

Can we identify the category of imagined phoneme from EEG?

Jerrin Thomas Panachakel¹, Kanishka Sharma¹, Anusha A S¹, and Ramakrishnan A G¹

Abstract—Phonemes are classified into different categories based on the place and manner of articulation. We investigate the differences between the neural correlates of imagined nasal and bilabial consonants (distinct phonological categories). Mean phase coherence is used as a metric for measuring the phase synchronisation between pairs of electrodes in six cortical regions (auditory, motor, prefrontal, sensorimotor, somatosensory and premotor) during the imagery of nasal and bilabial consonants. Statistically significant difference at 95% confidence interval is observed in beta and lower-gamma bands in various cortical regions. Our observations are inline with the directions into velocities of articulators and dual stream prediction models and support the hypothesis that phonological categories not only exist in articulated speech but can also be distinguished from the EEG of imagined speech.

Clinical relevance—Identification of neural correlates of imagined speech helps in developing better prompts for imagined speech based brain-computer interfaces (BCI) leading to improvements in both accuracy and degrees of freedom. BCIs play a significant role as technology aids for differently abled individuals and for patients with disorders of consciousness (DoC). It also helps in better understanding the neural correlates of psychotic disorders such as schizophrenia and paranoia where auditory hallucinations is a major symptom.

I. INTRODUCTION

Imagination is one of the earliest human abilities which has enabled vivid description of visuals and even ideas which often got converted into some invention by mankind. Speech production begins at the level of imagination and then it activates the motor parts in human brain. In imagined, covert, or inner speech, the subjects imagine speaking without any intentional movement of their articulators [1]. During speech imagery, articulatory planning occurs in premotor cortex. Since motor movements are not intended, the information flow is terminated at M1 area [2], [3]. Nevertheless, a motor efference copy is sent to inferior parietal cortex for somatosensory estimation [4]. The perceptual efference copy generated at the inferior parietal cortex is sent to posterior superior temporal gyrus (pSTG) and superior temporal sulcus (STS), leading to an activation in the auditory cortex [5].

This study investigates whether phonological categories are evident in imagined speech. Moving further from such previous studies [6], [7], we try to understand how imagination of prompts from different phonological categories lead to differences in neural activation. The observed differences are interpreted based on various theories of attention modulation, speech production etc. Nasal and bilabial consonants are the two phonological categories chosen for study, since their articulation require different motor actions and hence

these may lead to difference in the functional connectivity in the premotor cortex, where motor planning takes place. Also, due to the perceptual efference copy sent to pSTG and STS, a significant difference is expected in the auditory cortex also. Mean phase coherence (MPC) [8] is used as the metric for identifying the differences in neural activation in various cortical regions.

II. PHONOLOGICAL CATEGORIES CONSIDERED

Phonemes are classified according to their place of articulation and their manner of articulation. The place of articulations deals with the point of contact where an obstruction occurs in the vocal tract between an articulatory gesture, an active articulator, and a passive location. The manner of articulation is the configuration and interaction of the articulators (speech organs such as the tongue, lips, and palate) when making a speech sound.

Two phonological categories are chosen in this study, *viz.* nasal consonants and bilabial consonant. Nasal consonant is produced by an occlusive consonant produced with a lowered velum, allowing air to escape freely through the nose whereas a bilabial consonant is a labial consonant articulated with both lips. These two phonological categories are chosen due to their difference in the place and manner of articulation. It is hypothesized that this difference can lead to difference in neural activation in different cortices of human brain when prompts from these categories are imagined.

III. CORTICAL REGIONS CONSIDERED

A large area of neocortex along with subcortical regions are involved in speech production. The information processing occurs at somatosensory, motor, auditory and prefrontal cortices [9] and hence these regions are considered in this study. In addition, preomotor cortex is considered based on its role in motor planning and sensorimotor cortex based on its role in motor representation [10].

