1-1-2015

Electroencephalogram Based Causality Graph Analysis in Behavior Tasks of Parkinson's Disease Patients

Abdulaziz Saleh Almalaq
University of Denver

Electroencephalogram based Causality Graph Analysis in Behavior Tasks of Parkinsons Disease Patients

A Thesis

Presented to

the Faculty of the Daniel Felix Ritchie School of Engineering and Computer Science

University of Denver

in Partial Fulfillment

of the Requirements for the Degree

Master Of Science

by

Abdulaziz S. Almalaq

August 2015

Advisor: Dr. Jun Zhang

Author: Abdulaziz S. Almalaq
Title: Electroencephalogram based Causality Graph Analysis in Behavior Tasks of Parkinsons Disease Patients
Advisor: Dr. Jun Zhang
Degree Date: August 2015

Abstract

Electroencephalographic (EEG) signals of the human brains represent electrical activities for a number of channels recorded over a the scalp. The main purpose of this thesis is to investigate the interactions and causality of different parts of a brain using EEG signals recorded during a performance subjects of verbal fluency tasks. Subjects who have Parkinson's Disease (PD) have difficulties with mental tasks, such as switching between one behavior task and another. The behavior tasks include phonemic fluency, semantic fluency, category semantic fluency and reading fluency. This method uses verbal generation skills, activating different Broca's areas of the Brodmann's areas (BA44 and BA45). Advanced signal processing techniques are used in order to determine the activated frequency bands in the granger causality for verbal fluency tasks. The graph learning technique for channel strength is used to characterize the complex graph of Granger causality. Also, the support vector machine (SVM) method is used for training a classifier between two subjects with PD and two healthy controls. Neural data from the study was recorded at the Colorado Neurological Institute (CNI). The study reveals significant difference between PD subjects and healthy controls in terms of brain connectivities in the Broca's Area BA44 and BA45 corresponding to EEG electrodes. The results in this thesis also demonstrate the possibility to classify based on the flow of information and causality in the brain of verbal fluency tasks. These methods have the potential to be applied in the future to identify pathological information flow and causality of neurological diseases.

Acknowledgements

I would like to take the opportunity to express my genuine gratitude and appreciation to my academic advisor, Dr. Jason Jun Zhang, who has continually given me his suggestions and guidance. Without his support and persistent help in regard to my studies and research, this thesis would not have been possible. Also, I wish to thank the committee members Dr. Caroline Li, Dr. Ronald Delyser, and Dr. Adam Hebb, who have spent their precious time to review my research and provide critical analysis of my work. In addition, I am very thankful for the contribution of Sara Hanrahan, PhD who assisted me in understanding the basics of EEG signals and many neurological aspects of the brain. Also, I would like to thank the Colorado Neurological Institute (CNI) for the contribution of providing me datasets for this thesis. Lastly, I owe a very important debt to Xiaoxiao Dai, who is a PhD candidate at University of Denver, for her collaboration in the use of the MVGC toolbox and explanation of the fundamental tools for signal processing analysis.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction and Literature Review

1.1 Electroencephalographic Recordings

In 1875, Richard Caton measured and recorded electrical brain signals from an animal scalp [1], and the first study of human brain Electroencephalographic (EEG) signals was conducted by Hans Berger in 1920 [1]. EEG is a technique which is widely used to measure electric signals for brain activity [2]. Using EEG, the electrical activity of the brain is recorded by a number of metal electrodes, which are placed over the surface of a human scalp [3]. The placement of the metal electrodes is standardized by an international system called the 10/20 system [4]. They are placed with a fraction of distance in diameter of brain from front to back and from left to right, where they are separated by 10% or 20% [1]. An electrode label is a combination of a letter and a number where the letters refer to the brain areas, such as: T - Temporal lobe, F - Frontal lobe, C - Central lobe, P - Posterior lobe and O - Occipital Lobe [1]. However, even numbers are categorized to the right side of the brain, and the odd numbers refer to left side of the brain, respectively [1] as shown in Figure 1.1. In the last few decades, neuroscientists have been motivated to study causality of the brain; how one area affects another area of the brain during behavior tasks [2]. Connectivity analysis typically utilizes

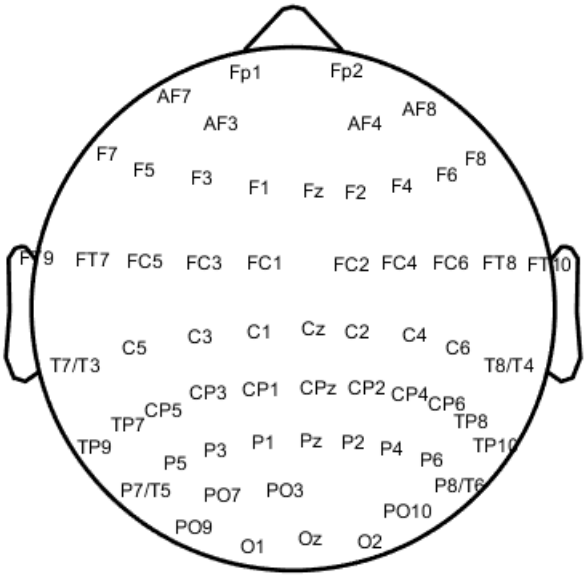


Figure 1.1: International System 10/20 for 61 channel EEG.

EEG signals [5], however, interference in EEG measurements may cause difficulty in estimation of information flow in the brain [6]. Connectivity can result from electrical interference from different sources, internally or externally [7]. Internal interference can appear when a subject moves, e.g., in the action of blinking eyes or head movement. External interference can arise from the noise of power lines (50 or 60 Hz), electrodes, or cable movement [7]. An approach to removal of blinking eye interference is measuring the signal to noise ratio (SNR) between periods when the subject is instructed to blink and periods with no observed blinking [7]. Removing head movement interference has been proposed by placing an accelerometer subjects and applying an independent component analysis method to separate the artifacts of EEG signals which are related to heading movement [8].

1.2 EEG Frequency Bands

Brain activity generate waves that are mixed with many frequency levels up to 500 Hz [9]. There are five major frequency bands that are clinically interesting. They are denoted from low to high frequencies respectively as in Table. 1.1 [9].

Delta Band	Theta Band	Alpha Band	Beta Band	Gamma Band
Up to 3 Hz	4 - 7 Hz	8 - 12 Hz	13 - 32 Hz	33 - 45 Hz

Table 1.1: EEG frequency bands.

The delta band has the slowest wave in terms of frequency, which is lower than three hertz, and has the highest amplitude. This band appears in infants, children under the age of one, and in adults during deep sleep. Theta band also has slow waves which fall in the frequency band from four to seven hertz; these are usually seen in children [9]. The alpha band is the frequency band from 8 to 12 Hz, and it is usually seen in adults; this appears when an individual is in relaxation mode or has eyes closed. The waves occur on both sides of the brain but are mostly higher in amplitude for the non-dominant brain side [9]. The Beta band has a small amplitude and fast frequency, ranging from 13 Hz to 32 Hz; this band is seen in anxious patients or those with open eyes. Also, this band can appear on both sides with symmetrical distribution. The gamma band is the fastest frequency band for an EEG signal, and it falls in the range of 33 Hz to 45 Hz. It normally appears in short-term memory and matching of recognized sound tasks or sensations. Sensory Motor Rhythm (SMR) is a lower end of the beta band which falls between from 12 to 15 Hz. This is very important in relation to the sensory motor area. SMR appears in the motor area when the EEG signals are recorded; SMR may be reduced by moving the arm or leg [9]. Normal EEG signals in adults who are awake exist in the alpha band and beta band [9].

