thresholding the change in signal amplitude after the cue relative to a baseline period. In addition, previously proposed alignment methods relied either on the peak [2] or the shape [8] of the neural response. To most accurately capture movement onset and maximize the contrast between movement and rest, we previously proposed alignment based on the increase in neural activity, more specifically the slope of the response [3]. We extend this approach by developing a toolbox that integrates both trial selection and alignment without the need for behavioral markers. Here, we introduce the TRial sElection and AlignmenT (TREAT) toolbox: a user-friendly automated neural event onset detection method for signals acquired with electrocorticography (ECoG) and electroencephalography (EEG). This toolbox facilitates offline analysis of unlabeled datasets, supporting both trial selection and trial alignment based on neural

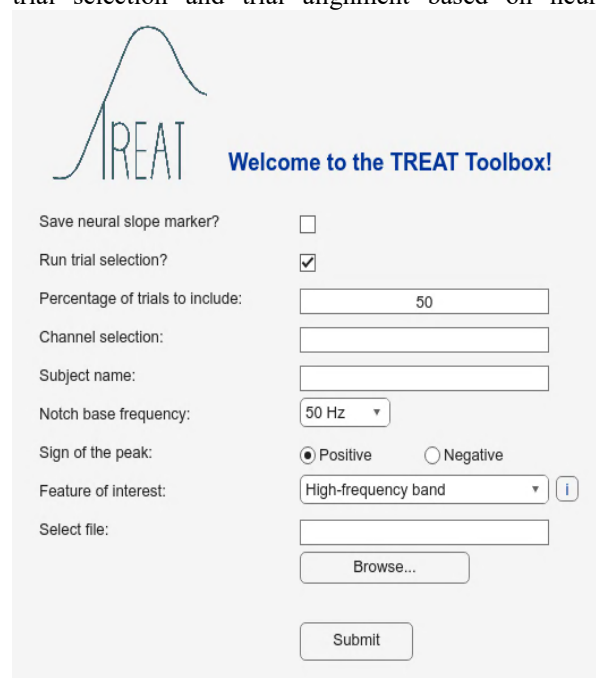


Figure 1: Graphical user interface of the TREAT toolbox.

features. We validated this approach in a high-density ECoG dataset from three able-bodied individuals performing a speech task, by comparing alignment with TREAT to alignment with a behavioral marker: the voice onset time (VOT).

MATERIALS AND METHODS

The TREAT toolbox: The toolbox consists of a graphical user interface developed in MATLAB 2025b (The MathWorks, Natick, USA) (Fig. 1). The interface allows the user to input raw data with .dat (BCI2000 [9]), .nev/.nsx (Blackrock Neurotech, Salt Lake City, USA) or .trc files (Micromed, Treviso, Italy) formats. The toolbox preprocesses the data, segments it into trials and performs trial selection (optional) and trial alignment (see Fig. 2 for an overview of this process).

Preprocessing: The user provides a selection of channels to be analyzed. This selection should exclude flat or noisy channels. Subsequently, the user can specify 1) the center frequency of the line noise (50 or 60 Hz), 2) the feature of interest, high-frequency band (HFB; 65-95 Hz), alpha band (8-12 Hz), beta band (13-30 Hz) or local motor potentials, as well as 3) the expected direction of the feature's effect (e.g., an increase in HFB power as a positive effect, or a decrease in alpha band power as a negative effect). The toolbox then applies a common average reference filter to remove noise shared by all signals, a band-pass filter (0.15-120 Hz; 3rd order Butterworth), a notch filter at the indicated notch base frequency (4th order Butterworth), and down samples the data to 250 Hz. Spectral features are extracted using a Gabor wavelet dictionary, and the local-motor potentials are extracted by applying a 375ms causal moving average on the time-domain data [10].

Trial segmentation and channel selection: The user determines the event code that corresponds to the relevant stimulus being presented to the participant (active cue, Fig. 2), based on a figure that displays all event codes in the data. The user subsequently determines the trial length (TL). The toolbox then extracts active (0s to TL) and rest trials (-TL to 0s) centered around the active cues. To determine which channels are included in subsequent analysis, the signed R^2 value is calculated for each channel by averaging the signal across time for all active and rest trials and regressing those values to a task-regressor (0 for rest and 1 for active trials). Channels are included if the response is significant ($p < 0.05$) and the sign of the R^2 corresponds to the direction of the effect selected by the user (see Preprocessing). At least two channels should be included for method to proceed. The signal is then averaged over the significant channels and smoothed with a 0.2 second window.