IV. EEG BANDS CONSIDERED

The EEG bands considered in this study are:

A. Alpha & Mu bands

Alpha band (8 Hz to 13 Hz) is usually associated with inhibitory control and attention including auditory attention [11]. An increase in the alpha is observed in the visual cortex when the sensory input to the visual cortex is disrupted [12] and also when the brain processes potentially distracting information [13], [14]. A reduction in alpha is related to an increased ability in discriminate visual targets [15]. An alpha suppression is also associated with speech preparation [16].

¹All authors are with Department of Electrical Engineering, Indian Institute of Science, Bangalore, India jerrinp@iisc.ac.in

Mu band lies in the same frequency band but is differentiated from alpha based on topography and responsivity. The Mu band is observed in the sensorimotor cortex and a suppression in its activity is associated with motor movements or their observation [17]. A suppression in the mu band was also reported by Tamura et al. [18] during speech imagery.

B. Beta band

Beta band (13 Hz to 30 Hz) is associated with motor tasks including motor imagery. An event-related synchronization (ERS) during motor planning and an event-related desynchronization (ERD) after task execution are reported in the beta band. This is true for actual [19] and imagined motor tasks [10]. This ERD/S is observed in sensorimotor cortex [19]. Beta activity is observed in the prefrontal cortex during motor inhibition [20]. Since speech production involves movements of the articulators, changes in the sensorimotor and prefrontal cortices may be observed during speech imagery also.

C. Gamma Band

Gamma activity arises from negative feedback among excitatory and inhibitory neurons. Beta-gamma components have an evolutionary basis which lies in the six-layered neocortex to maintain its functional stability under constant exposure of sensory stimuli [21]. Gamma band recorded from cortical and subcortical regions is found to be modulated in cognitive processing of different memory types and high level of vigilance [22]. Frontal and temporal areas communicate during speech production and increased coherence was found in high gamma range [23]. In this work, lower gamma band from 30 to 45 Hz is considered since the publicly available dataset is bandpass filtered from 1 to 50 Hz.

V. DATASET

This study is based on the publicly available Kara One EEG dataset [6] acquired from 14 subjects when they were imagining articulating 7 phonemic/syllabic prompts (/iy/, /uw/, /piy/, /tiy/, /diy/, /m/, /n/) and 2 phonetically-similar pairs of words (pat, pot, knew, and gnaw). EEG was acquired using 64-channel Neuroscan Quick-cap and SynAmps RT amplifier at a sampling rate of 1 kHz. “/piy/”, “pat” and “pot” are the bilabial prompts and “/n/”, “knew” and “gnaw”, are the nasal prompts. “/m/” is excluded from the analysis since it belongs to both categories.

VI. MEAN PHASE COHERENCE (MPC)

Mean phase coherence (MPC) is a measure of phase synchronisation between two EEG channels [8]. MPC is closely related to phase locking value (PLV) defined for the condition where phase differences between the studied channels are attributed to evoked activity [24]. PLV measures the phase synchronisation between two channels across different trials assuming that every trial is time-locked to a specific stimulus. This assumption does not hold good for EEG acquired during speech imagery since the imagination is not time-locked across trials albeit the presence of cues for the participant.

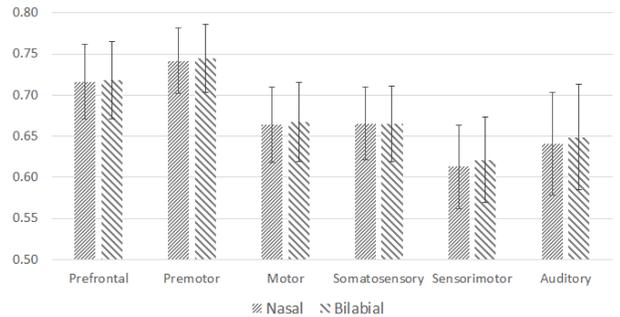


Fig. 1: Average MPC values (given along y-axis) of six cortical regions (given along x-axis) for alpha band and mu band. The difference in mean for none of the cortical region is statistically significant at 95% confidence interval.

