

## Exploratory methods for high-performance EEG speech decoding

Lindy Comstock<sup>1,2</sup>, Claudia Lainscsek<sup>3,4</sup>, Vinícius R. Carvalho<sup>5</sup>,

Eduardo M. A. M. Mendes<sup>5</sup>, Aria Fallah<sup>1</sup>, and Terrence J. Sejnowski<sup>3,4,6</sup>

<sup>1</sup>*Department of Neurosurgery, University of California, Los Angeles, Los Angeles, CA 90095, USA*

<sup>2</sup>*Department of Linguistics, Higher School of Economics, Moscow 101000, RF*

<sup>3</sup>*Computational Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA*

<sup>4</sup>*Institute for Neural Computation, University of California San Diego, La Jolla, CA 92093, USA*

<sup>5</sup>*Postgraduate Program in Electrical Engineering,*

*Federal University of Minas Gerais, Belo Horizonte, MG 31270-901, Brazil*

<sup>6</sup>*Division of Biological Sciences, University of California San Diego, La Jolla, CA 92093, USA*

(Dated: November 16, 2021)

State-of-the-art technologies in neural speech decoding utilize data collected from microwires or microarrays implanted directly into the cerebral cortex. Yet as a tool accessible only to individuals with implanted electrodes, speech decoding from devices of this nature is severely limited in its implementation, and cannot be considered a viable solution for widespread application. Speech decoding from non-invasive EEG signals can achieve relatively high accuracy (70-80%), but only from very small classification tasks, with more complex tasks typically yielding a limited (20-50%) classification accuracy. We propose a novel combination of technologies in which transcranial magnetic stimulation (TMS) is first applied to augment the neural signals of interest, producing a greater signal-to-noise ratio in the EEG data. Next, delay differential analysis (DDA) – a cutting-edge computational method based on nonlinear dynamics – is implemented to capture the widest range of information available in the neural signal, by incorporating both linear and nonlinear dynamics.

### I. INTRODUCTION

In recent years, the field of speech decoding has made significant advances that place the goal of real-time translation from neural signals within sight [28, 44]. At the same time, current technologies still suffer from major limitations that require innovative new solutions to overcome. Brain-computer interface (BCI) devices that utilize neural signals from invasive implants remain inaccessible for the majority of their target user population [40], whereas BCI devices that decode from electroencephalography (EEG) signals remain insufficiently fast with low accuracy and a limited range of classification abilities [32, 35, 36]. Ultimately, the most realistic chance for wide-spread application of speech decoding technology will require a non-invasive method that can achieve a similar classification accuracy accomplished by invasive methods. This pilot study aims to increase the accuracy and generalizability of speech decoding from non-invasive EEG neural signals by means of a novel combination of technologies: the application of transcranial magnetic stimulation (TMS) to augment the neural signal of interest, resulting in greater signal-to-noise ratio in the EEG data, and delay differential analysis (DDA), a cutting-edge computational method that incorporates both linear and nonlinear dynamics to capture the widest range of information available in the neural signal.

Currently, the best metrics for speech decoding from EEG signals are obtained by means of a convolutional neural network (CNN) [36]. However, these results only achieve high performance (78-89%) for simple binary classification tasks which register the presence or absence of certain categories of articulatory movements that produce speech sounds (phonemes). The classification of actual speech sounds or words yields insufficiently robust results to support a realistic technology: between 16-54% accuracy for detecting one out of a set of linguistic items (words or phoneme strings) [36]. While speech decoding from implanted electrodes succeeds at substantially more difficult classification tasks, the efficacy of such methods is boosted primarily by language prediction models, which can improve median word error detection rate up to 35% [29]. For example, words in a possible set of 50 items obtained a mean classification accuracy of 47% in post-hoc offline analyses [29]. This means that in the absence of facilitatory language prediction models, the actual speech decoding classifier performs at only slightly better than 20%, even with a data type that is more stable across sessions and possesses a significantly better signal-to-noise ratio.

Instability across sessions in the composition of the cortical activity being recorded is another key challenge to EEG speech decoding, compounded by the low signal-to-noise ratio of this data type and its propensity to generate artifacts in the process of data collection. This instability in combination with the complexity of some machine learning classifiers can require repeated recalibration of the model for each session or individual user, which is time-consuming and requires specialist knowledge. Development of an algorithm that can provide reliably accurate speech decoding across all users without frequent recalibration over time or personalization of model features for each individual user remains an area of development that has yet to be addressed in the speech decoding literature. A generalized or

“universal” model of this nature would substantially aid the usability of the system for the target population outside of the research community. It is also important to note that both of the approaches reviewed here are computationally intensive and achieve their best results only offline with the benefit of large amounts of training data. For effective real-time translation that can compensate for novel words which the algorithm may not have previously encountered, an alternative to the data- and time-intensive model of traditional neural networks should be developed.

In light of these challenges, we propose that DDA can serve as a method of analysis for the reliable, quick, and accurate classification of motor EEG signals which may be adopted by a wide profile of other BCI researchers as a tool in their own research. Preliminary results suggest that TMS may aid this process by evoking a more consistent cortical-wide neural response across individuals, such that a universal DDA model for speech decoding may be developed.

## II. BACKGROUND

DDA is a powerful classification tool that combines differential embeddings with linear and nonlinear nonuniform functional delay embeddings. DDA has several advantages in comparison with CNNs, the most successful method thus far for the analysis of EEG data in speech decoding. Both CNNs and DDA effectively process high-dimensional data; however, CNNs do so by means of a high computational resources, such that they can be slow to train for complex tasks and they require a substantial amount of training data prior to performing the classification task and to avoid overfitting. Moreover, CNNs may require data augmentation when there is substantial variation within the data type to be classified. The complex architectures of CNNs require precise tuning for the problem of interest and they have difficulty modeling long distance relationships and require short data inputs to perform well. This may be one reason why simple classification tasks with one phoneme or word inputs have been performed well by CNNs, but not more complex decoding problems. Brain data is inherently high-dimensional, but the neural processes in speech decoding differ substantially in nature from the text or vision recognition tasks to which CNNs are often applied. It is widely assumed that many neural mechanisms in the brain are nonlinear [2, 27]. Thus, the integration of nonlinear dynamics into a classification analysis allows information from the data to be detected which is not observable in traditional linear methods.

