

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/352194219>

Electroencephalogram (EEG) Based Imagined Speech Decoding and Recognition

Article · June 2021

DOI: 10.31258/Jamt.2.2.74-84

CITATION

1

READS

479

8 authors, including:



[Sani Saminu](#)

Hebei University of Technology

41 PUBLICATIONS 136 CITATIONS

[SEE PROFILE](#)



[Zhang Shuai](#)

Renmin University of China

11 PUBLICATIONS 49 CITATIONS

[SEE PROFILE](#)



[Isselmou Abd El Kader](#)

Hebei University of Technology

28 PUBLICATIONS 107 CITATIONS

[SEE PROFILE](#)



[Adamu Halilu Jabire](#)

Taraba State University

30 PUBLICATIONS 106 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Frequency and Pattern Reconfigurable Metamaterial MIMO Antenna for 5G Applications [View project](#)



Antenna design on millimeter wave band applications [View project](#)

Electroencephalogram (EEG) Based Imagined Speech Decoding and Recognition



Sani Saminu^{1,2,*} , Guizhi Xu¹ , Zhang Shuai¹, Abd El Kader Isselmou¹ , Adamu Halilu Jabire³ , Ibrahim Abdullahi Karaye¹, Isah Salim Ahmad¹ , Abubakar Abdulkarim⁴

¹State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology Tianjin 300130, China

²Biomedical Engineering Department, University of Ilorin, P.M.B. 1515, Ilorin-Nigeria

³Electrical and Electronics Department, Taraba State University, Jalingo-Nigeria

⁴Electrical Engineering Department, Ahmadu Bello University, Zaria-Nigeria

ABSTRACT: The recent investigations and advances in imagined speech decoding and recognition has tremendously improved the decoding of speech directly from brain activity with the help of several neuroimaging techniques that assist us in exploring the neurological processes of imagined speech. This development leads to assist people with disabilities to benefit from neuroprosthetic devices that improve the life of those suffering from neurological disorders. This paper presents the summary of recent progress in decoding imagined speech using Electroencephalography (EEG) signal, as this neuroimaging method enable us to monitor brain activity with high temporal resolution, it is very portable, low cost, and safer as compared to other methods. Therefore, it is a good candidate in investigating an imagined speech decoding from the human cortex which remains a challenging task. The paper also reviews some recent techniques, challenges, future recommendations and possible solutions to improve prosthetic devices and the development of brain computer interface system (BCI).

Key words: EEG, Imagined speech, Decoding, Recognition, BCI, Neuroprosthetic device

■ LIST OF ABBREVIATIONS

ANN, Artificial Neural Network
AR, Auto Regressive
BCI, Brain Computer Interface
CAD, Computer Aided Diagnosis
CD, Correlation Dimension
CAR, Common average reference
CSP, Common Spatial Pattern
CNN, Convolutional Neural Network
DBF, Deep Belief Network
DL, Deep Learning
DWT, Discrete Wavelet Transform
EEG, Electroencephalogram
ECoG, Electrocorticography
ELM, Extreme Learning Machine
fMRI, Functional Magnetic Resonance Imaging
FCM, Fuzzy C-means
FDR, Fisher Discriminant Ratio
FFT, Fast Fourier Transform
GMM, Gaussian Mixer Model
HMM, Hidden Markov Model
iEEG, Intracranial EEG
ICA, Independent Component Analysis
k-NN, k-nearest Neighbor
LDA, Linear Discriminant Analysis

MEG, Magneto Encephalography
MKL, Multiple Kernel Learning
MRF, Markov Random Field
MRI, Magnetic Resonance Imaging
MS, Multiple Sclerosis
NIRS, Near Infrared Spectrum Imaging
NB, Naive Bayes
PCA, Principal Component Analysis
PSD, Power Spectrum Density
PNN, Probabilistic Neural Network
RF, Random Forest
RLS, Recursive Least Squares
RFE, Recursive Feature Elimination
RQA, Recurrence Quantification Analysis
SCP, Slow Cortical Potential
SMR, Sensorimotor Rhythms
SVM, Support Vector Machine
SRS, Simple Random Sampling
SFS, Sequential Feature Selection
SL, Surface Laplacian
VEP, Visual Evoked Potential
WPT, Wavelet Package Transform
WT, Wavelet Transform

Received : March 18, 2021

Revised : May 17, 2021

Accepted : May 25, 2021

1. INTRODUCTION

Neuroimaging techniques have made a significant contribution in decoding a brain physiological phenomena as signals to control a BCI system. These phenomena include P300 evoked potential, slow cortical potential (SCP), visual evoked potential (VEP), and sensorimotor rhythms (SMR) to restore lost verbal communication for people with complete language system but deficit in verbal communication due to disease or injury [1,2]. Loss in verbal communication can be due to neurodegenerative disorganization that influence speech articulation and motor production such as aphasia and its variants [3,4]. In speech comprehension and production, one of the aim of neural prosthetic device is to bring back communication to those affected patients by characterizing the neural activity [5]. Over the years, various neuroimaging methods have been in used and can be classified into invasive and noninvasive techniques. Invasive methods required microelectrode arrays to be implanted inside skull in the brain, as such it involves surgery by expert surgeons in order to obtain a high precision skill. Although this method provides good signal to noise ratio but the formation of a scar tissue over the device due to reaction to the extraneous matter as well as complex surgery that make a permanent hole in the skull limit its application and as such causes a health risk to the patient which may not be worth acceptable [6].

Electrocorticography (ECoG) or intracranial EEG (iEEG) is a partially invasive method in which electrode arrays are implanted over the brain inside the skull, this method overcomes the problem of formation a scar tissue even though the signal strength is weak. Noninvasive techniques are the most widely used with EEG as most commonly acceptable neuroimaging technique. Other non-invasive methods are magneto encephalography (MEG) [7], near infrared spectrum imaging (NIRS), and functional magnetic resonance imaging (fMRI). The advantage of EEG neuroimaging method includes high temporal resolution, it is very portable, low cost, safer or low risks to the users.

