# Role of Brainwaves in Neural Speech Decoding

Debadatta Dash
*Electrical and Computer Engineering*
*University of Texas at Austin*
Austin, TX, USA
debadatta.dash@utexas.edu

Paul Ferrari
*Department of Psychology*
*University of Texas at Austin*
Austin, TX, USA
pferrari@utexas.edu

Jun Wang
*Communication Sciences and Disorders*
*University of Texas at Austin*
Austin, TX, USA
jun.wang@austin.utexas.edu

***Abstract*—Neural speech decoding aims at direct decoding of speech from the brain to restore speech communication in patients with locked-in syndrome (fully paralyzed but aware). Despite the recent progress, exactly which aspects of neural activities are characterizing the decoding process is still unclear. Neural oscillations have been associated with playing a key functional role in neural information processing and thus might provide significant insight into the decoding process. Previous research has investigated a limited range of neural frequencies for decoding, usually the high-gamma oscillations ($70-200$ Hz) in electrocorticography (ECoG) and lower-frequency waves ($1-70$ Hz) in electroencephalography (EEG). Hence, the exact contribution of specific frequency bands is still unclear. Magnetoencephalography (MEG) is a non-invasive method for directly measuring underlying brain activity and has the temporal resolution needed to investigate the role of cortical oscillations in speech decoding, which we attempted in this study. We used three machine learning classifiers (linear discriminant analysis (LDA), support vector machine (SVM), and artificial neural network (ANN) to classify different imagined and spoken phrases for finding the role of brainwaves in speech decoding. The experimental results showed a significant contribution of low-frequency Delta oscillations ($0.1-4$ Hz) in decoding and the best performance was achieved when all the brainwaves were combined.**

***Index Terms*—brainwave, magnetoencephalography, neural speech decoding, LDA, SVM, ANN**

## I. Introduction

Severe brain damage or amyotrophic lateral sclerosis (ALS) may cause locked-in syndrome, a state of paralysis but with cognitive awareness [1]. These patients lose their communication ability due to articulatory paralysis, leaving only the neural pathway as a medium for restoring a certain level of communication. Current brain-computer interface (BCI) spellers address this challenge by decoding attentional correlates from the brain while the patients focus on selecting letters randomly displayed on a keyboard [2]. The slow communication rate ($<10$ words/minute) of these BCIs is a major impediment for cultivating natural communication. Moving beyond the slow and laborious BCIs, current research is progressing towards finding a solution for fast communication by attempting to decode speech directly from the brain. These neural speech decoding paradigms or speech-BCIs have the potential to offer real-time communication assistance, thereby, improving the quality of life for these neurologically impaired patients.

This work was supported by the University of Texas System Brain Research Grant under award number 362221 and the National Institutes of Health (NIH) under award numbers R03DC013990 and R01DC016621.

Neural speech decoding algorithms use brain signals, acquired either invasively with electrocorticography (ECoG) or non-invasively with magneto-/electro-encephalography (M/EEG), to find patterns corresponding to different speech representations (phonemes/syllables/words/phrases). Considering the low-cost and easy data acquisition setup, EEG has been used extensively to decode both imagined as well as overt speech [3], [4]. ECoG has also been used in neural speech decoding starting from classifying different speech units [5], [6] to continuous neural speech recognition [7] and synthesis [8], [9]. However, the frequency ranges used in decoding vary within these two types of modalities, i.e., usually the high-gamma brainwave ($70-200$ Hz) in ECoG and $<70$ Hz in EEG. These differences seem justified as previous research has shown the large correlation of high-gamma ECoG activity with multi-unit firing rates [10] and the low quality of EEG signal in high-gamma band due to spatial distortion, volume conduction, and large contamination due to movement artifacts [11]. The major advantage of high-gamma signals is the higher spatial specificity [12], which is inherently absent in EEG signals. In short, these differences in brainwave frequencies between these two modalities, make it difficult to understand the distinct role of different brainwaves in decoding. Since brainwaves reflect a certain level of functional processing of the brain [13] and speech production has been characterized with the assimilation of various dynamic cognitive functions understanding the contribution of different brainwaves in decoding might provide a deeper insight into the neural features that are actually being decoded in speech-BCI studies.