Computationally, DDA boasts two key advantages over traditional machine learning analyses. Firstly, DDA requires minimal preprocessing, as it captures only the relevant dynamics for the classification problem at hand. This provides a crucial advantage, in that preprocessing may be subjective and conditioned by individual differences between participants and the artifacts incurred during data collection. Preprocessing can strip away information in the signal that is meaningful to the neural process under analysis. Eliminating this step in the pipeline also allows for a faster analysis. Secondly, DDA uses a small set of model features. In this regard, DDA has a clear advantage over CNNs. A sparse model imparts speed to the analysis, reduces the need for huge computational resources, and virtually eliminates the likelihood of overfitting the model. Figure 1 provides an overview of the two-step process by which DDA is carried out: (i) model selection, which involves choosing the model structure and parameters that best fit the overall dynamics of the EEG data. Model selection can be supervised or unsupervised. In this analysis, we take a supervised approach in order to optimize structure selection by random-subsampling cross validation, a generalization of bootstrapping. The second step involves data analysis, during which data is fitted to the chosen model. The corresponding coefficients are used as features for the classification problem.

DDA has been successfully applied to a range of data types, including dolphin echolocation data (object identification) [14], heart ECG data (differentiating heart diseases) [22, 23], EEG data (sleep staging, biomarker development for schizophrenia patients) [17, 20, 37] Parkinsonian movement data in combination with EEG data (biomarker development, disease progression assessment) [15, 16, 18, 21], and iEEG data from epilepsy patients (seizure prediction, dynamical state detection) [13, 19, 25]. In findings of particular relevance, DDA has successfully analyzed speech signals (audio signal recognition [12], emotion detection in speech [9]). This previous success illustrates the suitability of DDA for modeling neural data and physiological processes more generally.

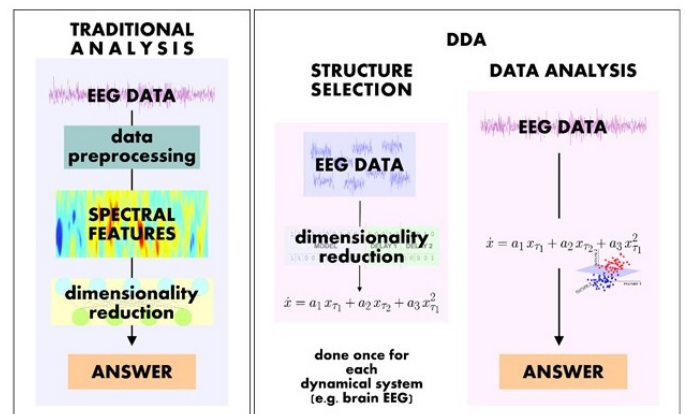


FIG. 1: Comparison of DDA and Machine Learning Analysis.

A more exploratory approach concerns the application of TMS. TMS acts upon cortical neurons by inducing an electrical field that depolarizes the membrane potential and pushes neurons past an excitation threshold. However, there remains considerable debate as to what exact effect TMS may have upon neural processes. TMS has been argued to act upon the same neurons implicated in volitional motor commands [1], an argument that is supported by task-related TMS-EEG studies. Findings from such studies suggest that task-specific stimulation elicits task-dependent neural responses and that the functional improvements seen in task performance are related to changes in cortical reactivity found close to the stimulated region [5, 34, 39]. For example, paired pulse TMS has been shown to improve subject performance on a phoneme discrimination task when motor areas involved in the production of two categories of phonemes were stimulated: reaction times and errors in the phoneme discrimination task decreased for each phoneme when the site associated with that phoneme was targeted [7]. The logical conclusion is that stimulation of the relevant motor area leads to either (i) facilitation of the behavioral response, corresponding to an increased ability to perceive the related phoneme, due to priming of the concordant motor representation, or (ii) inhibition of the alternative behavioral response, corresponding to a reduced ability to perceive the unrelated phoneme, due to priming of the discordant motor representation.

In either case, we argue that stimulation of the motor area should therefore evoke an optimal state for detecting the relevant neural signals necessary for decoding the phoneme of interest. Generally speaking, the theoretical assumptions of this approach are commonly accepted within the speech decoding literature. Motor circuits have been shown to be involved in the speech-perception process [33] and reliance upon neural activation within areas of the motor cortex associated with speech production is standard practice for speech decoding among BCI researchers [4, 30]. The exact nature of this optimal state is less clear and remains an important question in understanding the physiological response to TMS. Discharging action potentials that are near threshold may create a homogeneous state in the stimulated region, in effect reducing the noise that typically interferes with detecting the relevant signals. In this scenario, features may be isolated that can be used as inputs for the classification analysis by means of inhibiting competing neural signals. Alternatively, TMS may augment the motor signals relevant for the decoding task. When the motor cortex is stimulated, TMS elicits motor-evoked potentials (MEPs) and cortical responses [10] termed TMS-evoked potentials (TEPs) [8], as well as a range of induced neural oscillations and connectivity changes [42, 43]. These responses are stable over sessions performed a week apart [3], indicating they could serve as inputs for machine learning classifiers. Thus, analysis of the EEG signals collected during TMS may more readily allow for identification of cortical locations with a usable signal, helping to indicate areas of focus when using EEG prospectively.

We aim to test these assumptions in performing speech decoding by means of DDA on a dataset of phonemes (4 consonants and 5 vowels) that have been targeted with stimulus-specific TMS. We propose that the combined application of TMS and DDA will allow for superior speech decoding classification by (i) optimizing the stimulus-specific information in the neural signal, and (ii) analyzing both linear and nonlinear dynamics to access the widest possible range of useful information in the neural signal.