Trial selection: This is an optional step in the toolbox. The user enters in the graphical user interface the percentage of trials (T%) to be included in further analysis. The selected trials correspond to the T% of trials with the largest ratio (R_t , where t is a trial) of change in the feature of interest following the active cue ($\Delta y_t =$

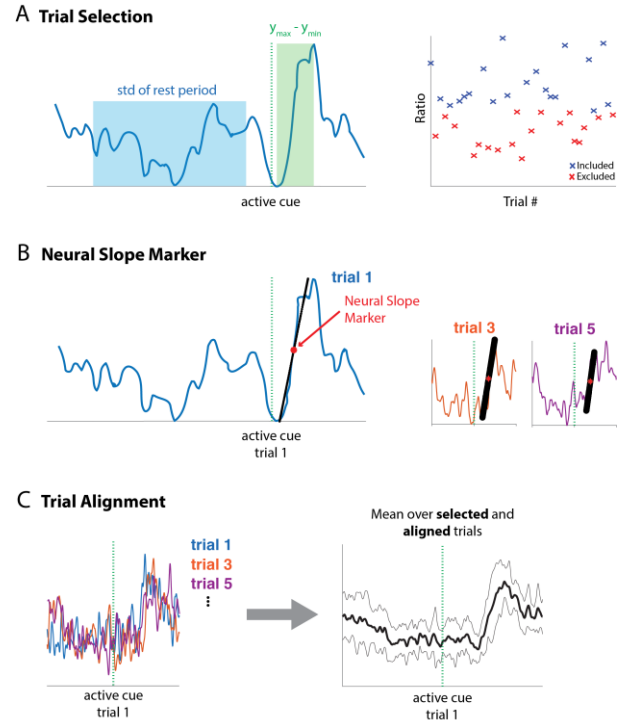


Figure 2: Example output of the TREAT toolbox for one participant. A) Trial selection computes a ratio of the range of active period versus the standard deviation of ‘rest’ period data per trial. A chosen percentage of trials with highest values is selected. B) Trial alignment slides a line segment across trials to identify the minimum distance between the line and the signal. C) Trials are aligned to a reference trial (trial 1).

$y_{max} - y_{min}$ in the time window 0.3 to 1.5 seconds) relative to the standard deviation of the rest period prior to the cue ($\sigma_{rest,t}$; -1.5 to -0.3 seconds) (Fig. 2A):

$$R_t = \frac{\Delta y_t}{\sigma_{rest,t}}. \quad (1)$$

Trial alignment: The alignment procedure builds on the method previously reported in [3] that uses the midpoint of the HFB increase as a marker for movement onset (Fig. 2B). In the current toolbox, this method is extended to other neural features as well. In short, per trial the method slides a line segment over the neural feature (averaged across significant channels) and determines the best fit by minimizing the distance between line and signal (Fig. 2B). It then uses the midpoint of the sliding line segment, hereafter referred to as the ‘neural slope marker’ (NSM), to align all trials (Fig. 2C). In the graphical user interface, the user chooses a time window after the active cue within which the NSM should be detected (Fig. 2B). To compute the slope of the line segment, TREAT first calculates the slope per trial (m_t) by dividing the signal amplitude change in the feature of interest (Δy_t) by the difference in the (time) samples corresponding to the minimum and maximum amplitude values (Δn_t), which is then averaged across trials.

$$m_t = \frac{\Delta y_t}{\Delta n_t} \quad (2)$$

When sliding the line segment over the time window, the method computes for each time point the distance between all the unique y-values of the signal and the line segment that are smaller than a threshold (ε : 0.5 seconds). This threshold ensures that distant time points are not included (for details see [3]). Finally, the midpoint of the line segment corresponding to the minimum distance is used as the NSM (Figure 2B, red point), and each trial is shifted such that its NSM is aligned with the NSM of the first selected trial (reference trial; Fig. 2C). The toolbox displays several steps in the process, including the trial-averaged signal, the best fit of the line segment in each trial and the distance between the two lines for each time point.