1.3 Brain connectivity

The relationships between different brain regions and the dynamics of brain interactions have been investigated in many research studies [10] by applying different effective connectivity for different types of brain recordings [10]. In earlier studies of activated motor areas for blood flow measurement when subjects perform simple or complex movement tasks, the areas of supplementary motor, lateral premotor, somatosensory, and cingulated motor areas were analyzed [11, 12, 13] . Functional Magnetic Resonance Imaging (fMRI) is widely used to detect activated brain areas and connectivity for brain areas. In the study of event-related coherence for the right and left sensorimotor areas between the contralateral sensorimotor for right finger movement, the dynamic interaction of separated regions is provided by event-related coherence [14]. The study of task-related coherence and task-related spectral power reveals the activity of cortical regions under the control of complex movement eight right-handed subjects in the alpha and beta band frequencies [15]. There are two identified types of brain connectivity. The first is functional connectivity, which investigates the correlation of temporal or time courses between neurological events; the second is effective connectivity, which indicates the influence of one system on another [16, 17, 18]. Neuroscientists have been concerned with investigating the dynamic causal of directionality of information flow for modeling and multivariate autoregressive modeling [19, 20, 18]. Granger causality is an important technique to explore dynamic causal relationships for two-time series [17]. The application is done for motor fatigue, which reveals that fatigue reduces brain connectivity in the motor area [21]. In the beta frequency band (14-30 Hz), the sensorimotor area is widely observed for human brains and non-human brains [22]. The beta band's relationship to the post central area and motor cortex are still poorly understood.

1.4 Granger Causality in Brain Analysis

One of the methods used to analyze multichannel EEG data sets is the Granger causality concept, which was first introduced by Clive Granger [23]. The Granger causality method is used to examine the connectivity for multivariate autoregressive models [24, 25] and has been performed for EEG recordings in [26, 27, 18]. Multivariate Granger causality is also adapted for fMRI and applies the graph concepts to investigate the temporal dynamics and brain causality [28]. By comparing Granger causality with dynamic causal modeling, fMRI data find that the causality uses the temporal priority as a directed causality measure (DCM), however, the DCM is disrupted on the network [28]. The Granger causality is used with healthy patients for emotion cognitive in the gamma band (30- 50 Hz) demonstrates that negative emotion face has larger causality than a face with positive emotions [29]. By applying independent component analysis and Granger causality for EEG of emotional states, the " network of brain " concept classifies the brain based on the causal connectivity brain network [30]. Using Granger causality analysis and graph theory for patients with Parkinson's disease in comparison to a control group of resting state-fMRI, researchers found that information flow of PD patients have less connectivity than the healthy control [31].

1.5 Parkinsons Disease

Tens of millions of people suffer from Parkinson's disease worldwide, and this number is increasing as the elderly population is growing [32]. PD patients suffer from progressive neurological disorders and difficulties, suchlike shaking and slowness of movement [31]. In earlier studies, the brain activity of Parkinson's disease has been investigated in both cortical and sub-cortical regions [33, 34]. EEG and MEG recordings of PD patients in the basal ganglia have shown a relationship between

electrical activity and pathological changes [35, 36, 37]. EEG results indicate higher abnormalities for PD than normal patients [38] when there is alteration of the cortical basal ganglia due to Parkinson's disease [39]. Neuroscientists and researchers are interested in pattern models that can determine the neural oscillations for PD patients [40]. Recent studies suggest that the most dominant frequency bands for sensorimotor areas are the alpha (8-12 Hz) and low beta (13-20 Hz) in people with movement disorder [41, 42, 43, 44]. An increase of oscillations and inter-regional coherence in the beta band (13-30 Hz) is due to the loss of dopamine in Parkinson's disease patients [42]. It is found that the loss of dopamine affects brain connectivity and networks [45]. One of the standard treatments of PD is a dopamine pro-drug, Levodopa; however, long term usage (five to ten years) of this drug results in motor complication for 80% of PD patients [32]. For the EEG power spectral analysis, delta and theta frequency bands have no significant differences after L-dopa intake for non-demented patients [46]. An EEG frequency analysis for three groups of ten subjects of dementia PD, non-demented PD, and a normal control group reveals that there is a significant decrease in alpha amplitude for demented PD patients unrelated to motor weakness [47]. However, the amplitude increases in the delta and theta ranges in comparison to the other groups. Furthermore, the increase happens in the theta and delta ranges with more severe motor weakness of non-demented PD patients [47].

1.6 Motor Activity and Verbal Fluency of Parkinson's Disease

The basal ganglia is suggested to control the primary motor cortex in the brain [48] and contributes to PD symptoms [48]. Earlier studies of causality of the motor cortex in the human brain relate to motor disturbance [49, 50, 37] and its relevance

for PD patients [51, 52, 37]. The loss of dopamine in the frontal-striatal cortex for PD patients affects verbal processing and plays a significant role in functioning of lexico-semantic fluency [53]. In order to test this, PD patients performed a phonemic fluency where they are asked to generate a list of words that starts with a given letter, as well as provide words in semantic categories for semantic fluency [54]. One study examined PD and control subjects with four verbal fluency tasks in which the number of words and the frequency of those words was tested for the control patients once, and twice for PD patients (when they were on and off medication). The results demonstrate that those PD patients who are off the dopamine medication generate fewer words than those PD patients who are on their medication. The frequency of words for PD patients who are either on or off medication is not significantly different; however, differences are significantly noticed between PD patients who are off medication and healthy control patients [53].

Finger Movement Behavior Investigation

In the study of finger movement behavior, we assess cortical information flow in two subjects with Parkinsons disease [55]. Connectivity was measured by applying a Granger Causality algorithm to EEG- data collected during a left and right hand movement task. The sensorimotor rhythm (12-15 Hz) was extracted from the EEG data and further analysis was performed on the upper extremity motor planning and sensorimotor integration areas of the left and right brain. The extracted graph features were classified by machine learning techniques for the right and left electrodes. We observed increased connectivity in the left and right motor planning areas (F3 and F4) contralateral to the behavioral task side. Recognition of the left and right motor planning areas had a rate of 83.3% for F3 and 91.7% for F4. Similarly, we observed increased connectivity in the left and right sensorimotor integration areas (C3 and C4) contralateral to the behavioral task side. Recognition of the left and

right motor planning areas had a rate of 91.7% for C3 and 91.7% for C4. Our results demonstrate the ability to classify the flow of information in the brain. These methods will be applied in the future to identify pathological information flow for disease such as Parkinson's disease. These promising results of research have already been accepted by Asilomar conference 2015 [56].

1.7 Verbal Fluency and brain regions

Usually, verbal fluency tests are performed to examine cognitive disorders for PD patients [57, 58, 59], even in early stages of PD, verbal fluency problems are observed in most PD patients [57]. The frontal cortices of the brain are considered to be the active areas for verbal fluencies in PD and healthy patients [60, 61, 62, 57]. Early studies of fMRI and blood flow in the cerebral area during verbal fluency tasks indicate activity in the left superior temporal cortex [63]; performance of verbal fluency is lower in demented PD patients than non-demented PD subjects [64, 65]. The impairment of phonemic fluency test is less in PD patients than semantic fluency loss [60, 66, 67]. Both semantic and phonemic fluency tests demonstrate significant activity in the left inferior frontal cortex [68, 67, 69]. fMRI studies report that the brain activation is also found in the left inferior and middle frontal areas [70, 71]. Furthermore, the loss of the phonemic fluency is impaired in the frontal lobe; whereas, the damage of the semantic fluency is activated more in the temporal area than the phonemic fluency [70, 72, 73]. Also, some studies suggest that semantic fluency can occur in the area of left inferior frontal gyrus [69, 74, 70, 75]. Speech production demonstrates activation in Broca's area in the left Brodmann's area (BA) 45 and 44 [75, 76, 77].

	Left hemisphere	Right hemisphere
BA44	F7, F5, F3, Ft7, Fc5	F6, F8, Fc4, Fc5, Ft8
BA45	Af7, F7, F5, Ft7, Fc5	Af4, Af8, F6, F8, Ft8
BA4	Fc3, Fc1, C3, C1, Cp3	Fc2, Fc4, C2, C4, C6
BA6	F3, F1, Fc3, Fc1, C3	F2, F4, Fc2, Fc4, C4

Table 1.2: Brodmann's areas for BA44, BA45, BA4 and BA6.

The activation of the Primary Motor Cortex in Brodmann's area is defined as A4; the Premotor cortex and Supplementary Motor Cortex activation area are in BA6 [76]. The nodes of electrodes that are categorized in each of Brodmann's areas belong to the verbal fluency and motor activity for the left and the right sides of the brain and are addressed in [76], as in the Table 1.2.

