INTERSPEECH 2020
October 25–29, 2020, Shanghai, China

Neural Speech Decoding for Amyotrophic Lateral Sclerosis

Debadatta Dash[1,2], *Paul Ferrari*[3,4], *Angel Hernandez*[5], *Daragh Heitzman*[6], *Sara G. Austin*[2], *Jun Wang*[2,7]

[1]Department of Electrical and Computer Engineering, University of Texas at Austin
[2]Department of Neurology, Dell Medical School, University of Texas at Austin
[3]MEG Lab, Dell Children's Medical Center
[4]Department of Psychology, University of Texas at Austin
[5]Division of Neurosciences, Helen DeVos Children's Hospital
[6]MDA/ALS Center, Texas Neurology
[7]Department of Speech, Language, and Hearing Sciences, University of Texas at Austin

debadatta.dash@utexas.edu, pferrari@utexas.edu, angel.hernandez@helendevoschildrens.org, dheitzman@texasneurology.com, sgaustin@ascension.org, jun.wang@austin.utexas.edu

Abstract

Amyotrophic lateral sclerosis (ALS) is a motor neuron disease that may cause locked-in syndrome (completely paralyzed but aware). These locked-in patients can communicate with brain-computer interfaces (BCI), e.g. EEG spellers, which have a low communication rate. Recent research has progressed towards neural speech decoding paradigms that have the potential for normal communication rates. Yet, current neural decoding research is limited to typical speakers and the extent to which these studies can be translated to a target population (e.g., ALS) is still unexplored. Here, we investigated the decoding of imagined and spoken phrases from non-invasive magnetoencephalography (MEG) signals of ALS subjects using several spectral features (band-power of brainwaves: delta, theta, alpha, beta, and gamma frequencies) with seven machine learning decoders (Naive Bayes, K-nearest neighbor, decision tree, ensemble, support vector machine, linear discriminant analysis, and artificial neural network). Experimental results indicated that the decoding performance for ALS patients is lower than healthy subjects yet significantly higher than chance level. The best performances were 75% for decoding five imagined phrases and 88% for five spoken phrases from ALS patients. To our knowledge, this is the first demonstration of neural speech decoding for a speech disordered population.

Index Terms: ALS, BCI, neural speech decoding, MEG, PCA

1. Introduction

Amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig's disease, is an idiopathic, fatal, and fast progressive neurodegenerative disease characterized by the degeneration of upper and lower motor neurons (at the spinal or bulbar level) disrupting the ability of the brain to control voluntary motor function leading to dysphagia, dysarthria, and impaired limb and respiratory function [1]. Regardless of the site of onset, progressive bulbar motor deterioration is common to most ALS patients which leads to the impairment of speech and thereby substantially diminishes the communication ability and shortens survival [2]. Reestablishing communication for these patients can significantly improve their quality of life. Current augmented and assistive communication (AAC) devices help these patients by using their residual movements, e.g., cheek twitches or eyeball movements, to navigate a cursor on alphabet displays to spell out words, although these devices tend to be slow, error-prone, and laborious [3]. Unfortunately, most patients with ALS eventually advance to a state of complete paralysis called locked-in syndrome (LIS), which can be also caused by severe brain damage. LIS is characterized by quadriplegia and anarthria but with preservation of consciousness [4], i.e., awake but selectively deafferented having no means of producing speech, limb or facial movements. Current AAC devices fail for these patients. The brain may be the only output pathway that can provide a level of communication.

Current EEG-BCI spellers provide a means of communication for these patients but the rates are slow (under 10 words per minute) [3]. Slow communication rate impacts the actual use of the device, as patients experience fatigue. Recent studies on neural speech decoding showed potential for faster communication rate, which can be achieved by decoding overt or covert speech directly from the brain and bypassing the corticospinal pathway. The efficacy of this decoding paradigm has been shown previously for invasive and non-invasive recordings using ECoG, MEG, and EEG, respectively [5–17]. However, these studies are limited to typical speakers with no communication disorders (epileptic patients in these ECoG studies have no speech disorders), whereas, the target group for this research is patients with LIS. It is still unclear whether if speech decoding can be successfully performed on the neural signals of a target population (with speech impairment).