Various BCI's application have made it possible for people to directly communicate between the brain and a computer to transfer messages from one's thought to the outside world to enable an individual to perform a non-muscular way of communication as well as to control his surroundings. When we perform a task, the brain generates a signal corresponding to the pattern activity [8, 9]. To explore and identify these patterns is a challenging task and key to the successful BCI system. Over the past decade, there are various BCI techniques that have been developed to assist people with severe communication deficit to restore their communication [10]. These neurotechnological devices ranges from a speller device, virtual key board, moving cursor on the screen to name a few [11,12]. Even though these type of techniques shows promising performance, the need for patient to learn how to adjust their brain activity in an artificial and trained manner such as detecting letters presented on a screen rapidly, rotating a cube, motions in order to operate an interface etc. limits their applications. Therefore, to improve those techniques and to provide other alternatives to those people,

a system which will allow people to communicate more naturally by directly translate or decode inner speech from brain signals is desirable [13,14].

Imagined speech or covert speech is the ability to produce representation of inner speech without any outside speech stimulation and self-generated verbal speech, to understand its underlying mechanism remain a great challenge by researches and also difficult to investigate inner neuronal process because of absence of behavioral output as well as complexity in time-lock exact events with neural activity during imagined speech [15,16]. A lot of effort have been in place to understand neural representation during imagined speech to improve neuroprosthetic devices and to develop various alternative approaches in analyzing neural signal features during imagined speech. Example is the work reported in [17,18] where they used imagined speech to identify different subjects, which proves the variability of EEG signal in performing the same task by different subjects. Also, EEG responses were decoded during imagination and perception of music [19].

This review paper provides a recent progress on decoding and investigating a neural processes, recognition, and monitoring during imagined speech for improving neuroprosthetic assistive communication devices. The paper review studies that have used EEG neuroimaging approach as this method enable us to monitor brain activity with high temporal resolution, it is very portable, low cost, and safer as compared to other approaches. We highlighted various imagined speech decoding and experimental paradigms to explore some challenges that are encountered during covert speech decoding. To develop a natural speech and realistic neuroprostheses device, future directions and new trends in tackling technical challenges have been explored and considered.

The remaining parts of this paper is organized as follows, section two briefly provide the properties of EEG signals, while section three introduce the general decoding model for characterizing brain activity and provides a review of studies that employed EEG signals in decoding imagined speech. Section four highlights future direction, challenges, and possible solutions for a successful and realistic brain machine interface system. Lastly, section five concludes our discussion and findings.

1.1 Imagined Speech Decoding System

This section provides the general overview of imagined speech decoding system. The system consists of the following stages as shown in Fig. 1. 1. Data acquisition 2. Pre-processing 3. Feature extraction 4. Classification and 5. Performance analysis and evaluation.



Fig. 1. General overview of imagined speech decoding and classification system

Preprocessing is a process of artifacts and noise elimination in the recorded EEG data which is contaminated by external and internal factors such as environmental, power line interference, muscle artifacts (EMG), eyes

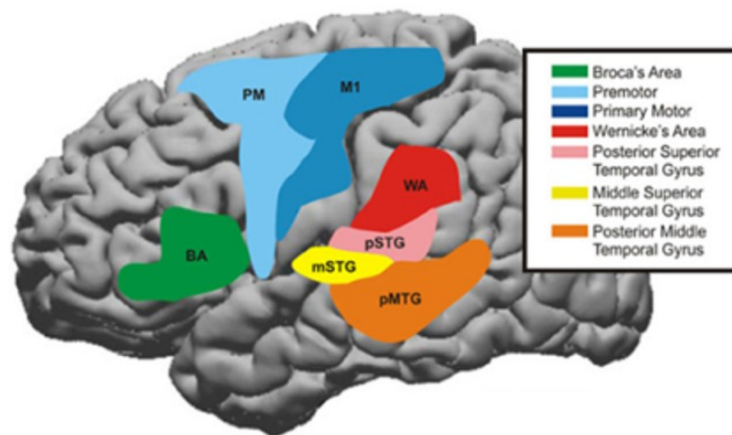


Fig. 3. Functional Brain network areas that involve in production, planning, and perception of speech [32].

blinking (Oculogram) etc. Preprocessing in imagined speech is an important step that influence the classifier performance by removing the unwanted signals. Relevant and significant features are extracted, the most commonly applied method for BCI system are common-spatial patterns [20], autoregressive coefficients [21], and spectro-temporal features [22]. In the classification stage, many classifiers have been employed for decoding imagined speech, the most popular ones are; support vector machine (SVM), linear discriminant analysis (LDA), and random forest (RF) classifiers. Summary of various techniques for preprocessing, feature extraction, and classification with their advantages and disadvantages are provided in section 3 and appendix I, II, and III respectively.

1.2 Data Acquisition

Data acquisition in speech decoding system is a process of collection and recording of neural data from the subjects participated in an experiment. Imagine speech dataset can be recorded either through invasive or noninvasive. Invasive EEG are called Electrocorticography (ECoG) while noninvasive method is by using EEG. This method is commonly applied as it has low risk and safer to the subjects [23]. During the experiments, the subjects have to imagine a set of choosing phonemes, syllables, phrases or sentences while their EEG signals are recorded. The experimental protocol involves pre-trial period in which the subjects were prepared and acquainted with the experimental target. Following the pre-trial, the subjects were asked to imagine the set of stimuli in pre-defined time interval. Example of time-locked experiment conducted in [7] is shown in fig. 2.

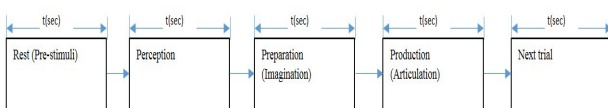


Fig. 2. Example of time-locked experiment in imagined speech decoding

Some datasets used in imagined speech decoding are publically available to be accessed by interested researchers such as [24], while some are acquired upon request from the owners.