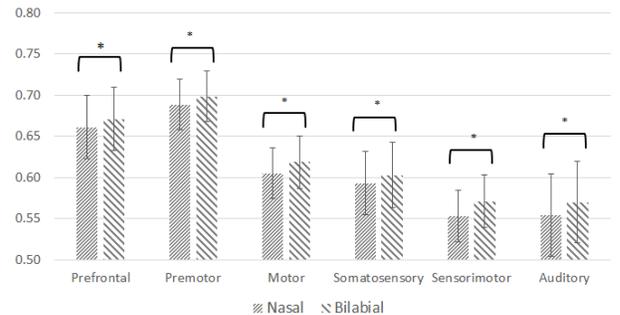


Fig. 2: Average MPC values (y-axis) of six cortical regions (x-axis) for beta band. The differences for all the cortical regions are statistically significant at 95% confidence interval.

MPC across two electrodes i and k are defined as,

$$MPC_{i,k} = \frac{1}{N} \left| \sum_{n=0}^{N-1} e^{-j(\phi_i(n) - \phi_k(n))} \right| \quad (1)$$

where N is the number of samples, $\phi_i(n)$ and $\phi_k(n)$ are the instantaneous phases of channels i and k at the n -th sample time, obtained using Hilbert transform. The value of MPC lies between $[0, 1]$; a value close to zero indicates that the phase differences between the signals are random whereas a value of one means that the two signals are phase synchronized during most of the time interval considered [25]. MPC has been used in epilepsy and sleep studies.

VII. RESULTS

The EEG signal is filtered into difference bands using an FIR filter and the MPC values between all pairs of channels in each cortical region considered are computed for each band and averaged. The obtained values are shown in Figs. 1, 2 and 3. Statistical significance is tested with dependent two-tailed Student’s t-test at 95% confidence interval. Normality is tested for each group using Shapiro-Wilk test at 95% confidence interval. No statistically significant differences are observed in the MPC values for the alpha band in any of the cortical regions when the participants imagined articulating nasal and bilabial prompts. Contrary to this, all the six cortical regions show statistically significant difference in

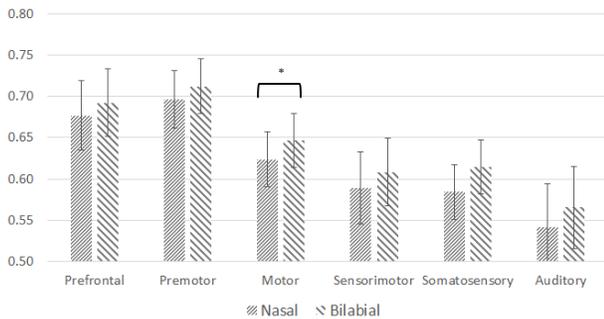


Fig. 3: Average MPC values (y-axis) of six cortical regions (x-axis) for gamma band. Asterisk denotes statistically significant differences at 95% confidence interval.

the MPC values in beta band. In the gamma band, statically significant differences are observed only in the motor cortex.

VIII. DISCUSSION

To our knowledge, this is the first study comparing the EEG based functional connectivity for nasal and bilabial consonants of imagined speech. The significant difference between these consonants may imply different mechanisms inducing functional brain activation. We discuss below our findings under different frequency bands.

A. Alpha & Mu bands

The absence of significant differences in any of the cortices in the alpha band between nasal and bilabial speech imagery can be due to the fact that alpha activity is associated with inhibition, attention etc. and these factors remain the same during the imagination of both nasals and bilabials. A probable exception is in the sensorimotor due to mu band suppression [18]. However, it may not result in a significant difference in MPC values since articulation of both nasals and bilabials require movements which may not lead to differences in the mu band, which can be captured by EEG.