### III. METHODS

**Experimental design.** Participants listened to consonant phonemes played aurally and identified each with a button-press response. Areas in the motor cortex associated with the production of each category of phonemes were stimulated: lip and tongue regions for bilabial (/b/,/p/) and dental (/t/,/d/) phonemes, respectively (see Figure 2). To aid intelligibility, each phoneme was followed by one of five vowels (/i/,/ɜ/,/ɑ/,/u/,/oʊ/). Two single pulses were administered, separated by a short 50 ms interpulse interval. Pulses were timed so that the last TMS pulse occurred 50 ms prior to consonant presentation. In the stimulated trials, the two pulses were delivered at 110% of the FDI rMT. In each of 4 blocks, participants completed 80 trials, 60 with TMS and 20 random catch trials. Random unstimulated catch trials serve as a reference to evaluate the effect of TMS by indicating the baseline classification accuracy in decoding speech from EEG.

**Participants.** Healthy adults between the ages of 20 and 40 were recruited by means of fliers from the UCLA community. Eligibility criteria included no prior or concurrent diagnosis of any neurological (e.g., epilepsy, Tourette’s syndrome), psychiatric (e.g., schizophrenia), or developmental (e.g., ADHD, dyslexia) disorders; and no structural brain abnormalities (e.g., aneurysm). Participants provided written informed consent and were paid for two sessions. In the first session, an MRI scan was conducted. The scan was used to target areas in the motor cortex associated with phoneme production. In the second session, participants performed a phoneme discrimination task.

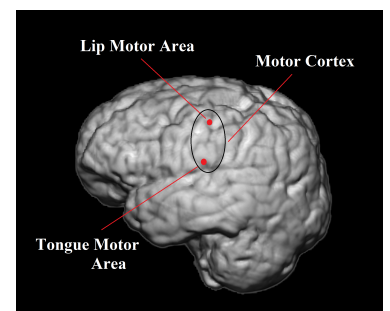


FIG. 2: TMS Targets.

**Data collection.** EEG data collection was conducted in the Neuromodulation Division of the Semel Institute for Neuroscience and Human Behavior at UCLA. TMS was carried out by means of a Magstim Super Rapid Plus1 stimulator, and the stimulation targets were identified using the Visor 2 neuronavigation system (ANT Neuro). Subject-specific targets were generated for each participant by means of transforming targets previously identified in the literature for the lip and tongue motor areas [7] to the subject-space of each participant MRI scan. The TMS coil was positioned using frameless stereotaxy. Coil orientation was maintained at 45 degrees with respect to the interhemispheric fissure. Electrode positions were digitized and registered to individual subject MRIs using the ANT Neuro Xensor (ANT Neuro). EEG signals were then bandpass-filtered 0.1-350 Hz, sampled at 2000 Hz, and referenced to an additional forehead electrode. All electrode impedances were kept  $< 5 \text{ k}\Omega$ . All data collection measures were approved by the IRB.

**Motor thresholding.** Each subject's motor threshold was determined in the presence of a physician to select the appropriate intensity of stimulation. Motor evoked potentials (MEPs) were elicited in the first dorsal interosseus (FDI) muscle of the dominant hand at the minimum amount of stimulation needed to evoke a MEP in a hand muscle after a single pulse over the primary motor cortex. Single TMS pulses were delivered in the motor cortex contralateral to the dominant hand. The intensity of the stimulation was gradually lowered until reaching a level at which 5 out of 10 MEPs in a hand muscle have an amplitude of at least 50 microvolts.

#### IV. ANALYSIS

**Delay differential analysis (DDA).** DDA combines differential embeddings with linear and nonlinear nonuniform functional delay embeddings [31, 38, 41]. The use of two delay variables relates the current derivatives of a system to current and past values of the system, [11, 13, 26]. The DDA model maps experimental data onto a set of natural embedding coordinates. The general DDA model with two delays and three terms is

$$\dot{u}(t) = \sum_{i=1}^3 a_i u(t - \tau_1)^{m_i} u(t - \tau_2)^{n_i} \quad (1)$$

where  $u(t)$  is a time series,  $m_i, n_i, \tau_{1,2} \in \mathbb{N}_0$  and a degree  $m_j + n_j \leq 4$  of nonlinearity. We use DDA models with two time delays and three terms to reduce complexity. Note, that the delays are independent from each other and have no physical meaning for a nonlinear DDA model [24]. They are specific to the question of research, here phoneme identification. In DDA we use the coefficients  $a_i$  and the least square error  $\rho$  as independent parameters or features. The model is fixed and is not updated during the analysis. DDA and its models have several advantages over the high dimensional feature spaces of other signal processing techniques: (i) due to sparsity of the model, the risk of overfitting is greatly reduced; (ii) it is insensitive to noise since such a tiny model concentrates on the overall dynamics of the system and cannot additionally model noise; (iii) it is computationally fast; (iv) there is no need for pre-processing except normalization to zero mean and unit variance, which will cause the model to ignore amplitude information and concentrate on system dynamics.

DDA is a two-step process. (i) For a new class of data (e.g. EEG data) the best DDA model (i.e. the coefficients  $m_i$  and  $n_i$  as well as the delays  $\tau_{1,2}$  that best fit the overall dynamical properties of the system) need to be selected. This can be done by supervised (maximizing the classification performance) or unsupervised (minimizing the least square error) structure selection from a list of candidate models [25, 26]. The performance is evaluated by the area under the receiver operating characteristic (ROC) curve (AUC or  $A'$ ). This is essentially a plot of the true positive rate against the false positive rate. This step is done once and does not have to be repeated for new data from the same data class (e.g. EEG data). For the question of interest (here phoneme classification), the best delays were selected. There is one model that has been found to represent the overall nonlinear dynamical structure of EEG data:

$$\dot{u}(t) = a_1 u(t - \tau_1) + a_2 u(t - \tau_2) + a_3 u(t - \tau_1)^2. \quad (2)$$

This model appears unique to EEG data [25]. (ii) After the DDA model is established, data can be analyzed by fitting the data to the model and estimating the features  $(a_1, a_2, a_3, \rho)$ . We estimate the parameters from the data without any pre-processing or filtering except normalizing each data window to zero mean and unit variance.