1.8 Motivation and Objectives

From the previous sections, the knowledge of different aspects and methods of brain connectivity is evident, but there are still questions regarding brain causality during performance of multi-behavior tasks. There is still a great need to explore brain causality of multichannel EEG signals during different behavior tasks for patients with Parkinson's disease and healthy control subjects. Notwithstanding the results from studies which measure directed connectivity during task behavior performance, the signal processing tool of Granger's causality is one of the best methods to extract brain interactions of EEG recordings for movement tasks and verbal fluency; this method has been applied to many EEG recordings of different behavior tasks [26, 27, 18]. The main goal of this research is to extract the strongest causal and activated regions of EEG signals during different verbal fluency tasks, such as phonemic fluency, reading fluency, semantic fluency, and category semantic fluency

for PD patients and healthy controls. These activated regions have been recommended for different studies of fMRI research for verbal fluency and motor tasks in [70, 71, 75, 76, 77, 78, 79] which propose that the activation for verbal fluency and motor cortex falls in the regions of (BA45, BA44) and (BA4, BA6), respectively [78, 79]. The secondary objective is to exploit the feasibility of classifying PD vs. healthy control (HC) based on the causality analysis by applying a method of machine learning for each task of verbal fluency. In order to achieve the proposed goal, the next procedure occurs:

1. Signal processing and filtering of Granger causality averaged by 45 trails of verbal fluency tasks, which determines the most activation frequency band during causality experiment for alpha band, beta band and gamma band.

2. Granger causality realization averaged by 5 random trails of common verbal tasks, in which the most causal and activated channels are identified before the onset of speaking by 600 ms for each participant. Features channels resulted from Granger causality experiments are proposed from the measured Granger causality strengths for all tested channels.

3. Graph learning classification is utilized in order to classify feature channels strength between PD vs. HC. Kernel Support Vector Machine (SVM) and soft-margin technique take place in order to separate data points that are not separated perfectly by linear SVM.

4. The experiment results of all the verbal fluency tasks are summarized using confusion matrices.

Although the activation regions and connectivity of different behavior tasks have been studied and stated in previous literature, there are few studies which illuminate the relationship of brain causality and activated regions by using the

brain signals of EEG. The primary concentration of this thesis is the analysis of data and signal processing which identifies Granger causality method and graph learning techniques in order to understand the causal relations of the brain during verbal fluency tasks. The thesis outlined as in the following:

- Chapter 1 introduces Electroencephalogram signals and briefly explains the signal processing of EEG signals. It also reviews and states of the findings from studies which relate to brain connectivity and Granger causality methods applied to different brain signals. Then, it declares the motivation of this thesis and the thesis outline.

- Chapter 2 reviews the Granger causality method based on the multivariate autoregressive models. Furthermore, the chapter defines graph learning techniques and feature extraction of Granger causality by the vertex strength method. Also, it discusses the classification techniques of support vector machines by illustrating the soft margin method.

- Chapter 3 describes the verbal fluency tasks methods of recordings for the datasets recording of EEG signals and presents the area of interest during verbal fluency tasks. It also demonstrates the Granger causality signal processing and causality relation between channels for each fluency task.

- Chapter 4 presents graph-learning methods of feature extraction in order to extract the channel strength of Granger causality results. Moreover, it illustrates feature classification of the resulted strengths of Granger causality based on verbal fluency activated channels for each fluency task. Then, the chapter presents the results of classifications with the results of earlier established techniques for extracting the activation region of verbal fluency tasks.

- Chapter 5 discusses the advantages and disadvantages of studied methods and

presents the conclusion for the thesis. It also demonstrates the suggestion of future work.

Chapter 2

Causality and Machine Learning Methodology

2.1 Causality Method

2.1.1 Bivariate Autoregressive Models of Granger Causality

In the neurological analysis, the EEG data sets are considered as a multivariate time series since they are recorded from multi-electrodes on the patient's scalp. The neural interactions and directionality can be extracted from multivariate time series, which provide the fundamental frameworks to analyze the data sets in order to understand the neural systems [80]. Granger causality provides the mathematical frameworks for multivariate time series [80]. Granger causality is a standard tool for statistical computation and for determining the directional influence or interactions of systems variables [81]. The concept is based on a statistical study of two data sets of time series $X_1(t)$ and $X_2(t)$ where the historical information of $X_1(t)$ can improve the prediction of future values of $X_2(t)$ in the bivariate autoregressive model. The variable $X_1(t)$ is a Granger cause of variable $X_2(t)$ when the past information of $X_1(t)$ helps predict the future $X_2(t)$ and the two-time series. The time domain of

unconditional Granger causality can be referenced in full regression of the two time series, as in the following [82]:

$$\begin{aligned} X_1(t) &= \sum_{i=1}^{p} A_{11}(i)X_1(t-i) + \sum_{i=1}^{p} A_{12}(i)X_2(t-i) + E_1(t) \\ X_2(t) &= \sum_{i=1}^{p} A_{21}(i)X_1(t-i) + \sum_{i=1}^{p} A_{22}(i)X_2(t-i) + E_2(t) \end{aligned} \tag{2.1.1}$$

When A_{ij} are the regression coefficients from variables $X_j(t)$ to $X_i(t)$ of the full regression time series and E_i are residuals, white noise components, or prediction errors of each signal [83]. The number of lagged observations is denoted by p which also indicates the model order[83]. The prediction of a signal is based on the past values of its own signal and the other signal [81]. Furthermore, the prediction is contributed by each lagged observation for the two signals. Furthermore, Granger causality is defined as the past values of X_2 which help future values of X_1 over the prediction of its own past values [83]. The two time series are expressed, as in the following matrices, in order to extract the covariance matrix:

$$\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \sum_{i=1}^{p} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} X_1(t-i) \\ X_2(t-i) \end{pmatrix} + \begin{pmatrix} E_1(t) \\ E_2(t) \end{pmatrix} \tag{2.1.2}$$

The residual covariance matrix of the white noise E gives the accuracy of a prediction process for the two time series [83]. The covariance matrix is expressed in the following [83]:

$$\sum = \mathrm{Cov} \begin{pmatrix} E_1(t) \\ E_2(t) \end{pmatrix} = \begin{pmatrix} \sum_{11} & \sum_{12} \\ \sum_{21} & \sum_{22} \end{pmatrix} \tag{2.1.3}$$

The unconditional Granger causality is based on the unconditional dependence of the first signal X_1 on past values of the second signal X_2, which makes the regression

coefficients of the other signal equal to zero. The reduced regression forms for the two signals $X_1(t)$ and $X_2(t)$ depend on omitting past values of the other signal [83]. Furthermore, the reduced regression prediction depends only on its own past values, as evident in the following: [83].

$$\begin{aligned} X_1(t) &= \sum_{i=1}^{p} A_{11}(i) X_1(t-i) + {E_1}'(t) \\ X_2(t) &= \sum_{i=1}^{p} A_{22}(i) X_2(t-i) + {E_2}'(t) \end{aligned} \tag{2.1.4}$$

Now A_{ii} are the reduced regression coefficients and E_i' are the reduced regression residuals. The following expressions define the residual covariance matrices of X_1, the reduced regression form, and the X_2 reduced regression form respectively :

$$\begin{aligned} \sum {}'_{11} &= \mathrm{Cov}({E_1}'(t)) \\ \sum {}'_{22} &= \mathrm{Cov}({E_2}'(t)) \end{aligned} \tag{2.1.5}$$

Granger causality of the two signals stands to calculate the coefficients of full regression form and reduced regression form based on a better model of data. There are two famous model selection criteria; both the Akaike information criterion (AIC) and Bayesian information criteria (BIC) are widely used in the Granger causality method to determine the appropriate model order. The data modeling criteria of Maximum likelihood (ML) theory provides the fundamental framework of analysis [83]. The framework analysis is considered as a log-likelihood ratio statistic that is the measurement of an appropriate comparison for models of full regression and reduced regression [83]. The appropriate log-likelihood ratio test, which selects the lag length of a vector autoregressive (VAR) model, defines the statistic of Granger causality for null hypothesis of zero causality [83] :

$$
\begin{aligned}
H_0 &: A_{12,1}(t) = A_{12,2} = A_{12,3} = ... = A_{12,p} = 0 \\
H_1 &: A_{12,1}(t) \neq A_{12,2} \neq A_{12,3} \neq .... \neq A_{12,p} \neq 0
\end{aligned}
\tag{2.1.6}
$$

The log-likelihood ratio test is proportional to the generalized variance of VAR models. The mathematical expression of the Granger causality for the bivariate VAR models from X_2 to X_1 and from X_1 to X_2 are defined respectively, as in the following [83]:

$$
\begin{aligned}
\mathcal{G}_{X2\rightarrow X1} &= \ln \frac{|\sum{}'_{11}|}{|\sum_{11}|} \\
\mathcal{G}_{X1\rightarrow X2} &= \ln \frac{|\sum{}'_{22}|}{|\sum_{22}|}
\end{aligned}
\tag{2.1.7}
$$

The residual covariance matrices of VAR models of full regression and reduced regression are the fractions of the log-likelihood ratio of each signal. So, Granger causality between two signals of time series is based on the covariance of the reduced regression and the full regression computations. It is also can be considered as the computation of prediction error based on the interpretation of log-likelihood ratio test of VAR models. Furthermore, it measures the reduced values of prediction error when the past values of X_2 are involved in the model variables of X_1.