Although the role of brainwaves in speech processing has been studied before, the majority of those are for understanding perception [14], rather than speech imagination or production. With ECoG, the distinct contribution of high-gamma brain activity has been repeatedly shown [7], [10], but there is yet to be any similar investigation for speech imagination. There are a few EEG studies which have investigated the comparative performance of brainwaves in decoding such as Theta band dominance in decoding while speaking various words [15] or higher decoding performance by Beta band compared to Alpha and Theta in classifying two syllables [4], etc. Despite the differences in results, a clear investigation of the role of brainwaves is still lacking.

Magnetoencephalography (MEG) provides a direct and reliable representation of the functional characteristics of the brain activity, largely due to its relatively lower signal distortion and spatial smearing [16]. Also, high-gamma power changes have been shown to be observable with MEG [17], [18]. Thus, MEG might be a more suitable modality to investigate the role of distinct brainwaves in speech decoding including high-gamma frequencies. In our previous works on neural speech decoding, we used MEG signals with all brainwaves ($0.1 - 125\,\mathrm{Hz}$) to show the feasibility of speech decoding with high performance [19]–[24]. The high spatial and temporal resolution of MEG has also been shown to be effective in tracking the fast temporal dynamics and spatial neuroanatomical distribution in speech production [25]–[27]. Thus, we used MEG signals to investigate the role of each neural oscillations in decoding. To our knowledge, this is the first such investigation using MEG. We employed three machine learning classifiers namely support vector machine (SVM), linear discriminant analysis (LDA), and artificial neural network (ANN) to classify the MEG signals corresponding to imagination and articulation of five different phrases and found similar patterns of decoding performance with each brainwave across the three classifiers.

## II. Data Acquisition and Signal Processing

### *A. Data Collection*

Seven healthy subjects (age $= 41$y $\pm$ $14$y; 3 females) participated in this study with written consent. Institutional review board (IRB) approvals have been obtained for this study from the participating institutions. The MEG machine (MEGIN, LCC) used (Fig. 1) has 306 channels (204 gradiometers + 102 magnetometers). We used five short phrases as stimuli: *1. Do you understand me, 2. That's perfect, 3. How are you, 4. Goodbye, and 5. I need help*. We designed a time-locked protocol to collect the neuromagnetic signals during speech imagination and production. After a $0.5\,\mathrm{s}$ of pre-stimuli segment, these stimuli were shown on a screen, one at a time, for $1\,\mathrm{s}$ in a pseudo-randomized order. Then a fixation cross replaced the stimulus in the next stage of imagination/preparation, where the subjects were instructed to imagine the phrase and be prepared to speak. After $1\,\mathrm{s}$ of imagination, the screen went blank which heralded the subjects to overtly produce the previously shown phrase. A period of ($1.5 - 2.5\,\mathrm{s}$) was allotted for this articulation segment. This $4$-segment time-locked protocol constituted a trial, and we collected data for $100$ such trials for each phrase. We kept a non-movement baseline of $1\,\mathrm{s}$ within successive trials to suppress the overlapping of cognitive functioning between trials [28]. Jaw motion signals were collected via a custom-made air bladder with a pressure sensor. The audio output during articulation was recorded via a built-in microphone. Both jaw and audio signals were fed to the MEG ADC as separate channels. Considering the difficulty in verifying behavioral aspects of speech imagination [29], we acquired neural signals corresponding to both imagination and production within a single trial.

### *B. Preprocessing*

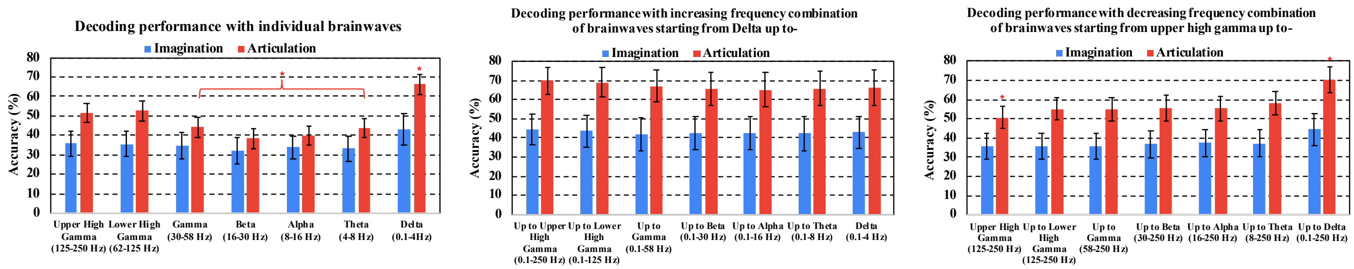
We acquired the MEG signals in $4\,\mathrm{kHz}$ sampling frequency, which were low-pass filtered below $250\,\mathrm{Hz}$ and resampled