Prior to conducting the classification analysis, the EEG data was spliced into individual items (consonant/vowel pairs) of roughly 250 ms in length. A 35 ms window from consonant onset was used to ensure that the auditory input from vowels did not contaminate the consonant classification problem. While alternative approaches may utilize a larger time window in order to capture the full range of evoked auditory potentials, this decision has a precedent in previous literature that found a short time window at phoneme onset to be the most informative for speech decoding [30].

In the current study, we identify one out of four possible consonant phonemes by means of a four-way classification in which the best delays  $\tau_{1,2}$  are selected for each consonant to be distinguished from the other three phonemes.

TABLE I: Phoneme targets for the four-way classifiers.

	/b/	/p/	/t/	/d/
$C_1$	1	0	0	0
$C_2$	0	1	0	0
$C_3$	0	0	1	0
$C_4$	0	0	0	1

The targets for the four classifiers  $C_{1,2,3,4}$  and the four phonemes are shown in Table I. We applied all four classifiers using a one-against-all approach in which each phoneme is treated as the positive class, one at a time. The classifier with the highest value is selected as the correct label. Performance is reported as the percentage of correctly identified phonemes.

An exact comparison of our method and that of previous authors, which would account for all possible experimental design factors, is not possible due to a lack of access to the data sets used by other research teams. However, the percentage of correctly identified phonemes (accuracy) is a common metric in speech decoding papers that can be approximately interpreted across studies. Precision and recall metrics are also frequently reported. Precision is the proportion of classified phonemes that match the correct label, whereas recall is the proportion of phonemes that were detected in the speech stream. While the difference between these two metrics is interesting for theoretical reasons [6], overall accuracy is more relevant concern for the effectiveness of a classification analysis in real world terms. Therefore, in this short report we will indicate classification accuracy and use this metric as an estimate for how analyses performed by different algorithms on different data sets can be evaluated.

## V. RESULTS

This study hypothesized that TMS would (i) improve the classification performance for the corresponding TMS target/phoneme pairs, and (ii) allow for better classification across participants. Two approaches were evaluated: a) individual models, in which model parameters were optimized for each participant under each target condition (sham, lip TMS, or tongue TMS) and for each channel; and b) a generalized model, in which the best overall model parameters were selected across participants and channels for each condition. Data comprised approximately 20 trials per condition, participant, and channel.

**Individual models.** We first identified the optimal model using cross-validation and tested this model on all available data. The best model parameters were selected on all trials for each condition, participant, and channel (80 total) using a 56/24-fold repeated random subsampling cross-validation, repeated 20 times for a one-vs-all classifier for each of the 4 phonemes. This resulted in 6 (participants) \* 3 (conditions) \* 61 (channels) \* 4 (phonemes) = 4392 classifiers.

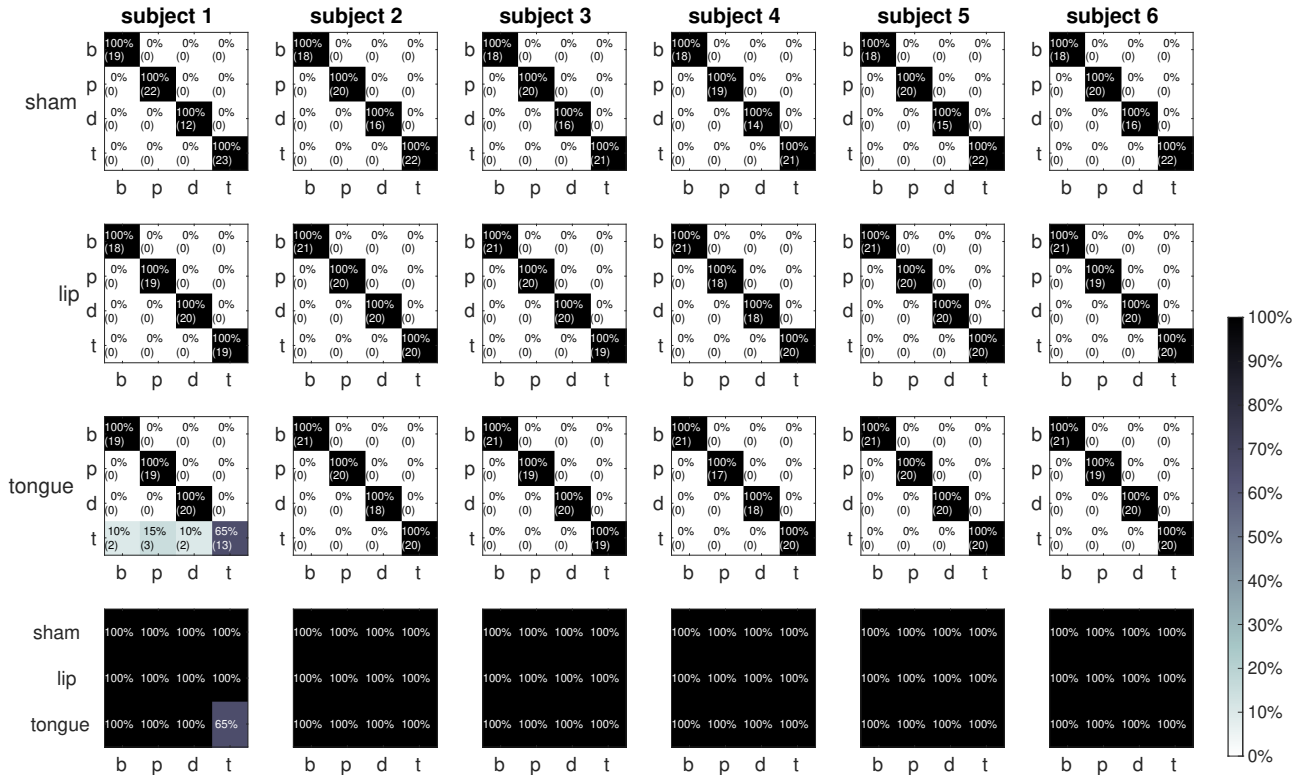


FIG. 3: Classification on all data for individual models for each subject, channel, and condition.