In this study, we investigated the neural decoding of imagined and spoken phrases from the neural signals of patients with ALS. We used magnetoencephalography (MEG) to record the neuromagnetic signals of both healthy and ALS subjects while they were imagining and speaking different short phrases. MEG has a high spatial and temporal resolution and thus is suitable to study the fast spatio-temporal dynamics of speech processing in the brain [18–21]. Previously, we have shown the efficacy of using MEG for neural speech decoding with high accuracy from healthy subjects [13–17]. Although current MEG machines are non-portable and costly, recent studies on optically pumped magnetometer based MEG systems [22–24] including decoding of imagination [25] and language [26] showed the potential of building next-generation, movable, wearable, and low-cost OPM-MEG mediated speech-BCIs in the near future.

Seven machine learning classifiers were used as decoders, which were to classify the neuromagnetic signal features corresponding to five imagined and spoken phrases. We used the spectral features of the neural signals, i.e., the band power of

http://dx.doi.org/10.21437/Interspeech.2020-3071

Figure 1: *An MEG unit with a participant with ALS. Unlike healthy controls who prefer to sit up, participants with ALS prefer to lay down in the machine due to their potential fatigue.*

delta, theta, alpha, beta, gamma, lower high gamma, and upper high gamma and their combination to train the decoders. Spectral features have been shown to contribute to robust decoding performance in various decoding experiments [10,11,13,14]. In addition, we analyzed the decoding performance by implementing PCA compression on the concatenation of spectral features.

2. Data Collection and Preprocessing

Ten individuals (7 healthy; 3 females; age= 41 ± 14 years and 3 ALS; 1 female; age= 52 ± 12 years) participated in the study with voluntary consent. The healthy subjects had no history of language or cognitive disorders. The ALS subjects were in early to mid-stage of the disease with mild but noticeably speech impairment. Although their speech intelligibility was still beyond 90% (word recognition accuracy based on listening tests), they spoke significantly slower (as slow as 132 words per minute) than healthy controls (typically 200 words per minute). One patient had bulbar onset, one had spinal onset, and the other had generalized ALS symptoms. Two identical Elekta Neuromag TRIUX MEG machines (MEGIN, LCC) at the Dell Children's Medical Center in Austin, TX and the Cook Children's Medical Center in Fort Worth, TX (Figure 1) were used. IRB approvals have been obtained from the corresponding institutions prior to this study. The machines are housed within a magnetically shielded room to restrict external magnetic field noise. The machines have 204 gradiometers and 102 magnetometers. ECG and EOG each with two bipolar integrated sensors were used to collect the heartbeat and eye-blink artifacts.

A time-locked protocol was designed for the speech imagination and production task. Stimulus onset was characterized by displaying a visual stimulus generated by a computer running the STIM2 software (Compumedics, Ltd.), and presented via a DLP projector situated at 90 cm from the subjects'. Five AAC phrases were used as stimuli (*1. Do you understand me, 2. That's perfect, 3. How are you, 4. Good-bye, 5. I need help*). The phrases were shown for one second, one at a time, in a pseudo-randomized order. Then the subjects imagined and prepared to speak the shown phrase for 1 s, after which they overtly produced the sentence at their natural speaking rate. This procedure was repeated 100 times per phrase per subject. To overcome the difficulty in verifying the timing of imagined speech, we designed our protocol to collect both imagination and production consecutively in the same trial under time constraints.

The signals were recorded at 4 kHz sampling rate which were low pass filtered with a 4^{th} order Butterworth low-pass filter with cutoff frequency 250 Hz and resampled to 1 kHz. Notch filters with cutoff frequencies 60 Hz and its harmonics were applied to remove the line noise. Through visual inspection, trials containing high amplitude artifacts and untimely articulated trials were removed with an average rejection rate of 25%. The first 60 trials for each phrase were used for analysis. Data only from gradiometer channels were used for decoding owing to their effectiveness in noise suppression over magnetometers. Noisy and unresponsive channels were removed. Data from 196 gradiometers were used in the decoding task.