One more reason to expect a difference in the mu band is based on the motor theory of speech perception. According to this theory, spoken words are perceived not by the sounds but by the articulatory gestures that produced them [26]. If this was true, we would have seen a significant difference in the mu band for both bilabials and nasals. This is because if sounds were indeed neurally encoded as articulatory gestures, imagination of these sounds would lead to imagination of these articulatory gestures. As shown by Gastaut and Bert [27], even the imagination of movements result in mu band suppression. The articulatory gestures involved in nasals and bilabials are significantly different. Our results are in line with several other studies [17] that have highlighted the failure of motor theory of speech production in explaining speech perception-production associations.

B. Beta Band

The statistically significant difference between the MPCs of bilabials and nasals in the premotor cortex in the beta band can be due to the differences in the motor planning needed to articulate the distinct prompts. The statistically significant

difference observed in the motor, sensory and sensorimotor cortices can be due to the sensorimotor estimation using motor efference copy [4]. The difference in the auditory cortex may be due to the perceptual efference copy [5]. These observations go inline with the dual stream prediction model for imagined speech proposed by [28]. The idea of efference copy was proposed by Tian and Poppel [5] to explain the activation of auditory cortex during speech imagery.

Beta band activity is observed in the prefrontal cortex (PFC) when an information from the working memory is erased [20] but this cannot explain the difference observed between imagination of nasals and bilabials. A possible explanation is the association of beta band oscillation in PFC with inhibitory motor control [20]. Since different articulators produce nasals and bilabials, the observed difference might arise from the inhibitory control of these articulators [3]. Also, D'Zmura et al. [29] have shown that beta band contains more information than alpha band in the case of speech imagery, which is in line with our results.

C. Gamma Band

Mean values of MPC are significantly different ($p < 0.05$) for the two types of consonants for higher frequencies at the motor cortex. Though the values are higher for bilabials in all the other regions, they are not significant ($p > 0.05$). Imagined speech shares areas of neocortex common to speech production, auditory perception and memory processing [23]. Gamma activity beyond 75 Hz is observed in superior-temporal lobe and temporoparietal junction explored in sentence-level rhythmic neural activity [30]. However, the data utilized for this analysis has frequencies only till 45 Hz. Hence, we cannot study whether the consonants differ in their processing at high gamma frequencies.

Since imagined and articulated speech share common structures and information processing, we expect to get higher frequencies such as gamma, which is involved in crossmodal binding of information. This is achieved by momentary synchronization of neurons from different modalities in the gamma-frequencies [31]. At the level of summation of neuronal firing, gamma acts as gain control [32] of precepts in order to provide a binding function to construct a meaningful inference of stimulus. Guenther [9] has modeled the speech motor movements based on speech sound map. This map seems to be analogous to somatosensory parcellation in sensorimotor homunculus. Based on directions into velocities of articulators model, he argued that the cortical organizational map of speech shares a model similar to those of motor movements of different body parts. Leszczyński et al. showed high frequency activation in sensory-motor and bilateral superior temporal cortices for overt speech [33].

Common neural basis of speech perception and imagery was found in left frontal and temporal areas using fMRI [34]. Common areas were found in speech generation as well as imagined speech; only the magnitude of activation was different [35]. This is partially supported by our results of difference in MPC values in motor cortex for phonemes differing in their articulation.

IX. CONCLUSION

We attempted to discriminate consonants in imagined speech based on the MPC values in different EEG bands. MPC in beta band in every cortical region considered in this study showed statistically significant differences between nasal and bilabials. Alpha and gamma bands (except at motor cortex) were not able to distinguish between the phonemes in terms of the connectivity of different cortical regions. In future, we will explore the intercortical differences and a broadband frequency range for analysis. We may utilize MPC as a feature to distinguish imagined speech at the phoneme level. In [36], we presented the results of using artificial neural networks and MPC for classifying bilabial and nasal prompts in imagined speech.