2.1.2 Multivariate Autoregressive Models of Granger Causality

The above mathematical expressions of bivariate time series from the previous section are basically defined by linear regression models. From the fundamental method of computing Granger causality, it can be rapidly extended to multivariate autoregressive (MVAR); this is a model theory for design computational efficiency and accuracy of vectors of time series [83]. The MVAR model can be defined for pth orders, as in the following [83]

$$\boldsymbol{X}(t) = \sum_{j=1}^{p} \boldsymbol{A}(j)\boldsymbol{X}(t-j) + \boldsymbol{E}(t), \tag{2.1.8}$$

Where $\boldsymbol{A}(j)$ is the matrix of real values of multivariate autoregressive coefficients $\boldsymbol{x}(t) = [X_1(t), ...X_N(t)]^T$ with N being the number of variables, and $\boldsymbol{E}(t)$ is the vector of white noise, which defines the residuals of MVAR models. In this case, the causality is defined as X_2 Granger causes X_1 if, and only if, past values of lagged observations improve the prediction of X_1 when other lagged observations of other variables X_3, X_4 ... X_N are counted.

Using $N = 3$ as an example, the conditional Granger causality can be defined from the dependency between X_1, X_2, and an additional set of variable X_3. This means that the Granger causality has a possible dependency of X_1 and X_2 on X_3. The conditional Granger causality can be expressed in full regression models for three distributed time series vectors, as seen in the following [83]:

$$\begin{aligned}
X_1(t) &= \sum_{i=1}^{p} A_{11}(i)X_1(t-i) + \sum_{i=1}^{p} A_{12}(i)X_2(t-i) + \sum_{i=1}^{p} A_{13}(i)X_3(t-i) + E_1(t), \\
X_2(t) &= \sum_{i=1}^{p} A_{21}(i)X_1(t-i) + \sum_{i=1}^{p} A_{22}(i)X_2(t-i) + \sum_{i=1}^{p} A_{23}(i)X_3(t-i) + E_2(t), \\
X_3(t) &= \sum_{i=1}^{p} A_{31}(i)X_1(t-i) + \sum_{i=1}^{p} A_{32}(i)X_2(t-i) + \sum_{i=1}^{p} A_{33}(i)X_3(t-i) + E_3(t).
\end{aligned} \tag{2.1.9}$$

The effect of X_3 is eliminated from the Granger causality realization for X_2 to X_1. So, the reduced regression form of X_1 is illustrated as the following [83]:

$$X_1(t) = \sum_{i=1}^{p} A_{11}(i)X_1(t-i) + \sum_{i=1}^{p} A_{13}(i)X_3(t-i) + {E_1}'(t). \tag{2.1.10}$$

Similarly, the likelihood ratio is computed for the covariance matrices of the full regression and the reduced regression forms with the condition of X_3. The null hypothesis test is obtained as in $H_0 : A_{12}(1) = A_{12}(2) = A_{12}(3) = ... = A_{12}(p) = 0$. As a result, the conditional Granger causality $X_2 \rightarrow X_1$ on condition X_3 is given in the following [83]

$$\mathcal{G}_{X2\rightarrow X1|X3} = \ln \frac{|\sum{}'_{11}|}{|\sum_{11}|}. \tag{2.1.11}$$

Therefore, the implication of X_3 to full and reduced regression model forms of X_1 can be assumed as the past values of X_2 improves the prediction of X_1, depending on its own past values, plus the past values of X_3. This conditional Granger causality of the three time series may be considered as multivariate Granger causality. So, multivariate Granger causality is suitable to account for many interactions in a multivariate system since its components work collaboratively with more complex layers than bivariate VAR modeling [83].

The first step to calculate MVAR of Granger causality is to determine the appropriate model order for MVAR. One may select the appropriate model order by ML theory, such as the Akaike information criteria (AIC), Bayesian information criterion (BIC), or cross-validation [83, 84]. The model order selection is beneficial to balance the number of MVAR parameters and should be appropriately selected to avoid over-fitting a finite data sequence [83]. The next step of Granger causality realization is to estimate the model parameters of MVAR by maximizing the likelihood ratio for MVAR models for both full regression form and reduced regression form [83]. Once all parameters have been estimated, Granger causality is computed

from the equation of conditional Granger causality between pairs within selected channels in a multivariate system in order to generate one graph g. The realization of Granger causality is a non-negative magnitude when it is considered positively biased [83].

For the statistical testing of Granger causality, the multivariate Granger Causality MVGC toolbox provides significance testing and confident interval computation routines e.g., permutation test and bootstrapping test which are based on the simulation of surrogate time series [83]. The surrogate data series can be generated from two main algorithms which are typical realization and constrained realization [85, 86, 87]. The first generates a surrogate data set from a mode that gives the best model fit for the original data and the second is an algorithm to generate surrogate data from the original data set [85]. The statistical bootstrapping can estimate the statistical measurement, e.g., mean, variance or median [87, 88]. The procedure of bootstrapping is started with replacing the samples of the interested data [87, 88]. Statistical distribution estimation is computed for the re-sampled data for the first replacement and then repeat the process for many times [88].

2.2 Graph Learning

Machine learning techniques and graph learning statistics have been commonly used in pattern recognition and data mining [89]. Since brain analysis graphs have been studied for research over many decades, the need for graph learning is essential to help understand brain functions better [89]. Application and development of machine learning techniques and graph representation for brain network analysis has been gaining attention from researchers in neuroscience. The goal of functional brain graph analysis is to classify and characterize brain activation patterns which result from a cognitive state change (within subject or across subjects) that may be related to experimental stimulations [89]. To interpret the integration and seg-

regation of brain activities, several topological measurements are utilized to extract features from the graphs generated from neural data (which in this case is the EEG data recorded in verbal fluency tasks). Complex network graphs compile abstract information of graph representations [90]. Each graph contains a finite number of nodes (vertices) and a finite number of connections (edges), which represent the interaction between network nodes [90].

In general, a finite number of vertices V and a finite number of edges E forms a graph $g = (V, E)$. In the literature of the complex networks, graph vertices are identified by an integer index $i = 1, 2, 3,..$ N, and graph edges can be identified as pairs (i, j) which denote the connectivity edge or link between vertex i and vertex j [91]. Also, it supposed that there is no self-connectivity for a single vertex; for example, there is no edge or link of (i, i) or (j, j), as well as for these pairs (i_1, j_1) or (i_2, j_2) where $i_1 \neq i_2$ and $j_1 \neq j_2$ [91]. The graph g is generated from the previous section of Granger causality, as in the following:

$$g = (V_g, E_g, \alpha_g, \beta_g), \tag{2.2.1}$$

where V_g is the vertices set, E_g is the edges set, and α_g is the vertex labeling function while β_g is the edge-labeling function. In this paper, EEG data channels are the graph vertices; the directional causal relationship between two channels is the graph edge; and the edge-labeling function is the Granger causality estimations $\mathcal{G}_{X2\rightarrow X1|X3}$, which is carried out in a previous section.