3. Methods

3.1. Features

Brainwaves play a key functional role in neural information processing and thus might provide significant insight into the decoding process. MEG provides a direct and reliable representation of the functional oscillatory characteristics of the brain activity, largely due to its high temporal resolution and low signal distortion compared to EEG [27–29], thereby, making it suitable to use the wide range of its oscillatory dynamics. Thus, in this study, we explored the efficacy of spectral features for which we computed the band-power of different brainwaves of the MEG signals. The frequency ranges of for power calculation were: delta ($0.3 - 4$ Hz), theta ($4 - 8$ Hz), alpha ($8 - 15$ Hz), beta ($15 - 30$ Hz), gamma ($30 - 59$ Hz), lower high gamma ($61 - 119$) Hz and upper high gamma ($121 - 250$ Hz). The spectral features were extracted from each of the 196 gradiometer signals for each trial making the feature dimension to be 196 for each trial. We also performed the decoding analysis by concatenating all the spectral features as the input to the decoders. Considering the high dimension of sensor and feature concatenation (1372 = 196 sensors $\times$ 7 brainwaves), we performed PCA on the concatenated features and reduced the feature dimension with a 90% variance threshold.

3.2. Decoders

Seven machine learning algorithms were used to classify the MEG features corresponding to different phrases. They were Naive Bayes (NB), K-nearest neighbor (KNN), decision trees (DT), Ensembles (EN), support vector machine (SVM), linear discriminant analysis (LDA), and artificial neural network (ANN). Naive Bayes classifier was designed with a normal distribution predictor and Gaussian Kernel. The KNN classifier was designed with a K value of 5 and Euclidean distance metric. For decision tree, the maximum number of splits were set to 5, with leaves being merged that originated from the same parent node yielding a sum of risk values greater or equal to the risk associated with the parent node and the optimal sequence of pruned trees were set to be estimated. The predictor selection method was set to 'Standard CART' that selects the split predictor that maximizes the split-criterion gain over all possible splits of all predictor and the split criterion was 'gdi' (Gini's diversity index). For ensembles, the aggregate method was bagging with random predictor selections at each split (random forest). The maximum number of learning cycles was set to 100 with a discriminant analysis weak learner. For SVM, a 2^{nd} order polynomial kernel was chosen based on cross-validation (CV). 'C' parameter and kernel scale were optimized based on the Bayesian