Validation data set: We validated the toolbox with a high-density ECoG dataset, of which the task was previously described in detail [4]. In short, three participants (P1: female, 36 years old; P2: female, 30 years old; P3: male, 19 years old) implanted with clinical ECoG grids for epilepsy monitoring, had an additional high-density ECoG grid (3-4 mm inter-electrode distance, 1-1.17 mm exposed electrode diameter) implanted over the sensorimotor cortex for research purposes. Only the high-density data is considered for the current analysis. Four channels with flat or noisy signals were excluded from analysis for P3. Each participant performed 39 trials, in which they produced the /i/ vowel 3-5 times. Voice onsets of the first vowel were determined with Praat software [11]. With the TREAT toolbox, we extracted the HFB and alpha band and compared the timing of the NSM to the timing of the voice onsets. We ran trial selection, including both 100% and 50% of the trials. To investigate potential temporal bias in the distribution of neural slope markers, we also investigated the distribution of NSM over time on the HFB trace of a baseline recording of participant P2 with randomly generated event codes (repeated 1000 times). To run the TREAT toolbox, we chose a trial length of TL = 3 seconds for all participants and frequency bands. The period to search for the NSM was selected per participant based on the timing of their neural responses (Tab. 1).

Table 1: Time periods for sliding the line segment relative to the cue (in seconds) per participant and neural feature.

<i>Participant</i>	<i>HFB</i>	<i>Alpha band</i>
P1	0.3 - 1.8	0.1 - 2.0
P2	0.1 - 1.8	0.1 - 2.5
P3	0.1 - 1.8	0.1 - 2.2

RESULTS

Overall, neural responses to the task were strong for P1 and P2, and less pronounced for P3, as can be seen by the high number of significant responding channels for HFB

and alpha bands (Tab. 2). With TREAT we compared no trial rejection with 50% trial rejection. We noted that a 50% trial rejection increased the signal amplitude change in the HFB by 12.00% for P1, 26.68% for P2 and 17.37% for P3, and in the LFB by 14.11% for P1, 42.51% for P2 and 67.96% for P3.

Table 2: Number of channels with significant positive (in case of the HFB) or negative (in case of the alpha band) R2 values ($p < 0.05$) together with the number of channels included in the analysis.

<i>Participant</i>	<i>HFB</i>	<i>Alpha band</i>
P1	39/128	66/128
P2	68/128	20/128
P3	15/60	12/60

In addition, when using 50% trial selection, the cut-off ratio (see Fig. 2A for an illustration) was larger than 1 for all participants and signal features, apart from the alpha band of P3 (Tab. 3). This indicates that for at least half of the trials, the change in HFB or alpha band power following the cue was larger than the standard deviation over the rest period. Selecting the 50% trials with ratios equal to or larger than the cut-off value resulted in more precise estimation of the NSM relative to VOT compared to including all trials, as reflected by the reduced spread around the means (Fig. 3).

Table 3: Cut-off ratio for 50% trial selection.

<i>Participant</i>	<i>HFB</i>	<i>Alpha band</i>
P1	7.50	5.27
P2	8.07	1.78
P3	4.36	0.98

Our results also show that the NSM occurred before voice onset time (VOT, Fig. 3) for all three participants. For the HFB, the average NSM across trials occurred 0.450 seconds (± 0.229) prior to VOT for P1, 0.407 seconds (± 0.407) for P2 and 0.017 seconds (± 0.711) for P3. For the alpha band, the average NSM across trials occurred 0.420 seconds (± 0.312) for P1, 0.528 seconds (± 0.576) for P2 and 0.008 seconds (± 0.801) for P3 prior to VOT. Repeatedly calculating the NSMs with randomly generated events over a baseline recording resulted in a bimodal distribution, with peaks at the beginning (0.3-0.4 seconds) and end (1.7-1.8 seconds) of the sliding window (Fig. 4A). In contrast, the NSMs acquired with the true events followed a single-peaked distribution (Fig. 4B).