Fig. 1. The MEG machine housed inside a magnetically shielded room (MSR)

to $1\,\mathrm{kHz}$. Line noise ($60\,\mathrm{Hz}$) and harmonics were removed with a notch filter. We epoched the MEG signals into trials from $-0.5\,\mathrm{s}$ to $5\,\mathrm{s}$, centered on stimulus onset. Through visual inspection trials containing large artifacts were discarded. Out of $100$ trials, on an average $75$ trials per phrase were retained per subject. Only gradiometer channels were used for analysis considering their high SNR over magnetometers. Flat and noisy channels were excluded from the data. Out of $204$ gradiometers, $196$ common valid sensors across subjects were considered for analysis. The MEG sensor signals were then decomposed with a discrete Daubechies-4 (db-4) wavelet into distinct brainwaves. After a 7-level decomposition, the reconstructed signal from the low pass approximation coefficient at $7^{th}$ level represented the Delta frequency band ($0.1 - 4\,\mathrm{Hz}$), and the reconstructed signals from the high pass coefficients from each level starting from level $7$ to level $2$ corresponded to Theta ($4 - 8\,\mathrm{Hz}$), Alpha ($8 - 16\,\mathrm{Hz}$), Beta ($16 - 30\,\mathrm{Hz}$), Gamma ($31 - 58\,\mathrm{Hz}$), lower high-gamma ($62 - 125\,\mathrm{Hz}$), and upper high-gamma ($125 - 250\,\mathrm{Hz}$) brainwaves respectively. The use of db-4 wavelet to generate distinct brainwaves has been employed previously in various MEG studies [22], [30].

## III. Experimental Results

To investigate the role of brainwaves in neural speech decoding, we performed a 5-class classification of the $196$-dimensional MEG signals corresponding to $5$ phrases during imagination and production, in single-trial level taking each brainwave separately and also taking their orderly combination (increasing/decreasing frequency range) as input. Even though on an average $75$ trials were retained per phrase across subjects, a total of $60$ single trials/phrase were used for unbiased analysis, since for a subject only $63$ trials were retained for a phrase. Considering the cognitive variance across subjects [23], [31], we only developed subject-dependent models. We extracted root mean square (RMS) features from the decomposed signals to train the decoders as these were found to be more significant compared to other statistical features in our prior works [20]–[22]. SVM was taken as the baseline decoder, considering their effectiveness in managing high-dimensional data [32]. A 5-fold cross-validation (CV) strategy was used




(a) (b) (c)

Fig. 2. Comparison of decoding performances with (a) individual brainwaves (b) combination of brainwaves in increasing order of frequency starting from Delta up to different brainwaves, (c) combination of brainwaves in decreasing order of frequency starting from upper high gamma up to different brainwaves. Error bars represent the standard error across 7 subjects. * denotes statistical significance with 1-tail t-test with $p < 0.05$.

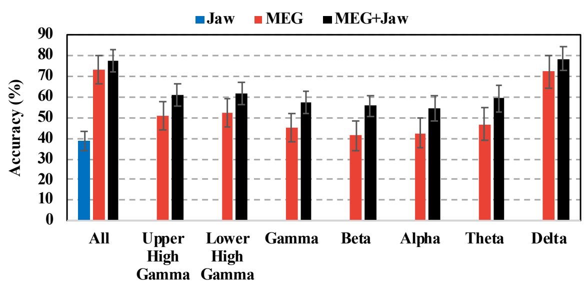



Fig. 3. Comparison of decoding performance with jaw only, MEG brain waves only, and MEG brainwaves with added jaw information. Error bars represent the standard error (SE) across 6 subjects.

to train an SVM with $2^{nd}$ order-polynomial kernel for each subject and the performance was evaluated by averaging the five CV accuracies across all 7 subjects. We experimented with several kernels for SVM including linear, polynomial (2 and 3), Sigmoid, and RBF and then the final kernel ($2^{nd}$ order polynomial) was chosen based on optimal cross-validation performance. Kernel scale and C parameter were tuned and selected based on Bayesian optimization search.