1.3 Preprocessing

To improve the efficiency of a classifier and also to reduce its computational complexity, the imagine speech data need to be preprocessed because not all the recorded data that are useful in classification stage as artifacts and noise are induced and contaminate the EEG data during the acquisition process. These artifacts include heart beat artifacts (ECG), eye winks (EOG), muscle movements (EMG), those caused by electrode faults, power line, and interference from equipment's and devices [25-27]. The artifacts related to EEG signals can be divided into two types depending on their source. Those that their source is internally from a biological activities of a body are called as interior artifacts, while those that their source is external are called as exterior artifacts. The summary of types of artifacts are listed in table 1. In addition, the EEG signal has a low signal to noise ratio, therefore it is highly imperative to preprocess the data which involves the process such as down sampling, windowing, and filtering.

Table 1 Type of artifacts in EEG signal

| Interior Artifacts | Exterior Artifacts |
|----------------------------|--|
| Blinking of the eye (EOG.) | Power line |
| Heart beat (ECG.) | Machine fault |
| Muscle movements (EMG.) | Faulty Electrode/Poor Placement |
| Skin resistance | Ventilation |
| Subject's Movement | Digital Artifacts (Loose Wiring, etc.) |

1.4 Feature Extraction

Feature extraction involves extracting relevant and significant features that characterize the neural signals of imagined speech to be fed to the classifier. It is a very important step as it also determines the classifier's accuracy by applying different techniques to ensure non-redundant and most discriminative information in the EEG signals are extracted. Domain approaches that are commonly employed in the feature extraction stage are time-domain, frequency-domain, time-frequency domain, and spatial domain. Other researchers combine two or more approaches in their imagine speech models termed as hybrid

techniques. Details of the feature extraction techniques are discussed in section 3 and appendix II.

1.5 Neural Correlates of Language

Early studies of speech recognition focused on identifying areas responsible for language processing in the human brain which uses patients that undergoing neurosurgery and those of neurological damage [28]. These studies began as early as 1861 by Paul Broca, a neurosurgeon who investigated the human brains using nine patients with lesions and find out that the language areas in humans are located at left hemisphere of the posterior frontal gyrus in the region called Broca's area. Some years after, Carl Wernicke identified an area known as Wernicke's area which he hypothesized that the posterior part of the left temporal lobe is also included in the language comprehension. Several researches have explored the speech and language processing neural regions and developed some functional models and their functional significance [29-31]. These models help in identifying areas that are involve in production, planning, and perception of speech. Fig.3. depicts these areas such as primary motor cortex, pre-motor cortex, Broca's area, Wernicke's area, primary auditory cortex etc.

2. EEG SIGNALS

Fig. 4 shows how the EEG is recorded noninvasively using electrodes placed on the scalp with signals displayed on a computer to depicts the electrical activity of the brain when electrodes detected electrical charges. However, in some specific applications, invasive electrodes can be used termed as intracranial EEG. These electrical recordings from the surface of the brain or even from the outer surface of the head reveals that there is continuous electrical activity in the brain [33]. Based on the state of the brain, the frequencies and amplitudes of the brain signals changes such as during sleep, wakefulness, in a disease state like dementia, epilepsy, sleep disorder, etc. or in mental state. Fig. 5 shows an example of normal EEG signals. EEG signal is measured as the potential difference over time, between the active electrode and the reference electrode. The international 10–20 system accessed from Brain Master Technologies Inc. [34] is shown in Fig. 6. The multichannel EEG sets contain up to 128 or 256 active electrodes. These electrodes are made of silver chloride (AgCl). A gel is used to creates a conductive path between the skin and the electrode for the flow of current. Electrodes that do not use gels, called 'dry' electrodes are made of materials such as titanium and stainless-steel.

2.1 Characteristics of EEG Signals

One of the most important scale in clinical EEGs for evaluating defects and in cognitive research is frequency. A recorded EEG has a frequency somewhere within the 0.01 Hz – 100 Hz range. The frequency content can be divided into five major bands known as delta, theta, alpha, beta, and gamma. Details on the frequencies associated with these bands are provided in Table 2.