REFERENCES

- [1] J. T. Panachakel, A. Ramakrishnan, and T. Ananthapadmanabha, "A novel deep learning architecture for decoding imagined speech from EEG," *arXiv preprint arXiv:2003.09374*, 2020.
- [2] J. T. Panachakel and R. A. Ganesan, "Decoding covert speech from EEG—a comprehensive review," *Frontiers in Neuroscience*, vol. 15, p. 392, 2021.
- [3] X. Tian and D. Poeppel, "Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation," *Frontiers in human neuroscience*, vol. 6, p. 314, 2012.
- [4] T. J. Whitford, B. N. Jack, D. Pearson, O. Griffiths, D. Luque, A. W. Harris, K. M. Spencer, and M. E. Le Pelley, "Neurophysiological evidence of efference copies to inner speech," *Elife*, vol. 6, p. e28197, 2017.
- [5] X. Tian and D. Poeppel, "Mental imagery of speech and movement implicates the dynamics of internal forward models," *Frontiers in psychology*, vol. 1, p. 166, 2010.
- [6] S. Zhao and F. Rudzicz, "Classifying phonological categories in imagined and articulated speech," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 992–996.
- [7] J. T. Panachakel, A. Ramakrishnan, and T. Ananthapadmanabha, "Decoding imagined speech using wavelet features and deep neural networks," in *2019 IEEE 16th India Council International Conference (INDICON)*. IEEE, 2019, pp. 1–4.
- [8] F. Mormann, K. Lehnertz, P. David, and C. E. Elger, "Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients," *Physica D: Nonlinear Phenomena*, vol. 144, no. 3–4, pp. 358–369, 2000.
- [9] F. H. Guenther, "Cortical interactions underlying the production of speech sounds," *Journal of communication disorders*, vol. 39, no. 5, pp. 350–365, 2006.
- [10] J. T. Panachakel, N. N. Vinayak, M. Nunna, A. G. Ramakrishnan, and K. Sharma, "An improved EEG acquisition protocol facilitates localized neural activation," in *Advances in Communication Systems and Networks*. Springer, 2020, pp. 267–281.
- [11] W. Klimesch, "Alpha-band oscillations, attention, and controlled access to stored information," *Trends in cognitive sciences*, vol. 16, no. 12, pp. 606–617, 2012.
- [12] S. Palva and J. M. Palva, "New vistas for a-frequency band oscillations," *Trends in Neuroscience*, vol. 30.
- [13] J. J. Foxe and A. C. Snyder, "The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention," *Frontiers in psychology*, vol. 2, p. 154, 2011.
- [14] M. Wöstmann, S.-J. Lim, and J. Obleser, "The human neural alpha response to speech is a proxy of attentional control," *Cerebral cortex*, vol. 27, no. 6, pp. 3307–3317, 2017.
- [15] H. Van Dijk, J.-M. Schoffelen, R. Oostenveld, and O. Jensen, "Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability," *Journal of Neuroscience*, vol. 28, no. 8, pp. 1816–1823, 2008.
- [16] D. A. Bridwell, S. Henderson, M. Sorge, S. Plis, and V. D. Calhoun, "Relationships between alpha oscillations during speech preparation and the listener N400 ERP to the produced speech," *Scientific Reports*, vol. 8, no. 1, pp. 1–10, 2018.
- [17] H. M. Hobson and D. V. Bishop, "The interpretation of mu suppression as an index of mirror neuron activity: past, present and future," *Royal Society Open Science*, vol. 4, no. 3, p. 160662, 2017.
- [18] T. Tamura, A. Gunji, H. Takeichi, H. Shigemasa, M. Inagaki, M. Kaga, and M. Kitazaki, "Audio-vocal monitoring system revealed by mu-rhythm activity," *Frontiers in psychology*, vol. 3, p. 225, 2012.
- [19] M. Zaepffel, R. Trachel, B. E. Kilavik, and T. Brochier, "Modulations of EEG beta power during planning and execution of grasping movements," *PLoS one*, vol. 8, no. 3, p. e60060, 2013.