### A. Experiment 1: Individual Brainwaves

To find the role of each brainwave in decoding, first, we trained the SVM with the RMS features extracted from individual oscillations separately. The feature dimension was 196 corresponding to the 196 gradiometer signals. The decoding was performed both for imagination and articulation. Figure 2(a) shows the result of this analysis. Low frequency Delta ($0.1 - 4$ Hz) oscillation provided the best classification accuracy for both imagination ($43.01\%\pm$ SE: $8.26\%$) and articulation ($66.09\%\pm$ SE: $9.04\%$). The decoding performance with Delta across 7 subjects was significantly higher than Theta, Alpha, Beta, and Gamma (1-tail t-test, $p < 0.05$). Accuracy during articulation with lower-high gamma ($52.87\%\pm$ SE: $5.84\%$) and upper-high gamma ($51.70\%\pm$ SE: $5.92\%$) were not statistically significantly lower than Delta (1-tail t-test, $p = 0.12$ & $p = 0.10$ respectively), but mean accuracies of these high-frequency brainwaves were about $15\%$ less. In case of imagination, there was no statistical significance obtained amongst the different brainwaves (1-tail t-test, $p > 0.05$), but, the mean accuracy obtained with Delta was about $8\% - 10\%$ higher than the rest. Performance obtained with all the brainwaves individually were significantly higher (1-tail t-test, $p < 0.05$) than the chance level ($20\%$).

### B. Experiment 2: Orderly Combination of Brainwaves

Considering the efficacy of Delta band obtained in the previous experiment, to reconfirm the significance of this band, first, we included one higher frequency band at a time in addition to the Delta-band and performed the decoding. In other words, first, we included Delta, and Theta bands (net frequency range: $0.1 - 8$ Hz) and performed the decoding. Then we used Delta, Theta, and Alpha (net frequency range: $0.1 - 16$ Hz) and so on up to the upper high-gamma frequency (net frequency range: $0.1 - 250$ Hz, i.e. all brainwaves combined). Instead of concatenating the RMS features of different bands, we reconstructed the signal up to the specific frequency with wavelets for this analysis for an unbiased input feature size of different combinations. The results are shown in Fig. 2(b) which shows a continuous increase in mean accuracy from Delta only ($0.1 - 4$ Hz) to Delta-upper high gamma frequency range ($0.1 - 4$ Hz), although the increment was not statistically significant (1-tail t-test, $p > 0.05$). A similar experiment with opposite direction of frequency band addition (from high to low frequency), starting from upper high gamma only ($125 - 250$ Hz) then upper high gamma + lower high gamma ($62 - 250$ Hz) until all brainwaves combined, i.e. up to Delta from upper high-gamma ($0.1 - 250$ Hz) yielded a similar result. With each addition of frequency band, the classification accuracy increased. However, the statistically significant increase in accuracy for articulated phrase decoding was observed when Delta band was added (1-tail t-test, $p < 0.05$, between upper high gamma only and All). For imagination, the increment was not significant (1-tail t-test, $p > 0.05$).

### C. Experiment 3: Individual Brainwaves with/without Jaw

To verify whether the low-frequency jaw motion is driving the high performance of Delta, we performed the analysis by taking jaw signal only, jaw information added to MEG signal of all brainwaves combined, and then by adding jaw

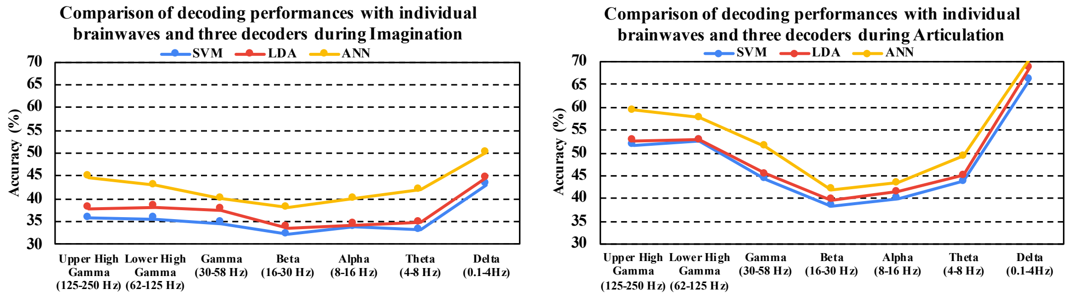



Fig. 4. Comparison of decoding performances with individual brainwaves across three decoders during Imagination (left) and Articulation (right)