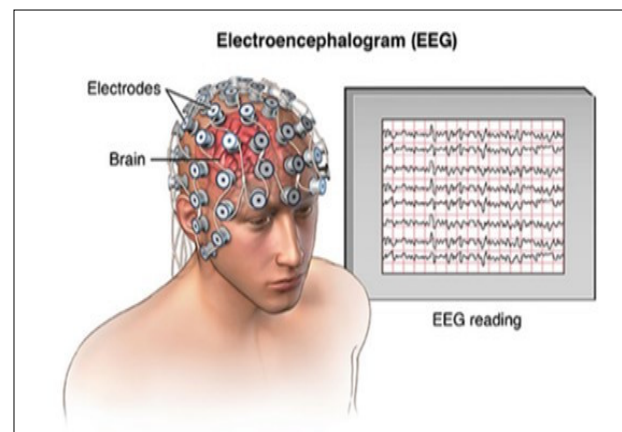


Fig. 4. An illustration of EEG recording

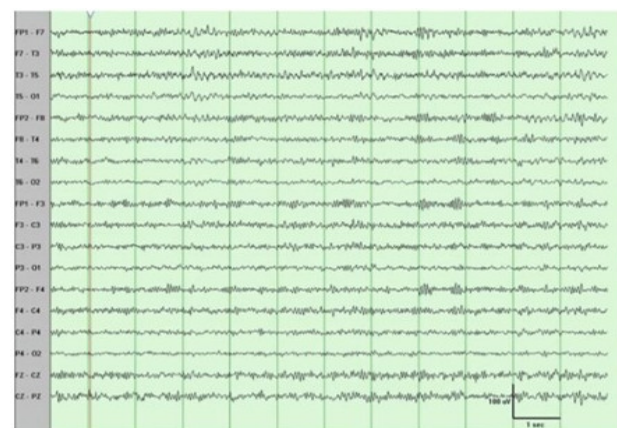


Fig. 5. Normal EEG signal

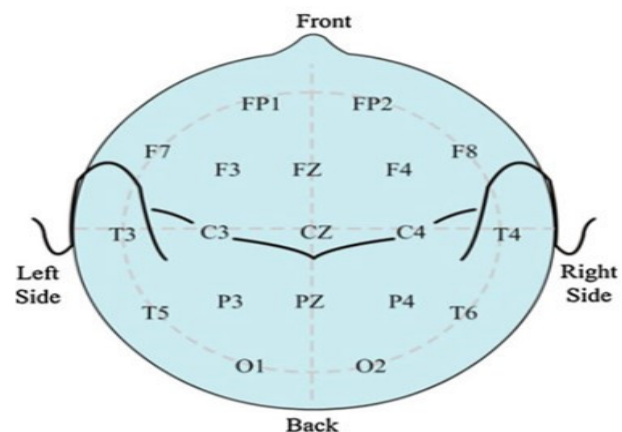


Fig. 6. International 10-20 system of EEG Recording

Table 2 EEG Frequency Bands

| Frequency Band Name | Frequency Bandwidth (Hz) |
|---------------------|--------------------------|
| Alpha | < 4 |
| Beta | 4 - 8 |
| Gamma | 8 - 12 |
| Delta | 12 - 30 |
| Theta | > 30 |

3. GENERAL DECODING MODELS

Sophisticated predictive models are required for targeting BCI application to decode cognitive functions in real time for researchers to use multivariable neural features in complex and rich behavioral conditions [35,36]. A regression framework is widely used modelling method to link neural processes, mental state, and stimulus features. For example, we can model the stimulus features at particular instance as a weighted sum of the neural processes as in equation (1).

$$Y(t) = \sum_p w(p) \cdot X(t,p) \quad (1)$$

Where, $Y(t)$ is the stimulus feature at time t , $w(p)$ is the weight for a given feature p , $X(t,p)$, is the neural processes at time instance t and feature p .

Classification can also be used as a decoding model in which from a finite set of options a neural activity can be recognized as a member to a discrete event type. Several machine learning algorithms can be use by both models such as support vector machine, neural networks, hidden markov models, and simple regression methods among others [37]. Some studies focused on summarizing the perception and imagination of speech and music into various models which relates the neuron's responses with an auditory stimulus. These models include the one proposed by Kaneshiro et al. [38], Geirnaert et al. [39], music perception models were proposed and reported in [40-42].

3.1 Decoding of imagined speech based on EEG

To understand the neural representation of imagined speech from low-level acoustic features to higher-level speech representations, evaluation the relationship between imagined speech stimulus and neural response is a great challenge. In view of that, various studies have demonstrated and highlighted the benefit of EEG recordings to classify imagined speech representations. Early work in this area is the work of [20,43] in which they classified 5 different words using hidden markov model (HMM) classifier. [20,44] used spatial filtering through common spatial filtering in decoding silent vowel speech with Support Vector Machine (SVM). [24] classified vowels using random forest after down sampling the data and reported an accuracy of 22.32%. Discrimination of imagined speech in EEG was proposed in [45] using Tensor decomposition. Hilbert transform and Hilbert spectrum methods were used by [46], [47], and [48] to decode imagined speech using EEG signal, both studies used two different syllables during the experiment but with four and seven subjects respectively. Wavelet transform was used for feature extraction with alternated least squares approximation and down sample the data for vowel classification, they obtained the accuracy of 59.70% using SVM classifier [45]. Multi-class classification of words was proposed in [49] using connectivity features.

Imagined speech was used for subject recognition using auto regressive (AR) coefficients with k-Nearest Neighbor (k-NN) and linear SVM was performed in [50-52]. EEG signals was used to characterize 10 English language imagined phonemes [52 53] with Naïve Bayesian and Linear Discriminant Analysis for feature extraction and

classification. Decoding Chinese characters based on EEG speech imagery was proposed by [54, 55], common spatial patterns (CSP) and SVM approach was employed for preprocessing and classification respectively. Japanese vowels were classified through EEG recordings using SVM [56]. Imagined speech classification based on Riemannian distance of correntropy spectral density was proposed in [57]. Word classification was performed by [58 59], English vowels [60-64], phonemic decoding [22], Spanish vowels [65] with different approaches in preprocessing, feature extraction and classification. In [66], extreme learning machine (ELM) was trained and tested on a raw EEG data classification and compared with several machine learning classifiers. ELM shows a promising result by outperforming other machine learning classifiers. Marthe et al. [19] proposed imagined music decoding and recognition model and hypothesed that the same linear neural decoding models used in imagined speech decoding can also be used to decode imagined music. Their proposed model achieved an accuracy of 69.2%.

Recent studies in the area of imagined speech through EEG recordings are tilting their attention towards new methods and recent machine learning techniques such as deep learning [67, 68]. These include a classification of imagined speech using regularized neural network [69], using Artificial Neural Network (ANN) to classify bilingual unspoken speech [70], and online classification of imagined speech for BCI based on EEG signals. Machine learning algorithms were applied to analyzed the similarity and differences among perception, production and imagination using EEG patterns [71]. Deep learning approach have been applied in other areas such vision recognition, image processing, and motor imagery for many years, but only recently researches begin to explore the benefit of this technique in imagined speech processing, decoding and analysis from EEG [72-75]. Convolutional Neural Network (CNN) algorithm with both deep and shallow architecture was experimented to classify word pairs of the EEG dataset with an average accuracy of 62.37% and 60.88% for the deep and shallow CNNs, respectively [66]. In another study, five vowels were classified after down sampling the data to 128 Hz, Independent Component Analysis (ICA) with Hessian approximation for artefact removal was deployed in preprocessing the data. Classification was performed using deep CNN with 32 layers. Five main vowels (a, e, i, o, u) and six different words were classified by Tamm et al. [76]. They proposed a low in computation with few number of layers using CNN model and achieved the accuracy of 23.98%. To improve CNN model, an optimized structure was proposed by Cooney et al. [77] by optimizing input layers to decode imagined speech using transfer learning.