2.2.1 Vertex Strength or Node Strength

Since the growth of complex networks has been escalating and has attracted many fields of research, the visualization of representations can be summarized with complex information in two-dimensional representation [89]. For many complex

networks, the large number of vertices makes it difficult to understand and interpret the differences between them. As a result, the measurement of node strength over all nodes of a graph produces an abstract representation that describes the graph overall [89]. Node strength is an intuitive realization of a complex graph, in order to present connectivity degree for a single node with respect to the total connections with other nodes [90].The strength of the i-th vertex concludes the number and weight of connections between itself and all the other vertices of graph g and is defined as

$$s_i = \sum_{j=1}^{N} \beta_g(i,j), \forall i, \tag{2.2.2}$$

where N represents the number of channels of EEG dataset. β_g is the edge labeling of the network, which is Granger causality estimation between two pairs, and the vector of vertex strength of the graph for all vertices is denoted as the following

$$\mathbf{s} = [s_1 s_2 ...., s_N]^T. \tag{2.2.3}$$

In other words, the strength vertices vector is considered as the feature vector of the graph g. Therefore, the feature vector of each Granger causality graph for EEG datasets can demonstrate or summarize the connectivity between each vertex (channel) with other vertices (channels).

2.2.2 Linear Support Vector Machine

Support vector machine (SVM) is a highly effective and robust supervised machine learning method [92]. It is one of the popular machine learning techniques for classifications and learning tasks. It can analyze and recognize different modalities of datasets and map the data into different feature classifications in the space, separating them into different areas [92]. Classifications are achieved through SVM

by finding the best possible hyperplane, which gives the largest separation margin between separated feature classes. In general, the training graph feature vector set with m vectors is defined as the following [92]

$$\mathcal{S} = \{(\mathbf{s}_i, c_i) | c_i \in \{-1, 1\}\}_{i=1}^{m} \tag{2.2.4}$$

where c_i is either -1 or 1, indicating which feature class $\mathbf{s}_i$ belongs to. The support vector points are data points which are closest to the hyperplane. The algorithm of SVM is to separate the data points by classifying them into $c_i = 1$ and $c_i = -1$ with the maximum margin of the hyperplane which is described as

$$\mathbf{h} \cdot \mathbf{s} - b = 0. \tag{2.2.5}$$

Here, b is the bias of the decision function, $\mathbf{s}$ is the vector of points, and $\mathbf{h}$ is the normal vector to the hyperplane [92]. The offset of the hyperplane is usually given as $\frac{b}{\|\mathbf{h}\|}$ from the origin in the hyperspace. The linear separation of SVM is where the data points are separated linearly in space, and two parallel lines of hyperplane are placed at the maximum separation region that does not have any points between them. This region is called separation margin. Theoretically, minimizing $\|\mathbf{h}\|$ results in a maximum region of margin that does not fall points into it. The expression of data points that fall in each region of c_i=1 and c_i=-1 is given in the following:

$$\begin{aligned} \mathbf{h} \cdot \mathbf{s}_\imath - b \geq 1 \text{ for } s_\imath \text{ in class } c_\imath = 1 \\ \mathbf{h} \cdot \mathbf{s}_\imath - b \leq -1 \text{ for } s_\imath \text{ in class } c_\imath = -1 \end{aligned} \tag{2.2.6}$$

The optimization problem of the linear SVM is to minimize $\|\mathbf{h}\|$ subject to the data points $s_\imath$ that fall in each region, as in the following:

$$c_i(\mathbf{h} \cdot \mathbf{s}_i - b) \geq 1, \quad \text{where } i = 1, 2, ...m \tag{2.2.7}$$

2.2.3 SVM Soft Margin

Soft margin method is applied since a perfect hyperplane might not exist to perfectly separate feature classes [93, 92]. This method is used for non-separable data sets and the sensitivity of outliers. It can separate the data points as cleanly as possible by choosing the best hyperplane; it can also maximize the distance of the margin to the data points in each class. The misclassification degree of the data points s_i is defined as slack variables ξ_i [93]. Therefore, the hyperplane equation of the soft margin method is given as [93]:

$$c_i(\mathbf{h} \cdot \mathbf{s}_i - b) \geq 1 - \xi_i, \quad \text{where } i = 1, 2, ...m \tag{2.2.8}$$

The optimization method of the soft margin is based on the trade-off between the maximum margin and the small error penalty. The optimization problem of the soft margin method introduces the misclassification degrees $\xi_i \geq 0$ and penalty function $(C \sum_{i=1}^{m} \xi_i)$, as in the following [92]:

$$\begin{aligned} &\arg \min_{\mathbf{h},\xi,b} \{\frac{1}{2}\|\mathbf{h}\|^2 + C \sum_{i=1}^{m} \xi_i\}, \\ \text{subject to} \quad &c_i(\mathbf{h} \cdot \mathbf{s}_i - b) \geq 1 - \xi_i, \\ \text{where} \quad &\xi_i \geq 0, \quad i = 1, 2, ...m \end{aligned} \tag{2.2.9}$$

From the optimization problem, one can conclude that points fall into the correct region within the margin when the slackness degree is $(0 \leq \xi_i \leq 1)$; however, the misclassified points appears when the slackness degree is $(\xi_i \geq 1)$.

Chapter 3

Granger Causality Analysis for Verbal Fluency Behavior

3.1 Verbal Fluency Recordings and Material

3.1.1 Procedure of Verbal Fluency Tests

Verbal fluency defined as the ability to generate as many as words as possible, given fluency conditions like phonemic fluency and semantic fluency. The procedure of verbal fluency tests is to allow participants to generate as many words as possible within sixty seconds by reason of certain fluency conditions. The verbal fluency tests conducted at Colorado Neurological Institute (CNI) examine PD patient and healthy control subjects. It is known that PD patients may have difficulties starting a movement and switching between movements. Furthermore, they may have difficulties with mental activities that require switching from one activity or movement to another. These subjects are asked to wear an EEG cap to record EEG signals and activity for different verbal fluency tasks. The purpose of verbal fluency tests is to understand connections between different cortical areas. The following are some definitions of verbal fluency that are used in this thesis:

- The phonemic fluency test involves subjects generating words that begin with a specified letter of the alphabet, such as F, A, and S; the subjects must state as many words as they think of within 60 s.

- The semantic fluency test requires subjects to generate words (again, as many as they think of in 60 s) that must be classified under a specified category, like animals.

- The category semantic fluency test asks subjects to generate semantic fluency words and switch between semantic categories, as many times as possible within 60 s. The aim of this test is to understand the difficulties for PD patients when they switch between categories.

- The reading fluency is a cortical measurement of subjects who have read the text in 60 s.

3.1.2 Experimental Materials and Tools

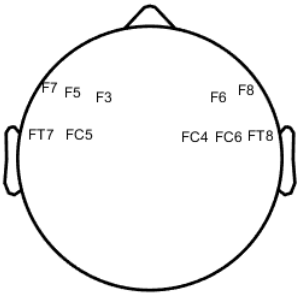
This study was approved by the HealthONE Institutional Review Board, and all subjects provided informed consent for study participation. Four participants were involved two Parkinson's disease patients and two healthy control subjects who volunteered from the Colorado Neurological Institute (CNI). The participants were asked to perform verbal fluency tasks and generate a list of words when a task cue was given. Each participant responded with approximately sixty trials of verbal fluency tasks. These tasks were divided into fifteen trial blocks for each fluency task command, such as: phonemic, semantic, category semantic, and reading fluency. The EEG was measured with 4 g.tec g.USBamp amplifiers and a BrainproductsActicap System. There were sixty-one electrodes placed on each subject's scalp, according to the international 10/20 electrode placement system. The data was recorded with a sampling rate of 4800 Hz. The datasets of PD patients were

defined as PD_1 and PD_2, with the datasets of healthy controls as HC_1 and HC_2, for more convenient reference of the data analysis.

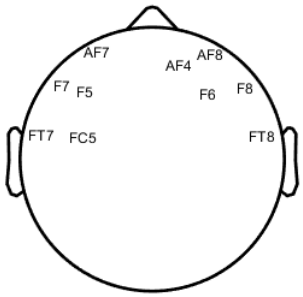
The work in this research primarily depended on signal processing analysis in the MATLAB environment. The MVGC MATLAB toolbox is utilized in the study, which was released by the University of Sussex in the United Kingdom [83]. This toolbox is based on the analysis of Granger causality for multivariate autoregressive models in the time domain and the frequency domain. This toolbox also has serious requirements to perform calculation of Granger causality graph [83]. The second MATLAB toolbox used was the EEGLAB for signal processing analysis in order to display the head map of Granger causality strengths.