information to individual bands. The results are shown in Fig. 3. The average CV accuracy across 6 subjects by taking only RMS features of jaw signal (1-dimensional feature) was $38.50\% \pm$ SE: $4.51\%$ which was significantly lower (1-tail t-test, $p < 0.05$, Wilcoxon signed-rank test, $p < 0.05$) than the average decoding accuracy obtained with RMS features of MEG signals (196-dimensional) with all frequency ($0.1 - 250\,\text{Hz}$) which was $73.30\% \pm$ SE: $6.83\%$. When jaw RMS feature is concatenated with MEG RMS features (197-dimensional) the mean accuracy significantly increased by about $4\%$ to $77.60\% \pm 5.34\%$ (Wilcoxon signed-rank test, $p < 0.05$). Further, when jaw RMS feature was concatenated with the RMS features of individual brainwaves, a significant performance improvement was observed for each band across all subjects (Wilcoxon signed-rank test, $p < 0.05$). However, the increment with Delta+Jaw from Delta only was the lowest ($6.12\%$) compared to the increment of about $12\%$ obtained when jaw information was added with other brainwaves.

### *D. Experiment 4: Individual Brainwaves with Three Decoders*

All of the previous results were with the $2^{nd}$ order polynomial SVM with 5-fold CV. To confirm whether the patterns observed with different brainwaves are significant or classifier (SVM) dependent we used 2 more classifiers: LDA and ANN, and repeated the analysis with individual brainwaves (*Experiment 1*). The same 5-fold cross-validation strategy was used for these 2 decoders with the same sample size and features. Automatic hyperparameter optimization was performed to find the best LDA parameters for Dirichlet distributions. For ANN, only a single hidden layer with 128 nodes was used, followed by a sigmoid activation and softmax each with 5 nodes. ANN was trained with scaled conjugate gradient optimization with backpropagation. The learning rate was fixed to 0.005, found based on coarse-to-fine tuning. The comparative results of the three decoders are shown in Fig. 4 (left) for imagination and (right) for production, as line plots to visualize the consistency in the patterns of accuracy across different brainwaves. As expected, SVM and LDA performed similarly and ANN outperformed LDA and SVM performance both during imagination and articulation. However, the interesting observation was in the consistency in the pattern of performance of different brainwaves irrespective of the three classifiers.

## IV. Discussion

Here, we used the classification accuracy as an index of contribution for different brainwaves in decoding. Although a more traditional way of finding the dominance is by visualizing the time-frequency representations on the averaged trials, it is difficult to do so for single-trial analysis. The consistency of higher performance obtained with Delta band for speech decoding (Fig. 2) has been observed before primarily for perceived speech [14] and the predominance of Delta in auditory perception of multi-word speech units has been shown [33]. Thus, perception of the subjects speech may be driving this high decoding performance for articulation. This also explains the contrasting results obtained in this study to an EEG study [4] where Beta band dominated classifying two syllables. Since we used multi-word sentences as stimuli, Delta band might be reflecting the combinatorial processes underlying the unification of words to sentences [14]. However, intriguingly, a higher performance by Delta was also observed during imagination decoding compared to the rest of the brainwave frequencies. This may suggest that similar neuro-cognitive constructs are occurring during imagination as articulation.

It might be argued that low-frequency articulatory jaw motion might be driving the decoding performance in the Delta band. To check on this hypothesis, we performed *Experiment* 3 which suggested that the brain signals alone carry additional information over that of transmitted by the jaw. These findings are consistent with our previous work with three subjects [20]. Adding the jaw information on top of MEG signals, both individual frequencies and the wide-band, increased the classification accuracy (Fig. 3) which suggests that the MEG brainwaves and jaw motion might entail complimentary information, and that the Delta band encompasses both jaw motion neural information vital for decoding. Further, the relative performances across different brainwaves were consistent with three different classifiers. This provides strong support in the relative contribution of different brainwaves, higher contribution of Delta and high gamma frequencies compared to the rest. However, combining all the frequencies always provided the best result. Instead of retaining each frequency-bands power for the multi-band comparison, we chose to widen the band prior to taking the power to keep the decoding feature dimensions equivalent across frequency bands. Since

Delta will always dominate the wide-band spectrum, it may have diminished the potential influence of the other bands. Additionally, in our experiments, we considered all the sensors for analysis. It is also possible that different sensors at different brainwaves might be dominating the speech imagination or production process. Thus, in future, we plan to analyze the role of the combination of brainwaves post-transform and across individual sensors and their combinations.