4. FUTURE CHALLENGES

For the successful implementation of BCI system using imagined speech based on EEG recordings, the future research direction must focus on key and challenging step which is to apply the different levels of speech processing and representation to imagined speech due to the absent of behavioral output and difficulty in monitoring the spectrotemporal structure. To time-lock a brain processes to

a behavioral state or a quantifiable stimulus is complex task as an experimenter cannot directly monitor the imagination. Therefore, some standards techniques/models that matched input-output data has to be employed. Also, several factors such as age, emotion, gender, dialect, and pronunciation affects natural speech which results in temporal irregularities. Other challenges explored from our study with their remedy include those associated with how to design a good experimental task, proper training of participants, generating sufficient amount of data, using effective and good electrodes as well as improve their design, employing unsupervised machine learning techniques among others. Recently, most of the researchers focused their attention on applying deep learning algorithm in decoding and recognition of imagined speech, therefore huge amount of data must be generated to train the network effectively. Also, the requirement of higher computational resources in deep learning models must be addressed to enable other researches to experiment and evaluate their models to realize precise, practical, and reliable non-invasive models. Finally, homogeneous performance comparison among the developed techniques is difficult as there is lack of standardization for evaluating their performance due to

different datasets with different sampling frequency, number of electrodes and other parameters.

5. CONCLUSIONS

In conclusion, this paper highlights the progress and explored the potential of using various decoding models and algorithms through collection of studies that used EEG signals recording to identified neural mechanisms related to complex speech production and functions. These models are also capable of recognizing and characterizing the important components of natural communication, that is speech and language perception and production directly from brain activity. Several studies reviewed in this paper reveals a promising result for classification of either phonemes, vowels and words from EEG brain signals but still shows that a lot of work need to be conducted to provide a realistic and efficient brain machine interface and neuroprosthetic devices. Recent machine learning techniques such as deep learning need to be further investigated, improvements in areas of experimental paradigm is paramount for better recognition, preprocessing, and feature extraction as they are the key aspects in development of new communication interfaces.

■ APPENDIX

Appendix 1 Summary of Performance Comparison of Preprocessing methods

| Method | Advantages | Disadvantages |
|--------------------|---|---|
| ICA | <ul style="list-style-type: none"> ◆ High calculation efficiency ◆ Works well for large data sized | <ul style="list-style-type: none"> ◆ Require more computational for decomposition |
| PCA | <ul style="list-style-type: none"> ◆ Reduction of feature dimension | <ul style="list-style-type: none"> ◆ Not as good as ICA |
| CAR | <ul style="list-style-type: none"> ◆ The best among all the reference methods | <ul style="list-style-type: none"> ◆ Needs sufficient head coverage |
| CSP | <ul style="list-style-type: none"> ◆ Good effect on EEG data processing based on moto imagination ◆ does not need to select a specific frequency band in advance. | <ul style="list-style-type: none"> ◆ sensitive to noise ◆ requires multiple electrodes |
| SL | <ul style="list-style-type: none"> ◆ Robust against the Artifacts | <ul style="list-style-type: none"> ◆ Sensitive to spline pattern ◆ Sensitive to Artifacts |
| Digital filtering | <ul style="list-style-type: none"> ◆ Easily remove electrical grounding noise | <ul style="list-style-type: none"> ◆ Needs signal and noise reside distinct frequency bands |
| Adaptive filtering | <ul style="list-style-type: none"> ◆ Good for the signals with overlapping spectra | <ul style="list-style-type: none"> ◆ Needs reference signal |

Appendix 2 Summary of Performance Comparison of Feature extraction methods

| Method | Advantages | Disadvantages |
|--------|--|--|
| PSD | <ul style="list-style-type: none"> ◆ Feature stability efficiency ◆ Works well for large data sized | <ul style="list-style-type: none"> ◆ Not suitable for unstable signals ◆ Unable to analyze time-domain signals |
| PCA | <ul style="list-style-type: none"> ◆ Lossless dimension reduction | <ul style="list-style-type: none"> ◆ Weak to handle complex set of data |
| ICA | <ul style="list-style-type: none"> ◆ fast calculation speed ◆ high efficiency ◆ Needs small duration of data | <ul style="list-style-type: none"> ◆ Needs more calculating for decomposition ◆ Not available for stationary signals |
| AR | <ul style="list-style-type: none"> ◆ Decrease spectra loss issues ◆ Gives well frequency resolution ◆ Applicable for stationary signals | <ul style="list-style-type: none"> ◆ Presents low performance once the estimated model is not suitable and its orders are mistakenly elected ◆ Not works for nonstationary signals |
| FFT | <ul style="list-style-type: none"> ◆ It is appropriate for narrowband signals, such as sine wave ◆ Works well for nonstationary signals | <ul style="list-style-type: none"> ◆ Sustains greatly noise sensitivity and does not have shorter duration data record |
| WT | <ul style="list-style-type: none"> ◆ Can analyze signal both in time and frequency domain ◆ Works well to analyze the nonstationary | <ul style="list-style-type: none"> ◆ Deficient particular mode to apply to prevalent noise |
| WPT | <ul style="list-style-type: none"> ◆ provide precise frequency resolution | <ul style="list-style-type: none"> ◆ Raise calculating time |

Appendix 3 Summary of Performance Comparison of Classification methods

| Method | Advantages | Disadvantages |
|--------|---|---|
| ANN | <ul style="list-style-type: none"> ◆ Handling multi-dimensional and multi-variety data ◆ High accuracy ◆ flexible structure ◆ Easily identifies trends and patterns | <ul style="list-style-type: none"> ◆ Requires large data sets to train the network ◆ Performance depends on the number of neurons in the hidden layer ◆ needs massive resources to function |
| k-NN | <ul style="list-style-type: none"> ◆ Easy to understand, ◆ simple to implement | <ul style="list-style-type: none"> ◆ High error-susceptibility ◆ Sensitive to irrelevant and redundant features ◆ requires large storage space |
| LDA | <ul style="list-style-type: none"> ◆ Easy to use ◆ Low computational complexity | <ul style="list-style-type: none"> ◆ Requires a linear model |
| SVM | <ul style="list-style-type: none"> ◆ works relatively well when there is a clear margin of separation between classes ◆ more effective in high dimensional spaces ◆ relatively memory efficient | <ul style="list-style-type: none"> ◆ not suitable for large data sets ◆ does not perform very well when the data set has more noise |
| NB | <ul style="list-style-type: none"> ◆ Easy to understand | <ul style="list-style-type: none"> ◆ Independent variables |
| DL | <ul style="list-style-type: none"> ◆ Features are automatically deduced and optimally tuned for desired outcome ◆ Robustness to natural variations in the data is automatically learned ◆ Massive parallel computations can be performed using GPUs and are scalable for large volumes of data ◆ flexible to be adapted to new problems in the future | <ul style="list-style-type: none"> ◆ It requires very large amount of data in order to perform better than other techniques ◆ It is extremely expensive to train due to complex data models ◆ difficult to be adopted by less skilled people |