3.1.3 EEG Data Preprocessing and Region of Interests

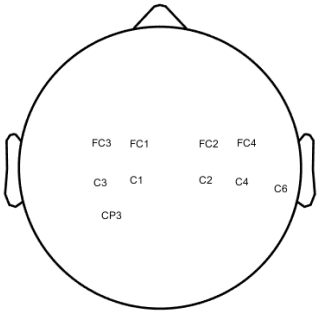
The preprocessing progresses start with down-sampling the signals to 240 Hz. Since the verbal fluency activity in the brain is considered to exist in the left frontal area, addressed in the BA44L and BA45L as in [69, 74, 70, 75] and listed as in Table 1.2, the electrode positions are illustrated in the head maps as in figure 3.1. From this Figure, one can see that EEG channels on the left side of hemisphere and the right side of hemisphere concentrate on the prefrontal areas. Also, there are some mutual nodes between the two areas. Also, the supplementary motor cortex and primary motor cortex channels for right hand action are defined as F3 and C3, respectively, and the F4 and C4 are the supplementary motor and primary motor cortices for left hand performance [51, 52, 37, 48]. The study of this thesis includes the channels which are only related to motor cortex in order to understand the relationship between verbal fluency and motor activity for PD patients and healthy controls. Moreover, the primary motor cortex and supplementary motor cortices are discussed in [76] and listed in Table 1.2. As a result, the illustration of the areas belong to BA4 and BA6 are demonstrated in Figure 3.1.



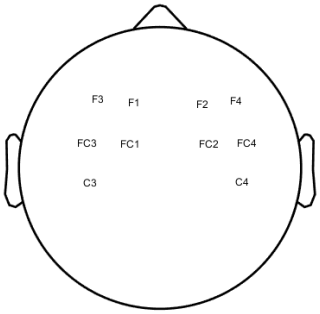
(a) Broca's area of BA44



(b) Broca's area of BA45



(c) Primary Motor Cortex of BA4



(d) Supplementary Motor Cortex BA6

Figure 3.1: Brodmann's area electrode positions of EEG channels.

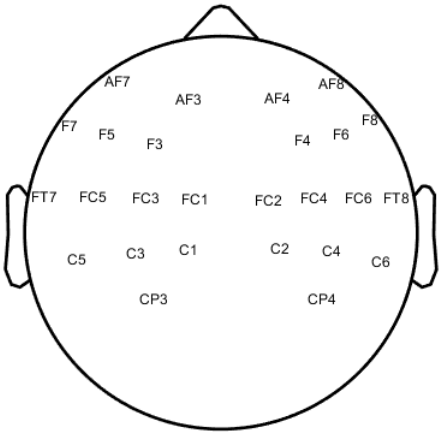


Figure 3.2: Region of interest for 26 EEG channels.

Overall, the channels of Figures 3.1a,3.1b, 3.1c, and 3.1d are merged in order to extract the region of interest.There are twenty-six channels in total, and they belong to prefrontal cortex channels and central cortex channels for both sides of the brain. The region of interest is given as the head map in Figure 3.2.

Typically, brain activity leading to a motor action starts between 150 ms and 200 ms before the response time; therefore, our last study of finger movement behavior task concentrates only on the motor actions for the PD patients. The window length of each trail was 200 ms before the response time. However, in this study, the activation of verbal fluency is proposed in [77] as in the inferior frontal gyrus which is the Brocas area after 400 ms of the stimulus onset. In this study, concentration is on the inferior frontal gyrus, the primary motor cortex, and the supplementary motor cortex where the window length is selected as 600 ms before the participant speaks.

3.2 Granger Causality Analysis

The method of Granger causality is briefly explained in the previous chapter, which depends on the past values of the second signal X_2 in order to improve the

prediction the first signal X_1 for bivariate autoregressive models. The MATLAB toolbox MVGC is used for Granger causality analysis where the datasets PD_1, PD_2, HC_1 and HC_2 consist of sixty-one channels that seriously affect the timing of Granger causality realization because of the large number of lag observations [83]. In regards to EEG channels, only the region of interest channels for analysis of Granger causality was extracted. The approach for finding Granger causality is to compute the causality graph or causality network for a certain task and then use the graph for vertex strength method of graph learning to identify the most activated or causal channel among all twenty-six channels. The computation of Granger causality have some constraints that should be considered, as the large number of trials with different tasks give undesirable results and signal interference. Also, the window length should not be too short, nor too long; this is a trade-off between data likelihood stationary seek and model fit accuracy [83]. The shorter window length helps the data to be stationary, whereas the large number of trails should be implemented by short windows. The longer window improves the computation of multivariate autoregressive model fit. The window length in this thesis depends on the verbal fluency activation. In the approach of Granger causality, the strength of channels that are mostly causal with other channels is interpreted by extracting the feature strength of channels. This method is proposed in the previous chapter for the graph learning technique. Next, the channel strength of the Granger causality is visualized with head maps. The investigation of verbal fluency research procedure by Granger causality begins by filtering signals in order to define the most effective frequency band for verbal fluency behavior. The details of Granger causality filtering are provided in the following section. Then, the focus is on the effective frequency band in order to apply the approach with each verbal fluency task. Also, we applied the channel strength technique for each Granger causality graph of a single test. As well, we demonstrated the channel strengths of each Granger causality

graph by head maps in order to compare most causal and activated channel between all verbal fluency tasks.

3.2.1 Granger Causality Filtering

Many earlier studies on verbal fluency test for fMRI have reported that the activated areas are in the left inferior frontal gyrus [69, 74, 70, 75]. The verbal fluency activation for EEG signals has been proposed in [77] as in the gamma band for Broca's area. The purpose in this study is to investigate the causality activation within three frequency bands (alpha, beta, and gamma) for all the verbal fluency tasks with forty-five trails for each of the datasets PD_1 and HC_1. Since the research includes a large number of trails for verbal fluency tasks, the window length is shortened due to the discussion in the previous section up to 300 ms before the participant speaking. The procedure of EEG Granger causality filtering starts with computing the Granger causality of each frequency band that results in a graph of Granger causality. Channel strength of the causality graph helps the visualization of complex graphs, such as the previous Granger causality graphs, in two-dimensional representation. The channel strength graph learning technique is used to extract the features of each channel for Granger causality. Finally, the channel strength of Granger causality in head maps after normalizing the strength values is illustrated in order to distinguish the difference between the three frequency bands.

3.2.2 Overview of Granger Causality Filtering

Granger causality filtering of verbal fluency has multiple procedure steps in order to determine the most effective frequency band. The procedure analyzes two datasets (PD_1 and HC_1) for verbal fluency analysis such as phonemic fluency, semantic fluency and category semantic fluency. The first step of analyzing each dataset is down-sampling the signals to 240 Hz. The second step is extracting the

region of interest (ROI) for EEG channels which contains twenty six channels in the BA 44, BA 45, BA 4 and BA 6. The third step is filtering the signal with band pass filter in the frequency band (alpha band, beta band, and gamma band). The fourth step is averaging the signals of multivariate channels with forty five trials for each computation of Granger causality and the window length of time epoch is 300 ms before subject speaking. The fifth step is applying Granger causality multivariate time series which produces a Granger causality graph for the averaged signals. The sixth step is extracting the graph features which are denoted as the channel strength in order to visualize the causality of the averaged signal. The seventh step is illustrating the normalized values of channel strengths for the region of interest (ROI) by head maps. The procedure of Granger causality looks for the best effective frequency band for causality in the Broca's area during verbal fluency tasks. The following flow chart demonstrates the procedure process for the Granger causality analysis of verbal fluency as in the Figure 3.3.

Alpha Band

EEG signals consist of five major frequency bands, as mentioned in Chapter 1. The Alpha band is a slow wave appearance of an EEG signal that appears usually for adults when they are relaxing or closing their eyes. The datasets PD_1 and HC_1 were filtered by a band pass filter from 8 Hz to 12 Hz and then generated granger causality for an average of forty-five trails of verbal fluency tests with 300 ms before participants speak.