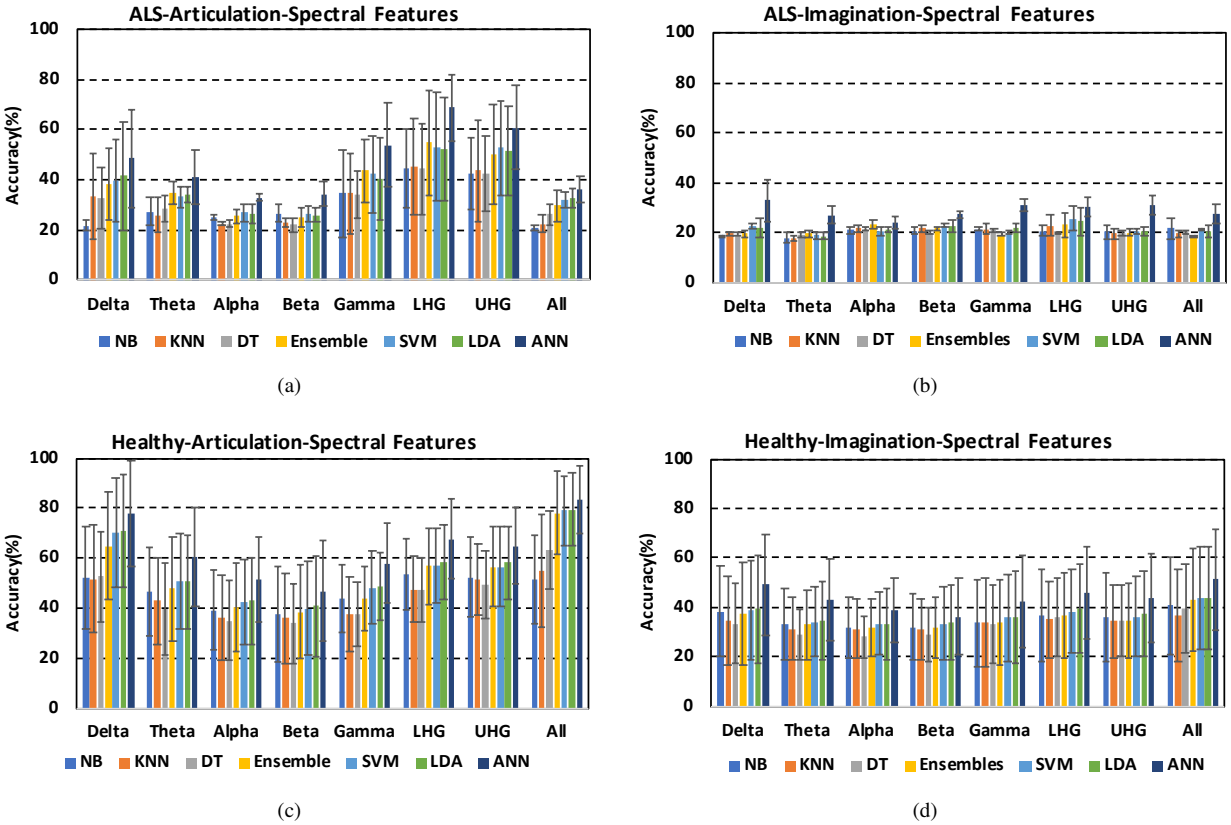


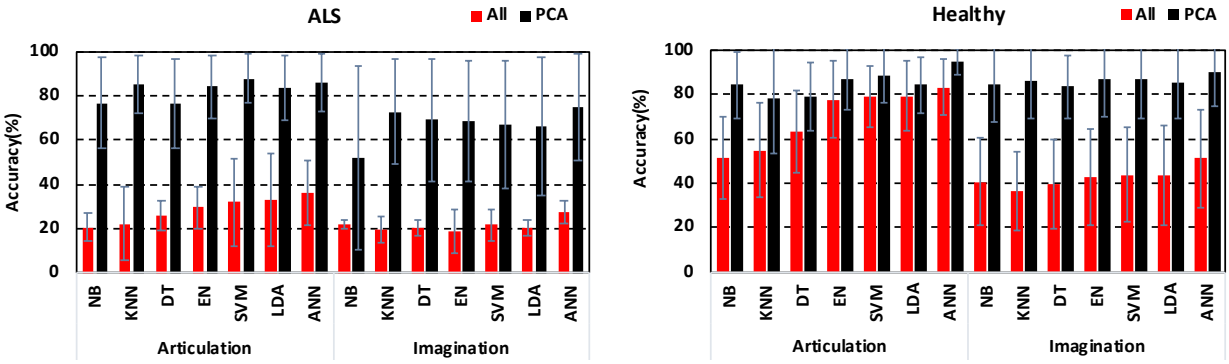
Figure 2: *Performances of healthy and ALS subjects during articulation and imagination using spectral features and seven decoders (a) ALS-Articulation (b) ALS-Imagination (c) Healthy-Articulation (d) Healthy-Imagination. LHG and UHG represent lower and upper high gamma brainwaves respectively. 'All' represents all spectral features concatenated. Error bars indicate standard deviation.*

optimization search. Considering the lower sample size than the feature dimension, we chose linear type discriminant analysis. The linear coefficient threshold and the amount of regularization were computed based on the Bayesian optimization search. The ANN decoder had one hidden layer of 128 nodes followed by a sigmoid and then a softmax layer each of 5 nodes. The training was performed using a scaled conjugate gradient optimizer with backpropagation with a learning rate of 0.01 (based on coarse to fine tuning) for a maximum of 100 epochs with early stopping validation patience of 6 epochs. Considering the high cognitive variance across subjects [15, 30, 31], here we performed subject-dependent decoding. We trained the decoders with two strategies, i.e., first, we implemented 5-fold cross-validation (12 trials/phrase in one fold) and second, we performed a sequential split of the data, where we used the first 48 recorded trials for training and the next 12 trials for testing because for ALS patients the chance of performance deterioration is more likely with time due to the higher motor load for articulatory compensation.