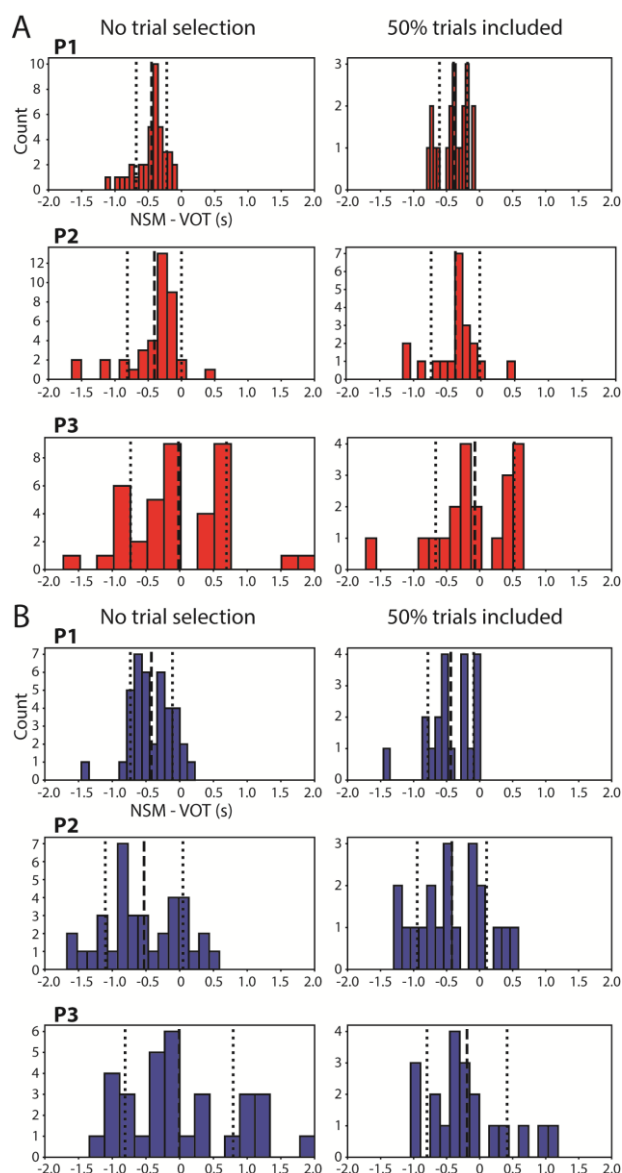


Figure 3: Neural slope marker (NSM) output relative to voice onset time (VOT). We ran the TREAT toolbox without trial selection (left column) and with 50% trial selection for both the high frequency band (A) and the alpha band (B). The dashed line indicates the mean value across trials; the dotted line indicates the standard deviation.

DISCUSSION

Here we introduced the TREAT toolbox, which selects trials based on the magnitude of the neural response to a task, and aligns these responses using the slope of the neural response. We found that the neural onset preceded VOT. This is expected, as the neural response over the sensorimotor cortex initiates the muscle contractions necessary for voice onset to occur [12]. Our results show that NSM centered approximately 400 milliseconds prior to VOT for both the HFB and alpha band for two out of three participants. This value is in line with previous reports that show larynx movement 250 milliseconds before VOT when repeatedly producing the /a/ vowel,

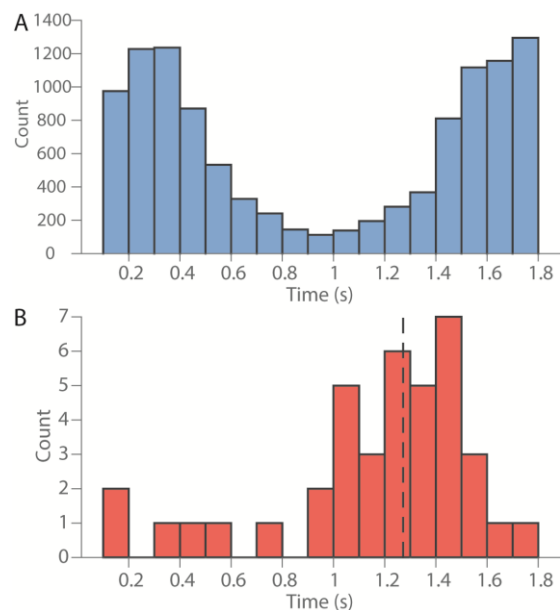


Figure 4: A) TREAT toolbox output with randomly generated event times for the HFB of a baseline recording for P2. 1000 iterations resulted in 293 runs with at least 2 electrodes significantly responding to the task. B) Neural slope marker output for the HFB of P2. The dashed line indicates the median value.