In general, the Alpha band does not show any effective Granger causality in the area of verbal fluency, located in the left frontal channels for the two data sets PD_1 and HC_1 as in Figure 3.4a and Figure 3.4b. Also, we can note from channel strengths of both datasets that the greatest channel strengths do not belong to the verbal fluency channels, as shown in Figure 3.5a and Figure 3.5b. Furthermore, the

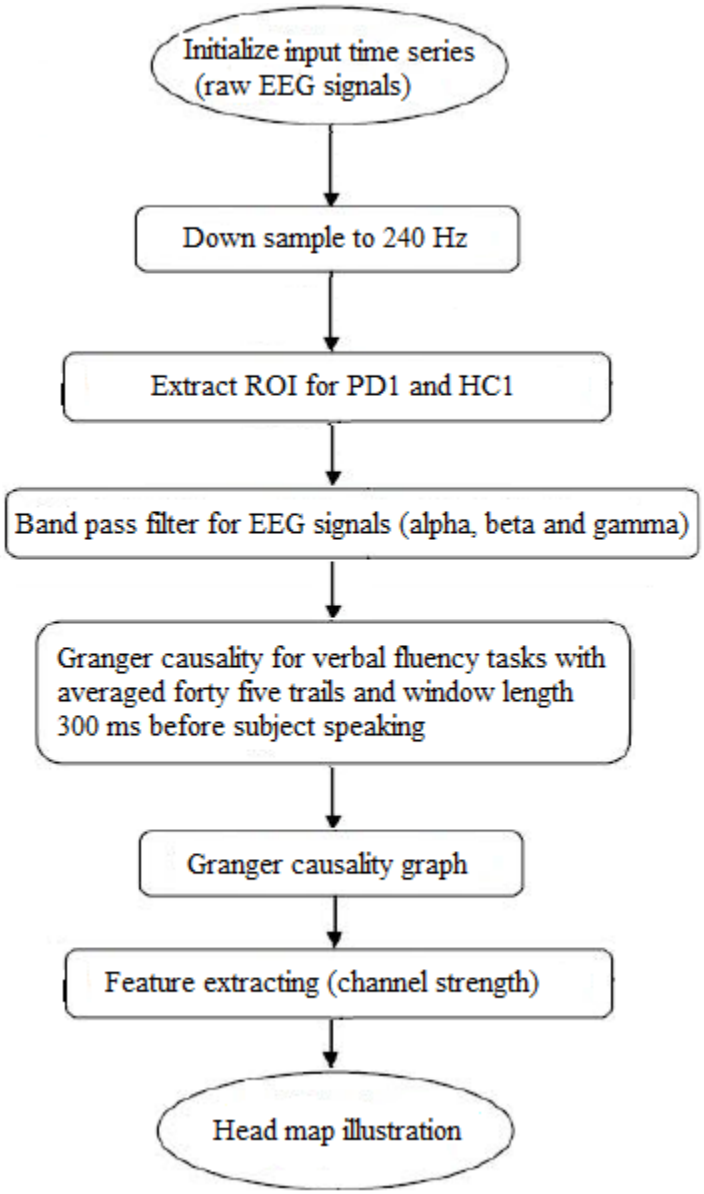
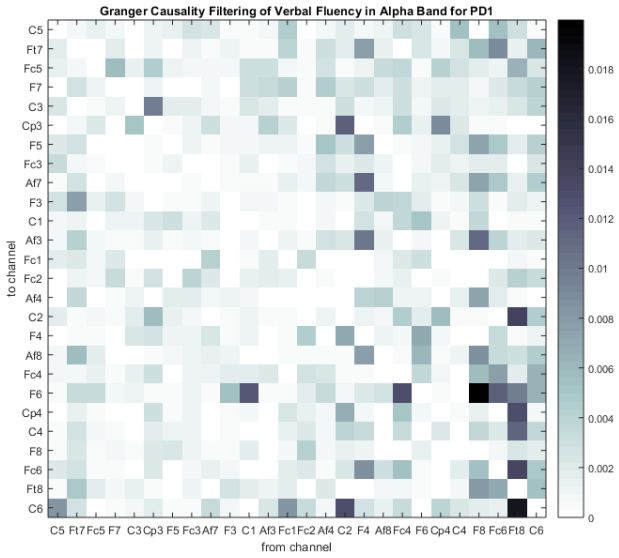
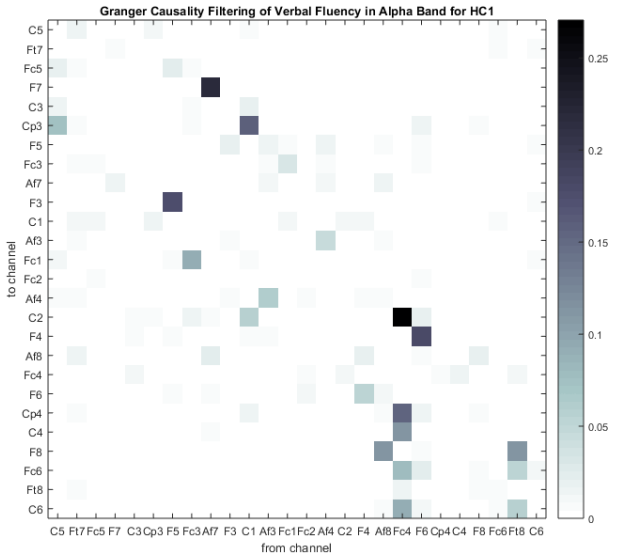


Figure 3.3: The flow chart for Granger causality filtering of PD_1 and HC_1.

channel strength with head maps of the region of interest of both datasets as in Figure 3.6a and Figure 3.6b is evident for Granger causality activation in the left inferior area.

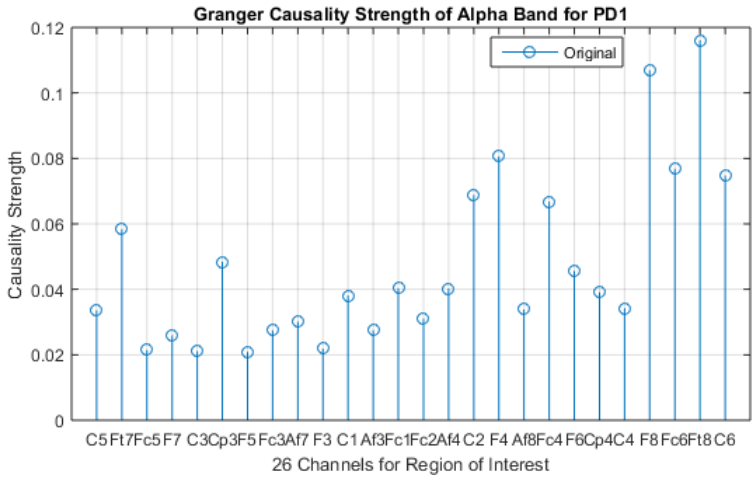


(a) for PD_1

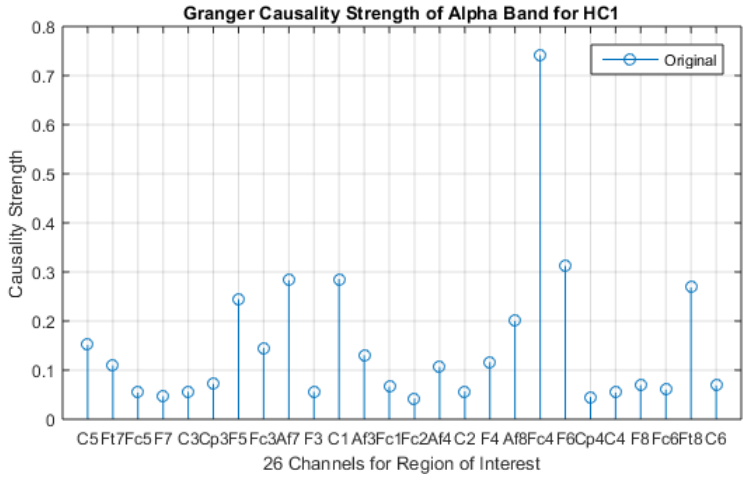


(b) for HC_1

Figure 3.4: Granger causality for 26 EEG channels in alpha band which is averaged for 45 trails and the window starts 300 ms before PD_1 and HC_1 subjects speak.