- [20] R. Schmidt, M. H. Ruiz, B. E. Kilavik, M. Lundqvist, P. A. Starr, and A. R. Aron, "Beta oscillations in working memory, executive control of movement and thought, and sensorimotor function," *Journal of Neuroscience*, vol. 39, no. 42, pp. 8231–8238, 2019.
- [21] W. J. Freeman, "Origin, structure, and role of background eeg activity. part 2. analytic phase," *Clinical Neurophysiology*, vol. 115, no. 9, pp. 2089–2107, 2004.
- [22] O. Jensen, J. Kaiser, and J.-P. Lachaux, "Human gamma-frequency oscillations associated with attention and memory," *Trends in neurosciences*, vol. 30, no. 7, pp. 317–324, 2007.
- [23] L. Lu, J. Sheng, Z. Liu, and J.-H. Gao, "Neural representations of imagined speech revealed by frequency-tagged magnetoencephalography responses," *NeuroImage*, vol. 229, p. 117724, 2021.
- [24] R. Bruña, F. Maestú, and E. Pereda, "Phase locking value revisited: teaching new tricks to an old dog," *Journal of neural engineering*, vol. 15, no. 5, p. 056011, 2018.
- [25] R. Q. Quiroga and S. Panzeri, *Principles of neural coding*. CRC Press, 2013.
- [26] B. Galantucci, C. A. Fowler, and M. T. Turvey, "The motor theory of speech perception reviewed," *Psychonomic bulletin & review*, vol. 13, no. 3, pp. 361–377, 2006.
- [27] H. J. Gastaut and J. Bert, "EEG changes during cinematographic presentation (moving picture activation of the EEG)," *Electroencephalography and clinical neurophysiology*, vol. 6, pp. 433–444, 1954.
- [28] X. Tian and D. Poeppel, "The effect of imagination on stimulation: the functional specificity of efference copies in speech processing," *Journal of cognitive neuroscience*, vol. 25, no. 7, pp. 1020–1036, 2013.
- [29] M. D'Zmura, S. Deng, T. Lappas, S. Thorpe, and R. Srinivasan, "Toward EEG sensing of imagined speech," in *International Conference on Human-Computer Interaction*. Springer, 2009, pp. 40–48.
- [30] J. Kingyon, R. Behroozmand, R. Kelley, H. Oya, H. Kawasaki, N. S. Narayanan, and J. D. Greenlee, "High-gamma band fronto-temporal coherence as a measure of functional connectivity in speech motor control," *Neuroscience*, vol. 305, pp. 15–25, 2015.
- [31] T. R. Schneider, S. Debener, R. Oostenveld, and A. K. Engel, "Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming," *Neuroimage*, vol. 42, no. 3, pp. 1244–1254, 2008.
- [32] B. Merker, "Cortical gamma oscillations: the functional key is activation, not cognition," *Neuroscience & Biobehavioral Reviews*, vol. 37, no. 3, pp. 401–417, 2013.
- [33] M. Leszczyński, A. Barczak, Y. Kajikawa, I. Ulbert, A. Y. Falchier, I. Tal, S. Haegens, L. Melloni, R. T. Knight, and C. E. Schroeder, "Dissociation of broadband high-frequency activity and neuronal firing in the neocortex," *Science advances*, vol. 6, no. 33, p. eabb0977, 2020.
- [34] A. Aleman, E. Formisano, H. Koppenhagen, P. Hagoort, E. H. De Haan, and R. S. Kahn, "The functional neuroanatomy of metrical stress evaluation of perceived and imagined spoken words," *Cerebral Cortex*, vol. 15, no. 2, pp. 221–228, 2005.
- [35] S. M. Smith, Y. Zhang, M. Jenkinson, J. Chen, P. Matthews, A. Federico, and N. De Stefano, "Accurate, robust, and automated longitudinal and cross-sectional brain change analysis," *Neuroimage*, vol. 17, no. 1, pp. 479–489, 2002.
- [36] J. T. Panachakel and A. Ramakrishnan, "Classification of phonological categories in imagined speech using phase synchronization measure," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021.