4. Results and Discussion

Figure 2 shows the average accuracies for 7 healthy participants and 3 participants with ALS during imagination and articulation obtained with the spectral features evaluated via seven classifiers. The accuracies for both imagination and articulation decoding for ALS were significantly lower than the healthy across all classifiers and features (1-tail t-test, $p < 0.05$) except (high) Gamma frequencies. The significantly lower performance in the case of ALS compared to healthy subjects indicates the difficulty in neural speech decoding for patients. The result in spoken (articulated) speech decoding was expected, considering the likely abrupt motor behavior such as articulatory compensation or diaphragmic pulsating during speech production exhibited by the patients. The results in speech imagination decoding for these patients were surprising, because the accuracies were as low as chance level (20%) indicating the difference in brain activity between healthy and ALS in speech imagination. This is possibly due to fatigue of the participants with ALS for imagination task (attention). As mentioned earlier, patients with ALS preferred to lay down in the MEG machine, while all healthy controls preferred to sit up during the experiment. Future studies with a larger number of subjects and in-depth neural pattern analysis, however, are needed to verify these findings and provide a better interpretation.

When comparing the decoding accuracies obtained with individual brainwaves, a consistent pattern was observed for both healthy and ALS subjects, which signifies the distinct role of brainwaves in neural speech decoding. For articulation decoding, the best performances were with high gamma frequencies ($61-119$ Hz; $121-250$ Hz) consistent with the previous MEG and ECoG findings [29, 32]. Delta band ($0.3-4$ Hz) also resulted in relatively better performance for both healthy and ALS subjects. Since we used multi-word sentences as stimuli, delta



(a) (b)

Figure 3: *Performances of ALS and healthy subjects during articulation and imagination using concatenated spectral features and PCA compressed features and seven decoders: (a) ALS (b) Healthy. Error bars indicate standard deviation.*

oscillations might be reflecting the combinatorial processes underlying the unification of words to sentences as observed in previous speech perception studies [33, 34]. It might be the self-perceived speech by the subjects that is driving the decoding process with this brainwave. In imagined speech decoding, the role individual brainwaves were not distinct for both the control and target population group except a slightly higher performance (about 3%) was obtained with lower high gamma features for ALS.

Figure 3 demonstrates the significant improvement in decoding performance with the compressed feature set that was obtained by PCA based dimension reduction. With a threshold of 90% variance, 10 principal components were selected, and the decoding was performed on the compressed feature set both for healthy and ALS data. The improvement in accuracies, both for imagination and spoken speech decoding was statistically significant across both the population group (1-tail paired t-test, $p < 0.05$). After compression, the articulation decoding performances for ALS were almost similar to healthy subjects, however, the final accuracy after PCA reduction during imagination decoding was still lower than healthy subjects across the evaluations of various classifiers. Nevertheless, this improvement suggests the efficacy of dimension reduction for decoding speech in both healthy and ALS. After PCA, the best classification accuracy for healthy subjects during imagination decoding was 90.10% and for articulation decoding, it was 94.91% with the ANN classifier. For ALS, after PCA, the best articulation decoding accuracy of 87.78% was obtained with the SVM classifier (ANN: 85.50%) and the best imagination decoding accuracy was with ANN classifier 74.57% (Figure 3).

Regarding the decoders, ANN outperformed the rest with the next better performance by both SVM and LDA and then satisfactory performances with Ensembles and Decision trees. The performances obtained with KNN and NB classifiers were the lowest. The better performances even with a shallow neural network over other machine learning classifiers suggest that these accuracies might further be improved with deep learning classifiers given enough sample size. However, with PCA all the classifiers performed similarly well suggesting the encouraged discriminating distribution after PCA compression (see supplementary materials, Figure S1, where t-SNE distributions of the features before and after PCA compression is shown). Please note, although the results reported here were with the cross-validation approach, the sequential split approach (training with first 48 trials and testing with next 12 trials per phrase) also resulted in similar performance (see supplementary materials, Figure S2), which indicates the absence of any performance deterioration within the first 60 trials. Considering the ultimate goal of speech-BCIs is imagined speech decoding for patients, these observations facilitate the need for better feature extraction strategies and decoders.