AUTHOR INFORMATION

Corresponding Author

*Email: sansam4k@gmail.com

ORCID

Sani Saminu : 0000-0002-5182-7150
 Guizhi Xu : 0000-0002-9637-0051
 Abd El Kader Isselmou : 0000-0002-0140-0749
 Adamu Halilu Jabire : 0000-0002-9955-7946
 Isah Salim Ahmad : 0000-0002-5634-1627
 Abubakar Abdulkarim : 0000-0003-0303-6327

REFERENCES

- [1] Palaniappan R.. Utilizing Gamma Band to Improve Mental Task Based Brain-Computer Interface Design. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2006; 14(3)
- [2] Nicolas-Alonso Luis Fernando, Gomez-Gil Jaime. Brain Computer Interfaces, a Review. *Sensors*. 2012 ; 12(2)
- [3] Ajiboye A Bolu, Willett Francis R, Young Daniel R, Memberg William D, Murphy Brian A, Miller Jonathan P, Walter Benjamin L, Sweet Jennifer A, Hoyer Harry A, Keith Michael W, Peckham P Hunter, Simeral John D, Donoghue John P, Hochberg Leigh R, Kirsch Robert F. Restoration of reaching and grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: a proof-of-concept demonstration. *The Lancet*. 2017; 389(10081)
- [4] Gilja Vikash, Pandarinath Chethan, Blabe Christine H, Nuyujukian Paul, Simeral John D, Sarma Anish A, Soric Brittany L, Perge János A, Jarosiewicz Beata, Hochberg Leigh R, Shenoy Krishna V, Henderson Jaimie M. Clinical translation of a high-performance neural prosthesis. *Nature Medicine*. 2015; 21(10)
- [5] Mustaqem Mustaqem, Kwon Soonil. CLSTM: Deep Feature-Based Speech Emotion Recognition Using the Hierarchical ConvLSTM Network. *Mathematics*. 2020; 8(12):1-19
- [6] Rashid Mamunur, Sulaiman Norizam, P. P. Abdul Majeed Anwar, Musa Rabi Muazu, Ab. Nasir Ahmad Fakhri, Bari Bifta Sama, Khatun Sabira. Current Status, Challenges, and Possible Solutions of EEG-Based Brain-Computer Interface: A Comprehensive Review. *Frontiers in Neurobotics*. 2020; 14
- [7] Dash Debadatta, Ferrari Paul, Wang Jun. Decoding Imagined and Spoken Phrases From Non-invasive Neural (MEG) Signals. *Frontiers in Neuroscience*. 2020; 14
- [8] Anumanchipalli Gopala K., Chartier Josh, Chang Edward F.. Speech synthesis from neural decoding of spoken sentences. *Nature*. 2019; 568(7753)
- [9] Luck Steven J.. An Introduction to the Event-Related Potential Technique, Second Edition. The MIT Press: Cambridge, MA, USA; 2014
- [10] Bashashati Ali, Fatourech Mehrdad, Ward Rabab K, Birch Gary E. A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. *Journal of Neural Engineering*. 2007; 4(2)
- [11] Vansteensel Mariska J., Pels Elmar G.M., Bleichner Martin G., Branco Mariana P., Denison Timothy, Freudenburg Zachary V., Gosselaar Peter, Leinders Sacha, Ottens Thomas H., Van Den Boom Max A., Van Rijen Peter C., Aarnoutse Erik J., Ramsey Nick F.. Fully Implanted Brain-Computer Interface in a Locked-In Patient with ALS. *New England Journal of Medicine*. 2016; 375(21)
- [12] Pandarinath Chethan, Nuyujukian Paul, Blabe Christine H, Soric Brittany L, Saab Jad, Willett Francis R, Hochberg Leigh R, Shenoy Krishna V, Henderson Jaimie M. High performance communication by people with paralysis using an intracortical brain-computer interface. *eLife*. 2017; 6
- [13] Xie Yu, Oniga Stefan. A Review of Processing Methods and Classification Algorithm for EEG Signal. *Carpathian Journal of Electronic and Computer Engineering*. 2020; 13(1)
- [14] Sharon Rini A., Narayanan Shrikanth, Sur Mriganka, Murthy Hema A.. An Empirical Study of Speech Processing in the Brain by Analyzing the Temporal Syllable Structure in Speech-input Induced EEG. *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2019.
- [15] Démonet Jean-François, Thierry Guillaume, Cardebat Dominique. Renewal of the Neurophysiology of Language: Functional Neuroimaging. *Physiological Reviews*. 2005; 85(1)
- [16] Hickok Gregory, Poeppel David. The cortical organization of speech processing. *Nature Reviews Neuroscience*. 2007; 8(5)
- [17] Moctezuma Luis Alfredo, Torres-García Alejandro A., Villaseñor-Pineda Luis, Carrillo Maya. Subjects identification using EEG-recorded imagined speech. *Expert Systems with Applications*. 2019; 118
- [18] Sharon Rini A, Murthy Hema A. Comparison of Feature-Model Variants for coSpeech-EEG Classification. *2020 National Conference on Communications (NCC)*. 2020.
- [19] Tibo Marthe. Decoding EEG responses during perception and imagination of music. 2020.
- [20] DaSalla Charles S., Kambara Hiroyuki, Sato Makoto, Koike Yasuharu. Single-trial classification of vowel speech imagery using common spatial patterns. *Neural Networks*. 2009; 22(9)
- [21] Song YoungJae, Sepulveda Francisco. Classifying speech related vs. idle state towards onset detection in brain-computer interfaces overt, inhibited overt, and covert speech sound production vs. idle state. *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings*. 2014.
- [22] Zhao Shunan, Rudzicz Frank. Classifying phonological categories in imagined and articulated speech. *2015 IEEE International Conference on Acoustics, Speech and Signal Processing*