Figure 3 shows modified confusion matrices in the top three rows that correspond to the three conditions and six subjects and the accuracy for each phoneme is summarized in the bottom row. The results illustrate that DDA can achieve essentially perfect classification. Moreover, perfect separation could be found with just one EEG channel, optimized for each sound and participant. Small data sets may be at risk of overfitting; however, with our sparse 3-term model, 20 trials should be sufficient and should not lead to overfitting. In the second analysis, we wanted to confirm this result in a second, more stringent classification task. Because we only had around 20 trials per subject, channel, and condition, we selected the best model on 5 trials from one channel and added 5 trials from a neighboring channel for structure selection. We then tested on the leave-out data from each channel separately.

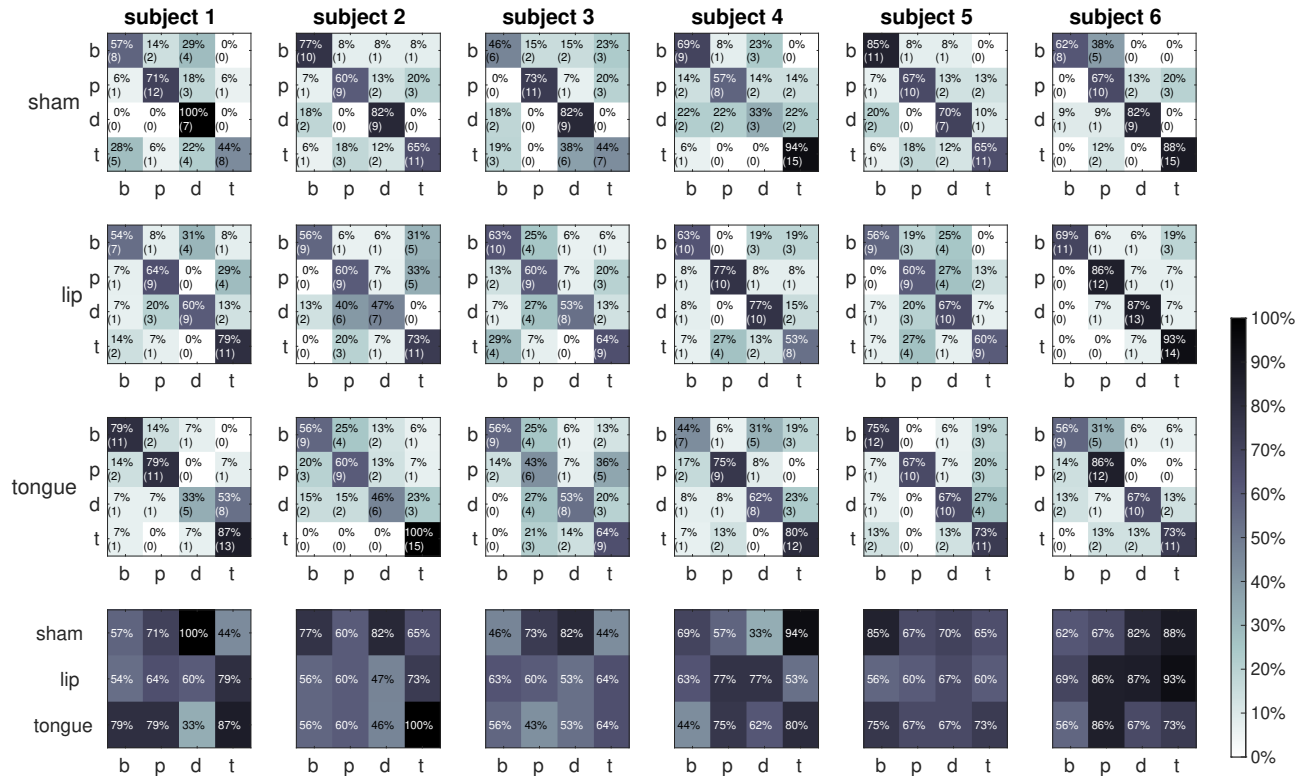


FIG. 4: Classification on leave-out data for individual models for each subject, channel, and condition.

Fig. 4 shows modified confusion matrices in the three top rows and the accuracy for each phoneme in the bottom row. We observe the potential facilitation of classification by TMS, although the sham condition also performs well in select participants. At this stage, we cannot state with certainty whether the hypotheses regarding TMS target/phoneme pairs are upheld. TMS was also theorized to increase the similarity in cortical response across participants. The possible effect of TMS will become more apparent when a generalized model is assessed. Already at this stage of the analysis, significant improvement in classification accuracy relative previous EEG speech decoding studies is observed.

**Generalized model.** The best model parameters were selected on 10 trials from each condition across channels and subjects for each condition and phoneme. In this case, there are four general classifiers (four phonemes) for each condition, but the best channel location depends on the subject. We find that in this more generalized approach, without individualized selection of all model parameters, superior classification is obtained: >71% correct for each phoneme in at least one of the three conditions. In many cases, classification reaches 100% accuracy (see Fig. 5). It is notable that a more generalized approach appears to outperform the fully individualized models. However, this analysis was also performed on a larger data set from all six subjects.

The two TMS conditions taken together tend to provide better classification accuracy than in the sham condition in most participants. This may support the ability of TMS to augment or isolate the neural signal of interest in EEG signals. Given the substantial variation in results across participants, more data is needed to better understand this effect. We argue that discrepancies between participants and in regard to our hypotheses may reflect variation in the optimal target location for each participant: the target coordinates were taken from the literature and transformed to subject-space during the neuronavigation targeting process. Nonetheless, considerable differences in the functional organization of the brain are found between individuals. Targets selected as the average coordinates of multiple subjects may therefore not reflect the cortical regions underlying motor representations for each individual with the same degree of accuracy.