and identified magnetoencephalography activity at least 125 milliseconds prior to movement onset [13], and others measuring EEG activity over the sensorimotor cortex 300 milliseconds prior to voice onset in picture naming tasks [12], [14]. Moreover, spoken phonemes have been shown possible to be decoded from a window 500 milliseconds prior to VOT [5].

Although our sample size is limited, our results indicate that the HFB may serve as a more precise movement-onset marker than the alpha band, which agrees with the hypothesis that HFB power reflects neural population firing underneath the electrodes [15], [16], [17], [18]. The spread of NSM values relative to VOT was larger for P3 than P1 and P2, which indicates that identified NSM were less reliable for P3. This may be related to the relatively poor neural response of P3 compared to the other two participants, as reflected by the lower number of significantly responding channels and cut-off ratios for trial selection. Selection of 50% trials with the strongest response to the active cue reduced the spread of NSM values relative to VOT.

Comparing with other alignment methods, aligning the slopes of the neural responses should increase both the signal amplitude and sharpen the onset in the average across trials, in contrast to aligning to the peak [2] or shape [8] of the responses, that increase the amplitude, but do not sharpen the signal onset per se.

Limitations and future directions: Running the toolbox through data with randomly generated event labels indicates that the random detection of a NSM is more likely at the beginning and the end of the sliding window. This result is expected as at the edges of the sliding window the number of points to compute the difference between the line segment and the signal that falls under

the threshold of 0.5 seconds is smaller than in the middle of the sliding window. Secondly, the participants pronounced the /i/ vowel multiple times within each trial. If they returned their tongues to resting position in between repetitions, each articulation could constitute a distinct movement onset, each accompanied by its own peak in HFB power [4]. As a result, the TREAT toolbox may have detected movement onset not at the first vowel articulation, but later in the trial, if that articulation elicited a stronger increase in HFB power. More extensive comparisons to behavioral markers are required to further validate our approach, including different movements.

Further optimization of the method will involve more stringent electrode selection, fixed line segment size across trials, and optimizing the various settings in the toolbox, such as the percentage of trials to include and signal smoothing, which may be particularly relevant for the HFB as that is an inherently noisy signal. In conclusion, the TREAT toolbox may serve as a versatile event-detection method when behavioral markers are absent, which may improve the interpretation of neural activity as well as the training of decoding models for BCIs.

ACKNOWLEDGEMENTS

This research was supported by the National Institute of Neurological Disorders and Stroke (CortiCom project, grant number: UH3NS114439), the Dutch Technology Foundation STW (PANDA project, grant number 19072), and Health~Holland Top Sector Life Sciences and Health (PPP Allowance made available to the Dutch Brain Foundation; NOTION project, grant number: PPS-2021-02).