## V. Conclusion

In this study, we investigated the role of individual brainwaves in decoding speech imagination and production. Significantly higher performance was observed for the Delta band which was validated with different experiments. Adding jaw information to individual brainwaves increased the decoding performance suggesting that they have complementary information. The relative performances across different brainwaves were found to be consistent after verifying with three different classifiers. The best performance was obtained when all the brainwaves are combinedly used for decoding. This study only used healthy participants and a similar investigation for ALS patients might provide better insight into the ultimate goal of speech-BCIs for neurologically impaired patients.

## References


[1] E. Smith and M. Delargy, "Locked-in syndrome," *BMJ*, vol. 330, no. 7488, pp. 406–409, 2005.

[2] J. S. Brumberg, A. Nieto-Castanon, P. R. Kennedy, and F. H. Guenther, "Brain-computer interfaces for speech communication," *Speech Commun.*, vol. 52, no. 4, pp. 367–379, Apr. 2010.

[3] C. Cooney, R. Folli, and D. Coyle, "Optimizing layers improves cnn generalization and transfer learning for imagined speech decoding from EEG," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Oct 2019, pp. 1311–1316.

[4] M. D'Zmura, S. Deng, T. Lappas, S. Thorpe, and R. Srinivasan, "Toward EEG sensing of imagined speech," in *Human-Computer Interaction. New Trends*, J. A. Jacko, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 40–48.

[5] S. F. Martin, P. Brunner, I. Iturrate, J. d. R. Millán, G. Schalk, R. T. Knight, and B. N. Pasley, "Corrigendum: Word pair classification during imagined speech using direct brain recordings," in *Scientific reports*, vol. 7, no. 44509, 2016.

[6] E. M. Mugler, J. L. Patton, R. D. Flint, Z. A. Wright, S. U. Schuele, J. Rosenow, J. J. Shih, D. J. Krusienski, and M. W. Slutzky, "Direct classification of all American English phonemes using signals from functional speech motor cortex," *Journal of Neural Engineering*, vol. 11, no. 3, p. 035015, may 2014.

[7] C. Herff, D. Heger, A. de Pesters, D. Telaar, P. Brunner, G. Schalk, and T. Schultz, "Brain-to-text: decoding spoken phrases from phone representations in the brain," *Front. in Neurosci.*, vol. 9, p. 217, 2015.

[8] M. Angrick, C. Herff, E. Mugler, M. C. Tate, M. W. Slutzky, D. J. Krusienski, and T. Schultz, "Speech synthesis from ECoG using densely connected 3D convolutional neural networks," *Journal of Neural Engineering*, vol. 16, no. 3, p. 036019, apr 2019.

[9] G. Anumanchipalli, J. Chartier, and E. F. Chang, "Speech synthesis from neural decoding of spoken sentences," *Nature*, vol. 568, pp. 493–498, 04 2019.

[10] N. Crone, L. Hao, J. Hart, D. Boatman, R. Lesser, R. Irizarry, and B. Gordon, "Electrocorticographic gamma activity during word production in spoken and sign language," *Neurology*, vol. 57, no. 11, pp. 2045–2053, 2001.

[11] E. M. Whitham, K. J. Pope, S. P. Fitzgibbon, T. Lewis, C. R. Clark, S. Loveless, M. Broberg, A. Wallace, D. DeLosAngeles, P. Lillie, A. Hardy, R. Fronsko, A. Pulbrook, and J. O. Willoughby, "Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG," *Clinical Neurophysiology*, vol. 118, no. 8, pp. 1877 – 1888, 2007.

[12] K. J. Miller, E. C. Leuthardt, G. Schalk, R. P. N. Rao, N. R. Anderson, D. W. Moran, J. W. Miller, and J. G. Ojemann, "Spectral changes in cortical surface potentials during motor movement," *Journal of Neuroscience*, vol. 27, no. 9, pp. 2424–2432, 2007.

[13] X.-J. Wang, "Neurophysiological and computational principles of cortical rhythms in cognition," *Physiological Reviews*, vol. 90, no. 3, pp. 1195–1268, 2010.

[14] A. Ksem and V. van Wassenhove, "Distinct contributions of low- and high-frequency neural oscillations to speech comprehension," *Language, Cognition and Neuroscience*, vol. 32, no. 5, pp. 536–544, 2017.

[15] P. Suppes, Z.-L. Lu, and B. Han, "Brain wave recognition of words," *Proceedings of the National Academy of Sciences*, vol. 94, no. 26, pp. 14 965–14 969, 1997.