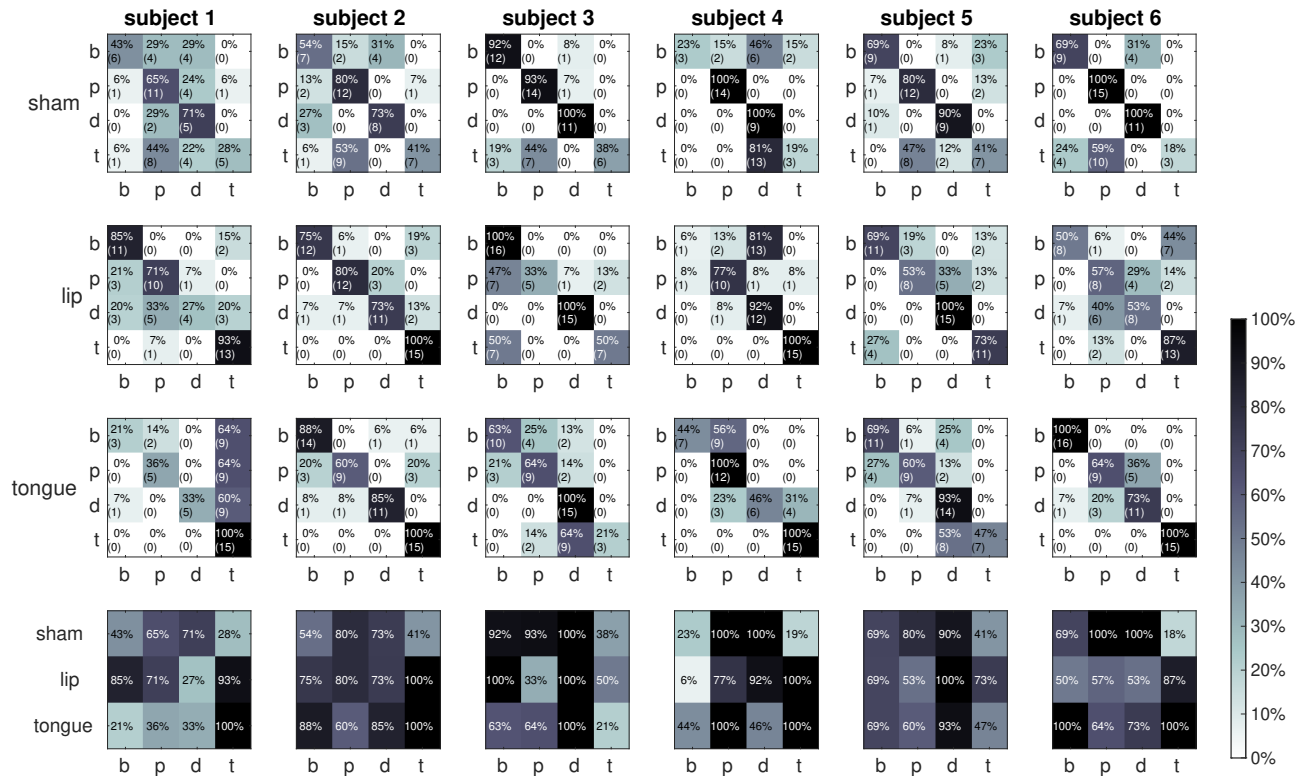


FIG. 5: Classification on leave-out data for generalized models for each condition and phoneme.

However, when we observe the minimum of the four classifiers we find greater support that TMS may create a more uniform neural response to phonemes across subjects. In this analysis, we identified when a phoneme could not be detected. The best classification performance on the leave-out data for each subject, condition, and channel is shown in Figure 6. Plots represent the signal obtained by each electrode. Each plot consists of a matrix of 24 cells: the six rows are the results for each of the six subjects across the 4 phonemes, shown in the columns. The results show clear suppression of /d/ in the lip TMS condition and of /b/ in the tongue TMS condition (i.e., the phoneme not associated with the target). There is greater consistency in the classification results obtained across participants and across all electrode channels in both of the TMS conditions than in the sham condition. Consonants with the articulatory feature of ‘voicing’ (vocal cord vibration) - /b/ and /d/ - show the greatest effect. Voicing requires greater motor activity, and thus this result supports our claim that TMS targets the motor cortex associated with phoneme articulation.

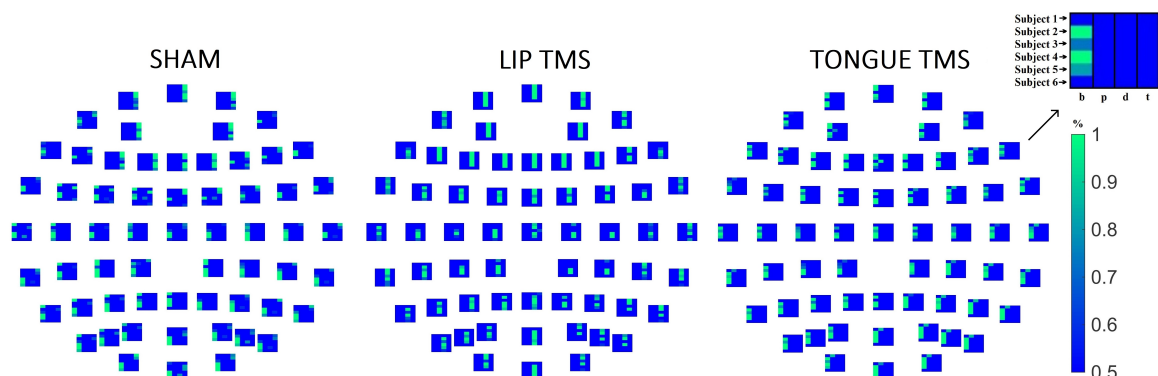


FIG. 6: Classification on leave-out data per electrode channel. Generalized model.