5. Conclusions

We showed the possibility of decoding imagined and spoken phrases from non-invasive neural (MEG) signals of patients with mild ALS. To our knowledge, this is the first demonstration of neural speech decoding from a population with neurological speech disorders. Spectral features (i.e., band-power of brainwaves) were used for decoding which reproduced the higher performance of high-gamma frequencies in speech decoding. Relatively lower decoding performance was observed for patients with ALS compared to healthy but significantly above chance level. PCA compression significantly improved the decoding performance for both the healthy and ALS group for both imagination and spoken speech decoding. This pilot study included data from three ALS patients only for one session. Analysis with a larger number of more severe ALS subjects (LIS) with multiple sessions is needed to verify the efficacy of this study. Further, a higher level of neurolinguistic understanding of the imagined speech would help design algorithms for higher decoding performance on imagined speech.

6. Acknowledgments

This work was supported by the University of Texas System Brain Initiative under award number 362221 and partly by the National Institutes of Health (NIH) under award numbers R03DC013990 and R01DC016621. We would like to thank Dr. Saleem Malik, Dr. Ted Mau, Dr. Mark McManis, Kristin Teplansky, Dr. Alan Wisler, Saara Raja, and the volunteering participants.

7. References

[1] M. C. Kiernan, S. Vucic, B. C. Cheah, M. R. Turner, A. Eisen, O. Hardiman, J. R. Burrell, and M. C. Zoing, "Amyotrophic lateral sclerosis," *The Lancet*, vol. 377, no. 9769, pp. 942 – 955, 2011.

[2] J. R. Green, Y. Yunusova, M. S. Kuruvilla, J. Wang, G. L. Pattee, L. Synhorst, L. Zinman, and J. D. Berry, "Bulbar and speech motor assessment in ALS: Challenges and future directions," *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, vol. 14, no. 7-8, pp. 494–500, 2013.

[3] J. S. Brumberg, K. M. Pitt, A. Mantie-Kozlowski, and J. D. Burnison, "Brain-computer interfaces for augmentative and alternative communication: A tutorial," *American Journal of Speech-Language Pathology*, vol. 27, no. 1, pp. 1–12, 2018.

[4] S. Laureys, F. Pellas, P. Van Eeckhout, S. Ghorbel, C. Schnakers, F. Perrin, J. Berre, M.-E. Faymonville, K.-H. Pantke, F. Damas *et al.*, "The locked-in syndrome: what is it like to be conscious but paralyzed and voiceless?" *Progress in brain research*, vol. 150, pp. 495–611, 2005.

[5] G. K. Anumanchipalli, J. Chartier, and E. F. Chang, "Speech synthesis from neural decoding of spoken sentences," *Nature*, vol. 568, no. 7753, pp. 493–498, 2019.

[6] 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," *Frontiers in neuroscience*, vol. 9, p. 217, 2015.

[7] 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, 2019.

[8] G. Krishna, C. Tran, J. Yu, and A. H. Tewfik, "Speech recognition with no speech or with noisy speech," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 1090–1094.

[9] A. R. Sereshkeh, R. Trott, A. Bricout, and T. Chau, "Online EEG classification of covert speech for brain–computer interfacing," *International journal of neural systems*, vol. 27, no. 08, p. 1750033, 2017.

[10] C. Cooney, R. Folli, and D. Coyle, "Mel frequency cepstral coefficients enhance imagined speech decoding accuracy from EEG," in *2018 29th Irish Signals and Systems Conference (ISSC)*. IEEE, 2018, pp. 1–7.

[11] X. Chi, J. Hagedorn, D. Schoonover, and M. D'Zmura, "EEG-based discrimination of imagined speech phonemes," *International Journal of Bioelectromagnetism*, vol. 13, no. 4, pp. 201–206, 2011.