[16] F. L. da Silva, "EEG and MEG: Relevance to neuroscience," *Neuron*, vol. 80, no. 5, pp. 1112 – 1128, 2013.

[17] D. Cheyne, S. Bells, P. Ferrari, W. Gaetz, and A. C. Bostan, "Self-paced movements induce high-frequency gamma oscillations in primary motor cortex," *NeuroImage*, vol. 42, no. 1, pp. 332 – 342, 2008.

[18] S. S. Dalal, A. G. Guggisberg, E. Edwards, K. Sekihara, A. M. Findlay, R. T. Canolty, M. S. Berger, R. T. Knight, N. M. Barbaro, H. E. Kirsch, and S. S. Nagarajan, "Five-dimensional neuroimaging: Localization of the timefrequency dynamics of cortical activity," *NeuroImage*, vol. 40, no. 4, pp. 1686 – 1700, 2008.

[19] J. Wang, M. Kim, A. W. Hernandez-Mulero, D. Heitzman, and P. Ferrari, "Towards decoding speech production from single-trial magnetoencephalography (MEG) signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2017, pp. 3036–3040.

[20] D. Dash, P. Ferrari, S. Malik, and J. Wang, "Overt speech retrieval from neuromagnetic signals using wavelets and artificial neural networks," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Nov 2018, pp. 489–493.

[21] D. Dash, P. Ferrari, S. Malik, A. Montillo, J. A. Maldjian, and J. Wang, "Determining the optimal number of MEG trials: A machine learning and speech decoding perspective," in *Brain Informatics*. Cham: Springer International Publishing, 2018, pp. 163–172.

[22] D. Dash, P. Ferrari, and J. Wang, "Decoding imagined and spoken phrases from non-invasive neural (MEG) signals," *Frontiers in Neuroscience*, vol. 14, p. 290, 2020.

[23] D. Dash, A. Wisler, P. Ferrari, and J. Wang, "Towards a Speaker Independent Speech-BCI Using Speaker Adaptation," in *Proc. Interspeech 2019*, 2019, pp. 864–868.

[24] D. Dash, P. Ferrari, D. Heitzman, and J. Wang, "Decoding speech from single trial MEG signals using convolutional neural networks and transfer learning," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2019, pp. 5531–5535.

[25] D. Dash, P. Ferrari, S. Malik, and J. Wang, "Automatic speech activity recognition from MEG signals using seq2seq learning," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, March 2019, pp. 340–343.

[26] N. Memarian, P. Ferrari, M. J. Macdonald, D. Cheyne, L. F. D. Nil, and E. W. Pang, "Cortical activity during speech and non-speech oromotor tasks: A magnetoencephalography (MEG) study," *Neuroscience Letters*, vol. 527, no. 1, pp. 34 – 39, 2012.

[27] D. Dash, P. Ferrari, S. Dutta, and J. Wang, "NeuroVAD: Real-time voice activity detection from non-invasive neuromagnetic signals," *Sensors*, vol. 20, no. 8, p. 2248, 2020.

[28] S. Baillet, *MEG, an Introduction to Methods*. Oxford University press, 01 2010, pp. 83–123.

[29] C. Cooney, R. Folli, and D. Coyle, "Neurolinguistics Research Advancing Development of a Direct-Speech Brain-Computer Interface," *iScience*, vol. 8, pp. 103–125, Oct 2018.

[30] A. S. Ghuman, J. R. McDaniel, and A. Martin, "A wavelet-based method for measuring the oscillatory dynamics of resting-state functional connectivity in MEG," *NeuroImage*, vol. 56, no. 1, pp. 69 – 77, 2011.

[31] D. Dash, P. Ferrari, and J. Wang, "Spatial and Spectral Fingerprint in the Brain: Speaker Identification from Single Trial MEG Signals," in *Proc. Interspeech 2019*, 2019, pp. 1203–1207.

[32] B. Ghaddar and J. Naoum-Sawaya, "High dimensional data classification and feature selection using support vector machines," *European Journal of Operational Research*, vol. 265, no. 3, pp. 993 – 1004, 2018.

[33] N. Ding and J. Z. Simon, "Neural coding of continuous speech in auditory cortex during monaural and dichotic listening," *Journal of Neurophysiology*, vol. 107, no. 1, pp. 78–89, 2012, pMID: 21975452.