REFERENCES

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002, doi: 10.1016/S1388-2457(02)00057-3.
- [2] W. G. Coon and G. Schalk, "A method to establish the spatiotemporal evolution of task-related cortical activity from electrocorticographic signals in single trials," *J. Neurosci. Methods*, vol. 271, pp. 76–85, 2016, doi: 10.1016/j.jneumeth.2016.06.024.
- [3] M. P. Branco, Z. V. Freudenburger, E. J. Aarnoutse, M. G. Bleichner, M. J. Vansteensel, and N. F. Ramsey, "Decoding hand gestures from primary somatosensory cortex using high-density ECoG," *Neuroimage*, vol. 147, pp. 130–142, 2017, doi: 10.1016/j.neuroimage.2016.12.004.
- [4] E. Salari, Z. V. Freudenburger, M. J. Vansteensel, and N. F. Ramsey, "Repeated Vowel Production Affects Features of Neural Activity in Sensorimotor Cortex," *Brain Topogr.*, vol. 32, no. 1, pp. 97–110, Jan. 2019, doi: 10.1007/S10548-018-0673-4.
- [5] N. F. Ramsey, E. Salari, E. J. Aarnoutse, M. J. Vansteensel, M. G. Bleichner, and Z. V. Freudenburger, "Decoding spoken phonemes from sensorimotor cortex with High-Density ECoG grids," *Neuroimage*, vol. 180, no. Pt A, p. 301, Oct. 2017, doi: 10.1016/J.NEUROIMAGE.2017.10.011.
- [6] S. Abrahams, "Neuropsychological impairment in amyotrophic lateral sclerosis–frontotemporal spectrum disorder," *Nature Reviews Neurology* 2023 19:11, vol. 19, no. 11, pp. 655–667, Oct. 2023, doi: 10.1038/s41582-023-00878-z.
- [7] B. Zhou, X. Wu, Z. Lv, L. Zhang, and X. Guo, "A Fully Automated Trial Selection Method for Optimization of Motor Imagery Based Brain-Computer Interface," *PLoS One*, vol. 11, no. 9, p. e0162657, Sep. 2016, doi: 10.1371/JOURNAL.PONE.0162657.
- [8] A. H. Williams *et al.*, "Discovering Precise Temporal Patterns in Large-Scale Neural Recordings through Robust and Interpretable Time Warping," *Neuron*, vol. 105, no. 2, pp. 246–259.e8, 2020, doi: 10.1016/j.neuron.2019.10.020.
- [9] G. Schalk and J. Mellinger, "A practical guide to brain-computer interfacing with BCI2000: General-purpose software for brain-computer interface research, data acquisition, stimulus presentation, and brain monitoring," *A Practical Guide to Brain-Computer Interfacing with BCI2000: General-Purpose Software for Brain-Computer Interface Research, Data Acquisition, Stimulus Presentation, and Brain Monitoring*, pp. 1–260, 2010, doi: 10.1007/978-1-84996-092-2/COVER.
- [10] G. Schalk *et al.*, "Decoding two-dimensional movement trajectories using electrocorticographic signals in humans," *J. Neural Eng.*, vol. 4, no. 3, p. 264, Jun. 2007, doi: 10.1088/1741-2560/4/3/012.
- [11] P. Boersma and D. Weenink, "Praat: doing phonetics by computer," 1992. [Online]. Available: <https://www.praat.org>
- [12] M. Laganaro, "Time-course of phonetic (motor speech) encoding in utterance production," *Cogn. Neuropsychol.*, vol. 40, no. 5–6, pp. 287–297, Aug. 2023, doi: 10.1080/02643294.2023.2279739.
- [13] H. Kothare *et al.*, "Temporal specificity of abnormal neural oscillations during phonatory events in laryngeal dystonia," *Brain Commun.*, vol. 4, no. 2, Mar. 2022, doi: 10.1093/BRAINCOMMS/FCAC031.
- [14] M. Hassan, P. Benquet, A. Biraben, C. Berrou, O. Dufor, and F. Wendling, "Dynamic reorganization of functional brain networks during picture naming," *Cortex*, vol. 73, pp.

- 276–288, Dec. 2015, doi:
10.1016/J.CORTEX.2015.08.019.
- [15] K. J. Miller *et al.*, “Spectral Changes in Cortical Surface Potentials during Motor Movement,” *Journal of Neuroscience*, vol. 27, no. 9, pp. 2424–2432, 2007, doi: 10.1523/JNEUROSCI.3886-06.2007.
- [16] K. J. Miller, L. B. Sorensen, J. G. Ojemann, and M. den Nijs, “Power-law scaling in the brain surface electric potential,” *PLoS Comput. Biol.*, vol. 5, no. 12, p. e1000609, Dec. 2009, doi: 10.1371/journal.pcbi.1000609.
- [17] N. E. Crone, D. L. Miglioretti, B. Gordon, and R. P. Lesser, “Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. II. Event-related synchronization in the gamma band,” *Brain*, vol. 121, no. 12, pp. 2301–2315, 1998, doi: 10.1093/brain/121.12.2301.
- [18] A. Dubey and S. Ray, “Cortical Electroencephalogram (EEG) Is a Local Signal,” *The Journal of neuroscience*, vol. 39, no. 22, pp. 4299–4311, May 2019, doi: 10.1523/JNEUROSCI.2917-18.2019.