## VI. CONCLUSION

While preliminary in scope, our results clearly indicate the superior capabilities of DDA for speech decoding. Even with the low signal-to-noise ratio of EEG data, the analysis achieved an astonishing 70-100% classification accuracy for all phonemes in at least one of the three conditions when tested on leave-out data. These results have a clear advantage over those obtained by previous EEG speech decoding studies (20%-54%). Moreover, DDA more than doubles classification accuracy while performing a more complex classification task with a minimal data set: in the first two analyses, superior performance was achieved with data from just one EEG channel. DDA even outperformed the metrics reported for studies utilizing invasive data (47%). In this case, although our analysis is less complex, it is performed without the added assistance of predictive language models. There also appears to be good evidence that TMS may assist in creating a consistent cortical response across participants, which in turn may facilitate the development of a generalized model. In select participants, TMS promoted greater classification accuracy. Additional research into subject-specific TMS coordinates will allow for the better understanding of how to harness the effects of TMS for speech decoding.

- 
- [1] Bestmann, S. and Krakauer, J. W. (2015). The uses and interpretations of the motor-evoked potential for understanding behaviour. *Experimental brain research*, 233(3):679–689.
  - [2] Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nature neuroscience*, 20(3):340–352.
  - [3] Casarotto, S., Romero Lauro, L. J., Bellina, V., Casali, A. G., Rosanova, M., Pigorini, A., Defendi, S., Mariotti, M., and Massimini, M. (2010). Eeg responses to tms are sensitive to changes in the perturbation parameters and repeatable over time. *PloS one*, 5(4):e10281.
  - [4] Chartier, J., Anumanchipalli, G. K., Johnson, K., and Chang, E. F. (2018). Encoding of articulatory kinematic trajectories in human speech sensorimotor cortex. *Neuron*, 98(5):1042–1054.
  - [5] Cipollari, S., Veniero, D., Razzano, C., Caltagirone, C., Koch, G., and Marangolo, P. (2015). Combining tms-eeG with transcranial direct current stimulation language treatment in aphasia. *Expert Review of Neurotherapeutics*, 15(7):833–845.
  - [6] Comstock, L., Tankus, A., Tran, M., Pouratian, N., Fried, I., and Speier, W. (2019). Developing a real-time translator from neural signals to text: An articulatory phonetics approach. *Proceedings of the Society for Computation in Linguistics*, 2(1):322–325.
  - [7] D’Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., and Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology*, 19(5):381–385.
  - [8] Fecchio, M., Pigorini, A., Comanducci, A., Sarasso, S., Casarotto, S., Premoli, I., Derchi, C.-C., Mazza, A., Russo, S., Resta, F., et al. (2017). The spectral features of eeg responses to transcranial magnetic stimulation of the primary motor cortex depend on the amplitude of the motor evoked potentials. *PloS one*, 12(9):e0184910.
  - [9] Gorodnitsky, I. and Lainscsek, C. (2004). Machine emotional intelligence: A novel method for spoken affect analysis. In *Proc. Intern. Conf. on Development and Learning ICDL 2004*.
  - [10] Gosseries, O., Thibaut, A., Boly, M., Rosanova, M., Massimini, M., and Laureys, S. (2014). Assessing consciousness in coma and related states using transcranial magnetic stimulation combined with electroencephalography. In *Annales françaises d’anesthésie et de réanimation*, volume 33, pages 65–71. Elsevier.
  - [11] Kremliovskiy, M. and Kadtko, J. (1997). Using delay differential equations as dynamical classifiers. *AIP Conference Proceedings*, 411:57.
  - [12] Lainscsek, C. (2021). Technical report on speech processing using dda. (an unpublished manuscript).
  - [13] Lainscsek, C., Gonzalez, C. E., Sampson, A. L., Cash, S. S., and Sejnowski, T. J. (2019a). Causality detection in cortical seizure dynamics using cross-dynamical delay differential analysis. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(10):101103.
  - [14] Lainscsek, C. and Gorodnitsky, I. (2003). Characterization of various fluids in cylinders from dolphin sonar data in the interval domain. In *Oceans 2003: Celebrating the Past...Teaming Toward the Future*, volume 2, pages 629–632. Marine Technology Society/IEEE.
  - [15] Lainscsek, C., Hernandez, M. E., Weyhenmeyer, J., Sejnowski, T. J., and Poizner, H. (2013a). Non-linear dynamical analysis of EEG time series distinguishes patients with Parkinson’s disease from healthy individuals. *Frontiers in Neurology*, 4(200).
  - [16] Lainscsek, C., Manuel E. Hernandez, M., Poizner, H., and Sejnowski, T. (2013b). Multivariate spectral analysis of electroencephalography data. In *6th Annual International IEEE EMBS Conference on Neural Engineering San Diego, California, 6–8 November, 2013*, page 1151.
  - [17] Lainscsek, C., Messenger, V., Portman, A., Sejnowski, T. J., and Letellier, C. (2014). Automatic sleep scoring from a single electrode using delay differential equations. In Awrejcewicz, J., editor, *Applied Non-Linear Dynamical Systems*, Springer *Proceedings in Mathematics & Statistics*, volume 93, pages 371–382. Springer.
  - [18] Lainscsek, C., Rowat, P., Schettino, L., Lee, D., Song, D., Letellier, C., and Poizner, H. (2012). Finger tapping movements of Parkinson’s disease patients automatically rated using nonlinear delay differential equations. *Chaos*, 22:013119.