- (ICASSP). 2015.
- [23] Bogue Robert. Brain-computer interfaces: control by thought. *Industrial Robot: An International Journal*. 2010; 37(2)
- [24] Pressel Coretto Germán A., Gareis Iván E., Rufiner H. Leonardo. Open access database of EEG signals recorded during imagined speech. *12th International Symposium on Medical Information Processing and Analysis*. 2017.
- [25] Sweeney K. T., Ward T. E., McLoone S. F.. Artifact Removal in Physiological Signals—Practices and Possibilities. *IEEE Transactions on Information Technology in Biomedicine*. 2012; 16(3)
- [26] EEG Signal Processing and Feature Extraction. Springer: Singapore; 2019.
- [27] Islam Md Kafiul, Rastegarnia Amir, Yang Zhi. Methods for artifact detection and removal from scalp EEG: A review. *Neurophysiologie Clinique/ Clinical Neurophysiology*. 2016; 46(4-5)
- [28] Ojemann GA. Cortical organization of language. *The Journal of Neuroscience*. 1991; 11(8)
- [29] Hickok Gregory, Poeppel David. The cortical organization of speech processing. *Nature Reviews Neuroscience*. 2007; 8:393-402.
- [30] Hickok Gregory. Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*. 2012; 13:135-145.
- [31] Herff Christian, Diener Lorenz, Angrick Miguel, Mugler Emily, Tate Matthew C., Goldrick Matthew A., Krusienski Dean J., Slutzky Marc W., Schultz Tanja. Generating Natural, Intelligible Speech From Brain Activity in Motor, Premotor, and Inferior Frontal Cortices. *Frontiers in Neuroscience*. 2019; 13
- [32] Leuthardt Eric C., Pei Xiao-Mei, Breshears Jonathan, Gaona Charles, Sharma Mohit, Freudenberger Zac, Barbour Dennis, Schalk Gerwin. Temporal evolution of gamma activity in human cortex during an overt and covert word repetition task. *Frontiers in Human Neuroscience*. 2012; 6
- [33] Azar Ahmad Taher, Balas Valentina E., Olariu Teodora. Classification of EEG-Based Brain-Computer Interfaces. *Advanced Intelligent Computational Technologies and Decision Support Systems*. Springer, Cham; 2013:97-106.
- [34] Inc. Brain Master Technologies. The international 10–20 system.
- [35] Kay Kendrick N, Gallant Jack L. I can see what you see. *Nature Neuroscience*. 2009; 12(3)
- [36] Geirnaert Simon, Francart Tom, Bertrand Alexander. An Interpretable Performance Metric for Auditory Attention Decoding Algorithms in a Context of Neuro-Steered Gain Control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2020; 28(1)
- [37] Naselaris Thomas, Kay Kendrick N., Nishimoto Shinji, Gallant Jack L.. Encoding and decoding in fMRI. *NeuroImage*. 2011; 56(2)
- [38] Kaneshiro Blair, Nguyen Duc T., Norcia Anthony M., Dmochowski Jacek P., Berger Jonathan. Natural music evokes correlated EEG responses reflecting temporal structure and beat. *NeuroImage*. 2020; 214
- [39] Geirnaert Simon, Vandecappelle Servaas, Alickovic Emina, de Cheveigné Alain, Lalor Edmund, Meyer Bernd T., Miran Sina, Francart Tom, Bertrand Alexander. Neuro-Steered Hearing Devices: Decoding Auditory Attention From the Brain. 2020.
- [40] Cantisani Giorgia, Essid Slim, Richard Gael. EEG-Based Decoding of Auditory Attention to a Target Instrument in Polyphonic Music. *2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*. 2019.
- [41] Di Liberto Giovanni M., Pelofi Claire, Shamma Shihab, de Cheveigné Alain. Musical expertise enhances the cortical tracking of the acoustic envelope during naturalistic music listening. *Acoustical Science and Technology*. 2020; 41(1)
- [42] Treder M S, Purwins H, Miklody D, Sturm I, Blankertz B. Decoding auditory attention to instruments in polyphonic music using single-trial EEG classification. *Journal of Neural Engineering*. 2014; 11(2)
- [43] Porbadnigk Anne, Wester Marek, Callies Jan Peter, Schultz Tanja. EEG-based Speech Recognition - Impact of Temporal Effects. 2009.
- [44] DaSalla Charles S., Kambara Hiroyuki, Koike Yasuharu, Sato Makoto. Spatial filtering and single-trial classification of EEG during vowel speech imagery. *Proceedings of the 3rd International Convention on Rehabilitation Engineering & Assistive Technology - ICREATE '09*. 2009.
- [45] García-Salinas Jesús S., Villaseñor-Pineda Luis, Reyes-García Carlos Alberto, Torres-García Alejandro. Tensor Decomposition for Imagined Speech Discrimination in EEG. In: Batyrshin Ildar, Martínez-Villaseñor María de Lourdes, Espinosa Hiram Eredín Ponce. *Advances in Computational Intelligence. MICAI 2018. Lecture Notes in Computer Science*. Springer, Cham; 2019:239-249.
- [46] D'Zmura Michael, Deng Siyi, Lappas Tom, Thorpe Samuel, Srinivasan Ramesh. Toward EEG Sensing of Imagined Speech. *Human-Computer Interaction. New Trends*. 2009.
- [47] Moses David A., Leonard Matthew K., Makin Joseph G., Chang Edward F.. Real-time decoding of question-and-answer speech dialogue using human cortical activity. *Nature Communications*. 2019; 10(1)
- [48] Deng Siyi, Srinivasan Ramesh, Lappas Tom, D'Zmura Michael. EEG classification of imagined syllable rhythm using Hilbert spectrum methods. *Journal of Neural Engineering*. 2010; 7(4)
- [49] Qureshi Muhammad Naveed Iqbal, Min Beomjun, Park Hyeon-Jun, Cho Dongrae, Choi Woosu, Lee Boreom. Multiclass Classification of Word Imagination Speech With Hybrid Connectivity Features. *IEEE Transactions on Biomedical*