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

[13] 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. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2020.00290

[14] 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)*, 2019, pp. 5531–5535.

[15] 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. [Online]. Available: http://dx.doi.org/10.21437/Interspeech.2019-3109

[16] 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*, S. Wang, V. Yamamoto, J. Su, Y. Yang, E. Jones, L. Iasemidis, and T. Mitchell, Eds. Cham: Springer International Publishing, 2018, pp. 163–172.

[17] 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)*, 2018, pp. 489–493.

[18] 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)*. IEEE, 2019, pp. 340–343.

[19] 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.

[20] P. G. Simos, J. Breier, G. Zouridakis, and A. Papanicolaou, "Identification of language-specific brain activity using magnetoencephalography," *Journal of Clinical and Experimental Neuropsychology*, vol. 20, no. 5, pp. 706–722, 1998.

[21] N. Memarian, P. Ferrari, M. J. Macdonald, D. Cheyne, F. Luc, 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.

[22] A. Borna, T. R. Carter, A. P. Colombo, Y.-Y. Jau, J. McKay, M. Weisend, S. Taulu, J. M. Stephen, and P. D. Schwindt, "Non-invasive functional-brain-imaging with an OPM-based magnetoencephalography system," *Plos one*, vol. 15, no. 1, p. e0227684, 2020.

[23] E. Boto, N. Holmes, J. Leggett, G. Roberts, V. Shah, S. S. Meyer, L. D. Muñoz, K. J. Mullinger, T. M. Tierney, S. Bestmann *et al.*, "Moving magnetoencephalography towards real-world applications with a wearable system," *Nature*, vol. 555, no. 7698, pp. 657–661, 2018.

[24] T. M. Tierney, N. Holmes, S. Mellor, J. D. López, G. Roberts, R. M. Hill, E. Boto, J. Leggett, V. Shah, M. J. Brookes *et al.*, "Optically pumped magnetometers: From quantum origins to multi-channel magnetoencephalography," *NeuroImage*, 2019.

[25] D. N. Barry, T. M. Tierney, N. Holmes, E. Boto, G. Roberts, J. Leggett, R. Bowtell, M. J. Brookes, G. R. Barnes, and E. A. Maguire, "Imaging the human hippocampus with optically-pumped magnetoencephalography," *NeuroImage*, vol. 203, p. 116192, 2019.

[26] T. M. Tierney, N. Holmes, S. S. Meyer, E. Boto, G. Roberts, J. Leggett, S. Buck, L. Duque-Muoz, V. Litvak, S. Bestmann, T. Baldeweg, R. Bowtell, M. J. Brookes, and G. R. Barnes, "Cognitive neuroscience using wearable magnetometer arrays: Non-invasive assessment of language function," *NeuroImage*, vol. 181, pp. 513 – 520, 2018.

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

[28] J. Kaiser, F. Walker, S. Leiberg, and W. Lutzenberger, "Cortical oscillatory activity during spatial echoic memory," *European Journal of Neuroscience*, vol. 21, no. 2, pp. 587–590, 2005.

[29] 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.

[30] D. Dash, P. Ferrari, and J. Wang, "Spatial and spectral fingerprint in the brain: Speaker identification from single trial MEG signals," *Proc. Interspeech 2019*, pp. 1203–1207, 2019.

[31] H. Watanabe, H. Tanaka, S. Sakti, and S. Nakamura, "Subject-independent classification of japanese spoken sentences by multiple frequency bands phase pattern of EEG response during speech perception," *Proc. Interspeech 2017*, pp. 2431–2435, 2017.

[32] A. Korzeniewska, P. J. Franaszczuk, C. M. Crainiceanu, R. Kuś, and N. E. Crone, "Dynamics of large-scale cortical interactions at high gamma frequencies during word production: event related causality (ERC) analysis of human electrocorticography (ECoG)," *Neuroimage*, vol. 56, no. 4, pp. 2218–2237, 2011.

[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.

[34] A. Kösem 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.