- [19] Lainscsek, C., Rungratsameetaeemana, N., Cash, S. S., and Sejnowski, T. J. (2019b). Cortical chimera states predict epileptic seizures. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(12):121106.
- [20] Lainscsek, C., Sampson, A. L., Kim, R., Thomas, M. L., Man, K., Lainscsek, X., Swerdlow, N. R., Braff, D. L., Sejnowski, T. J., Light, G. A., et al. (2019c). Nonlinear dynamics underlying sensory processing dysfunction in schizophrenia. *Proceedings of the National Academy of Sciences*, 116(9):3847–3852.
- [21] Lainscsek, C., Schettino, L., Rowat, P., van Erp, E., Song, D., and Poizner, H. (2009). Nonlinear DDE analysis of repetitive hand movements in Parkinson’s disease. In In, V., Longhini, P., and Palacios, A., editors, *Applications of Nonlinear Dynamics, Understanding Complex Systems*, pages 421–427. Springer.
- [22] Lainscsek, C. and Sejnowski, T. (2013a). Delay differential equation models of electrocardiograms. In *Proceedings of the International Conference on Theory and Applications in Nonlinear Dynamics; Seattle, 2012*.
- [23] Lainscsek, C. and Sejnowski, T. (2013b). Electrocardiogram classification using delay differential equations. *Chaos*, 23(2):023132.
- [24] Lainscsek, C. and Sejnowski, T. (2015). Delay differential analysis of time series. *Neural Computation*, 27(3):594–614.
- [25] Lainscsek, C., Weyhenmeyer, J., Cash, S. S., and Sejnowski, T. J. (2017). Delay differential analysis of seizures in multichannel electrocorticography data. *Neural computation*, 29(12):3181–3218.
- [26] Lainscsek, C., Weyhenmeyer, J., Hernandez, M., Poizner, H., and Sejnowski, T. (2013c). Non-linear dynamical classification of short time series of the Rössler system in high noise regimes. *Frontiers in Neurology*, 4(182).
- [27] Maksimenko, V. A., Pavlov, A., Runnova, A. E., Nedaivozov, V., Grubov, V., Koronovslii, A., Pchelintseva, S. V., Pitsik, E., Pisarchik, A. N., and Hramov, A. E. (2018). Nonlinear analysis of brain activity, associated with motor action and motor imaginary in untrained subjects. *Nonlinear Dynamics*, 91(4):2803–2817.
- [28] Mankin, J. G., Moses, D. A., and Changürrer, E. F. (2020). Machine translation of cortical activity to text with an encoded articulated speech. In *Proc. IEEE Int. Conf. Acoust.*, volume Apr. 2015, pages 992–996. Speech Signal Process. (ICASSP).
- [29] Moses, D. A., Metzger, S. L., Liu, J. R., Anumanchipalli, G. K., Makin, J. G., Sun, P. F., Chartier, J., Dougherty, M. E., Liu, P. M., Abrams, G. M., et al. (2021). Neuroprosthesis for decoding speech in a paralyzed person with anarthria. *New England Journal of Medicine*, 385(3):217–227.
- [30] Mugler, E. M., Patton, J. L., Flint, R. D., Wright, Z. A., Schuele, S. U., Rosenow, J., Shih, J. J., Krusienki, D. J., and Slutzky, M. W. (2014). Direct classification of all american english phonemes using signals from functional speech motor cortex. *Journal of neural engineering*, 11(3):035015.
- [31] Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S. (1980). Geometry from a time series. *Phys. Rev. Lett.*, 45:712.
- [32] Porbadnigk, A., Wester, M., and Jan-p Calliess, T. S. (2009). Eeg-based speech recognition impact of temporal effects.
- [33] Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., and Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, 103(20):7865–7870.
- [34] Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., and Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, 354(6316):1136–1139.
- [35] Rosinová, M., Lojka, M., Staš, J., and Juhár, J. (2017). Voice command recognition using eeg signals. In *2017 International Symposium ELMAR*, pages 153–156. IEEE.
- [36] Saha, P., Abdul-Mageed, M., and Fels, S. (2019). Speak your mind! towards imagined speech recognition with hierarchical deep learning. *arXiv preprint arXiv:1904.05746*.
- [37] Sampson, A. L., Lainscsek, C., Gonzalez, C. E., Ulbert, I., Devinsky, O., Fabó, D., Madsen, J. R., Halgren, E., Cash, S. S., and Sejnowski, T. J. (2019). Delay differential analysis for dynamical sleep spindle detection. *Journal of Neuroscience Methods*, 316:12–21. Methods and models in sleep research: A Tribute to Vincenzo Crunelli.
- [38] Sauer, T., Yorke, J. A., and Casdagli, M. (1991). Embedology. *Journal of Statistical Physics*, 65:579.
- [39] Silvanto, J., Muggleton, N., and Walsh, V. (2008). State-dependency in brain stimulation studies of perception and cognition. *Trends in cognitive sciences*, 12(12):447–454.
- [40] Sulligan, L. S., Klein, E., Brown, T., Sample, M., Pham, M., ..., P. T., and Goeringürrer, S. (2018). Keeping disability in mind: a case study in implantable brain-computer interface research. *Science and engineering ethics*, 24(2):479–504.
- [41] Takens, F. (1981). Detecting strange attractors in turbulence. In Rand, D. A. and Young, L.-S., editors, *Dynamical Systems and Turbulence, Warwick 1980*, volume 898 of *Lecture Notes in Mathematics*, pages 366–381. Springer Berlin/Heidelberg.
- [42] Tremblay, S., Rogasch, N. C., Premoli, I., Blumberger, D. M., Casarotto, S., Chen, R., Di Lazzaro, V., Farzan, F., Ferrarelli, F., Fitzgerald, P. B., et al. (2019). Clinical utility and prospective of tms-eeg. *Clinical Neurophysiology*, 130(5):802–844.
- [43] Veniero, D., Ponzio, V., and Koch, G. (2013). Paired associative stimulation enforces the communication between inter-connected areas. *Journal of Neuroscience*, 33(34):13773–13783.
- [44] Willett, F. R., Avansino, D. T., Hochberg, L. R., Henderson, J. M., and Shenoyürrer, K. V. (2021). High-performance brain-to-text communication via handwriting. *Nature*, 593(7858):249–254.