- Engineering*. 2018; 65(10)
- [50] Brigham Katharine, Kumar B. V. K. Vijaya. Subject identification from electroencephalogram (EEG) signals during imagined speech. *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. 2010.
- [51] Brigham Katharine, Kumar B. V. K. Vijaya. Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy. *2010 4th International Conference on Bioinformatics and Biomedical Engineering*. 2010.
- [52] AlSaleh Mashael, Moore Roger, Christensen Heidi, Arvaneh Mahnaz. Discriminating Between Imagined Speech and Non-Speech Tasks Using EEG. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2018.
- [53] Chi Xuemin, Hagedorn John B., Schoonover Daniel, D'Zmura Michael. EEG-Based Discrimination of Imagined Speech Phonemes. *International Journal of Bioelectromagnetism*. 2011; 13(4):201-206.
- [54] Wang Li, Zhang Xiong, Zhong Xuefei, Zhang Yu. Analysis and classification of speech imagery EEG for BCI. *Biomedical Signal Processing and Control*. 2013; 8(6)
- [55] Guo M.M., Qi Z.G., Wang L., Xu G.. Research on parameter optimization in speech rehabilitation system based on Brain Computer Interface. *Journal of Signal Processing*. 2018; 34(8):975-983.
- [56] Matsumoto Mariko. Silent speech decoder using adaptive collection. *Proceedings of the companion publication of the 19th international conference on Intelligent User Interfaces - IUI Companion '14*. 2014.
- [57] Bakhshali Mohamad Amin, Khademi Morteza, Ebrahimi-Moghadam Abbas, Moghimi Sahar. EEG signal classification of imagined speech based on Riemannian distance of correntropy spectral density. *Biomedical Signal Processing and Control*. 2020; 59
- [58] Arafat S Kazi Yaser, Kanade Sudhir S. Imagined Speech EEG Signal Processing For Brain Computer Interface. *International Journal of Application or Innovation in Engineering & Management*. 2014; 3 (7):123-126.
- [59] Mohanchandra Kusuma, Saha Snehanthu. A Communication Paradigm Using Subvocalized Speech: Translating Brain Signals into Speech. *Augmented Human Research*. 2016; 1(1)
- [60] Kamalakkannan R., Rajkumar R., Raj M. Madan, Shenbaga S.D.. Imagined speech classification using EEG. *Advances in Biomedical Science and Engineering*. 2014; 1(2):20-32.
- [61] Iqbal Sadaf, Uzzaman Khan Yusuf, Farooq Omar. EEG based classification of imagined vowel sounds. 2015.
- [62] Idrees Basil M., Farooq Omar. EEG based vowel classification during speech imagery. 2016.
- [63] Idrees Basil M., Farooq Omar. Vowel classification using wavelet decomposition during speech imagery. *2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN)*. 2016.
- [64] Min Beomjun, Kim Jongin, Park Hyeong-jun, Lee Boreom. Vowel Imagery Decoding toward Silent Speech BCI Using Extreme Learning Machine with Electroencephalogram. *BioMed Research International*. 2016; 2016
- [65] Rojas Diego A., Ramos Olga L., Jorge E. Saby. Recognition of Spanish Vowels through Imagined Speech by using Spectral Analysis and SVM. *Journal of Information Hiding and Multimedia Signal Processing*. 2016; 7(4):889-897.
- [66] Tan Ping, Sa Weiping, Yu Lingli. Applying Extreme Learning Machine to classification of EEG BCI. *2016 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*. 2016.
- [67] Tamm Markus-Oliver. Vowel Classification from Imagined Speech Using Machine Learning. 2020.
- [68] Patel Jigar, Pasha I.A., Krishna D. Hari. Classification of imagery vowel speech using EEG and cross correlation. *International Journal of Pure and Applied Mathematics*. 2018; 118(24):1-10.
- [69] Rezazadeh Sereshkeh Alborz, Trott Robert, Bricout Aurelien, Chau Tom. EEG Classification of Covert Speech Using Regularized Neural Networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 2017; 25(12)
- [70] Balaji Advait, Haldar Aparajita, Patil Keshav, Ruthvik T Sai, CA Valliappan, Jartarkar Mayur, Baths Veeky. EEG-based classification of bilingual unspoken speech using ANN. *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2017.
- [71] Sharon Rini A., Narayanan Shrikanth S., Sur Mriganka, Murthy A. Hema. Neural Speech Decoding During Audition, Imagination and Production. *IEEE Access*. 2020; 8
- [72] Saha Pramit, Fels Sidney. Hierarchical Deep Feature Learning for Decoding Imagined Speech from EEG. 2019.
- [73] Panachakel Jerrin Thomas, Ramakrishnan A.G., Ananthapadmanabha T.V.. A Novel Deep Learning Architecture for Decoding Imagined Speech from EEG. 2019.
- [74] Cooney Ciaran , Korik Attila , Folli Raffaella , Coyle Damien H . Classification of Imagined Spoken Word -Pairs Using Convolutional Neural Networks. 2019.
- [75] Roy Yannick, Banville Hubert, Albuquerque Isabela, Gramfort Alexandre, Falk Tiago H, Faubert Jocelyn. Deep learning-based electroencephalography analysis: a systematic review. *Journal of Neural Engineering*. 2019; 16(5)
- [76] Tamm Markus-Oliver, Muhammad Yar, Muhammad Naveed. Classification of Vowels from Imagined Speech with Convolutional Neural Networks. *Computers*. 2020; 9(2)

- [77] Cooney Ciaran, Folli Raffaella, Coyle Damien. Optimizing Layers Improves CNN Generalization and Transfer Learning for Imagined Speech Decoding from EEG. *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. 2019.



This article is licensed under a [Creative Commons Attribution 4.0 International License](#).