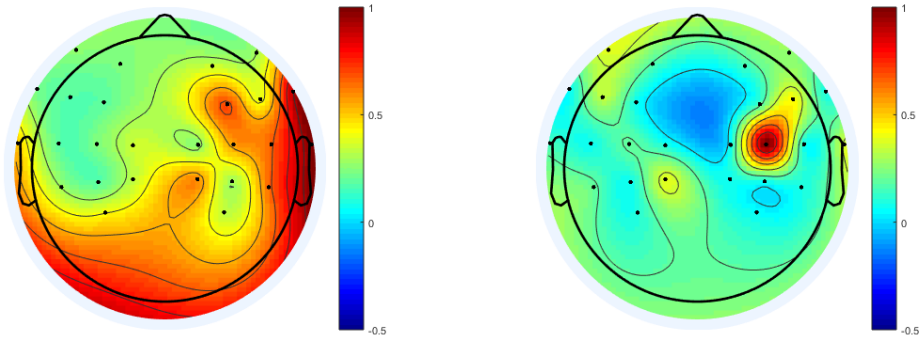


(a) for PD_1



(b) for HC_1

Figure 3.5: Channel strength causality for 26 EEG channels in alpha band for the Granger causality graph in the alpha band for subjects PD_1 and HC_1.

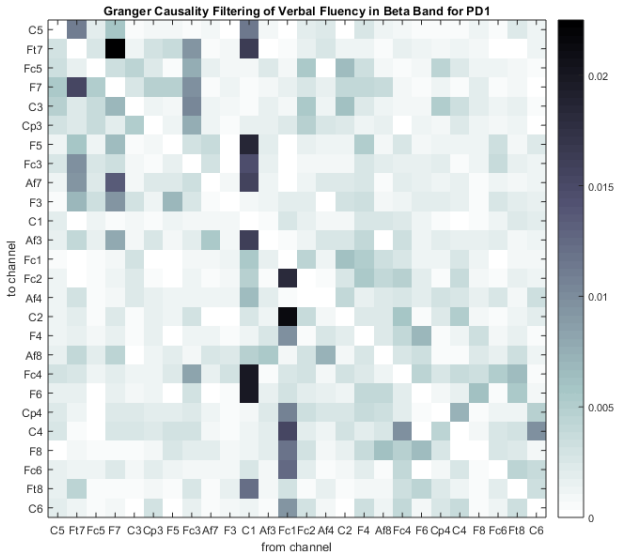


(a) for PD_1 (b) for HC_1

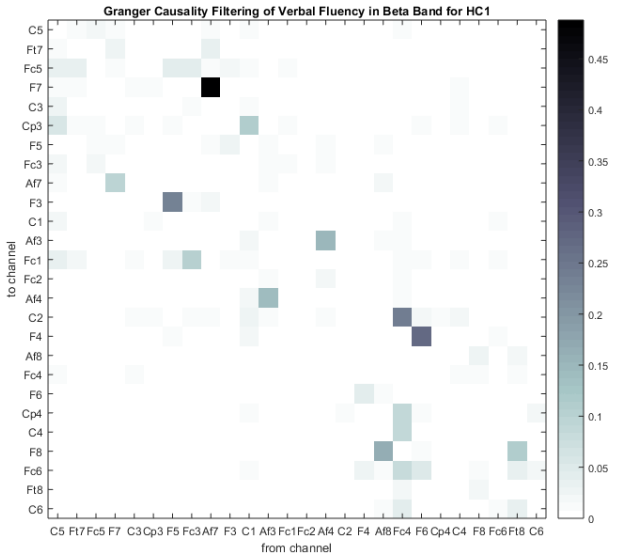
Figure 3.6: Head maps of Granger causality channel strength in alpha band where the values are normalized for visual comparison of Granger causality channel strengths between subjects PD_1 and HC_1.

Beta Band

The Beta band has faster frequency waves and appears when patients are anxious, or they have their eyes open. The band pass filter is used for the frequency band from 13 Hz to 32 Hz. The computation of Granger causality is also averaged for forty-five trials and the times window is 300 ms before subjects speak. In the Beta band, there is no considerable Granger causality in the area of verbal fluency channels for the two datasets, as shown in Figure 3.7a and Figure 3.7b. Also, the strongest channel strengths are not channels in Broca's area, as in Figure 3.8a and 3.8b. The head maps of Granger causality channel strength in the Beta band are given in Figure 3.9a and 3.9b.

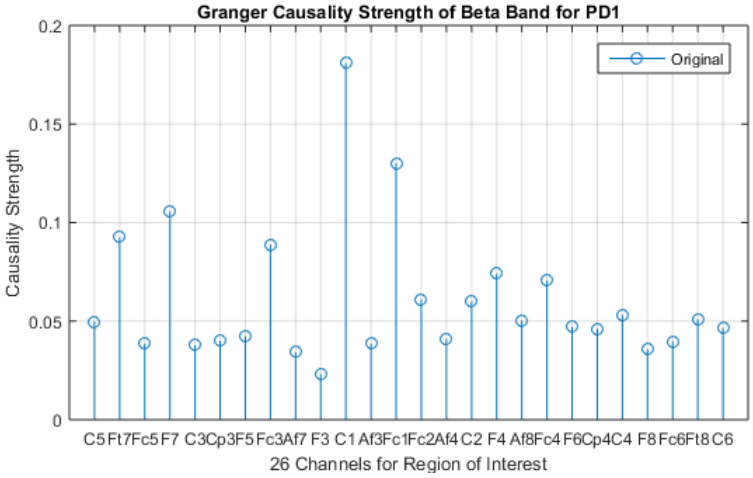


(a) for PD_1

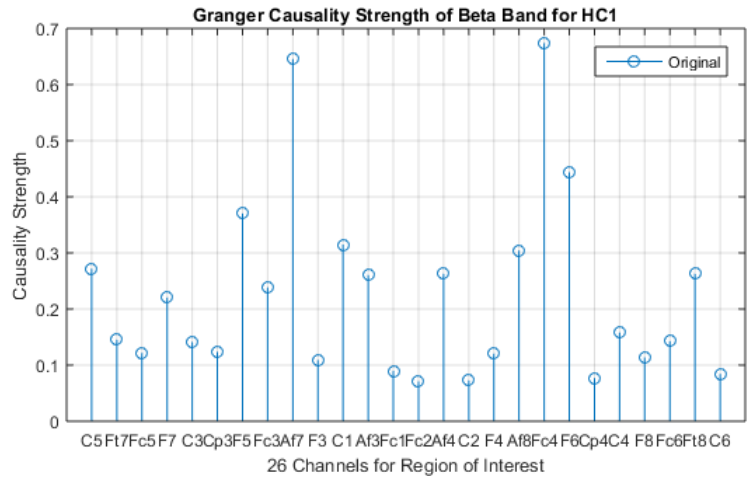


(b) for HC_1

Figure 3.7: Granger causality for 26 EEG channels in beta band where it is averaged for 45 trails and the time window is 300 ms before PD_1 and HC_1 subjects speak.

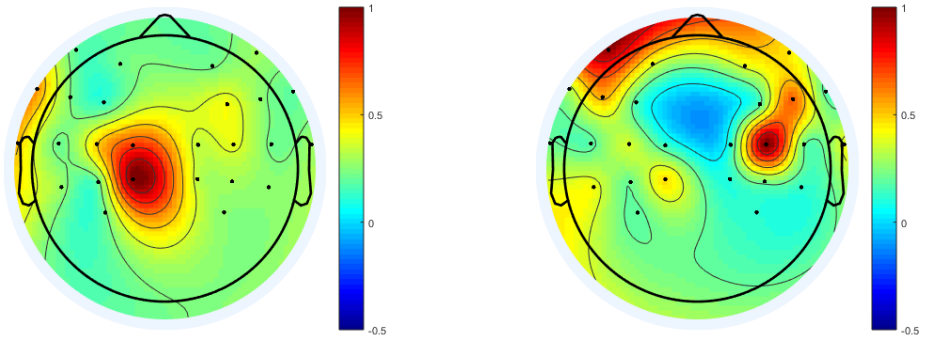


(a) for PD_1



(b) for HC_1

Figure 3.8: Channel strength causality for 26 EEG channels in beta band for the Granger causality graph in the alpha band for subjects PD_1 and HC_1.

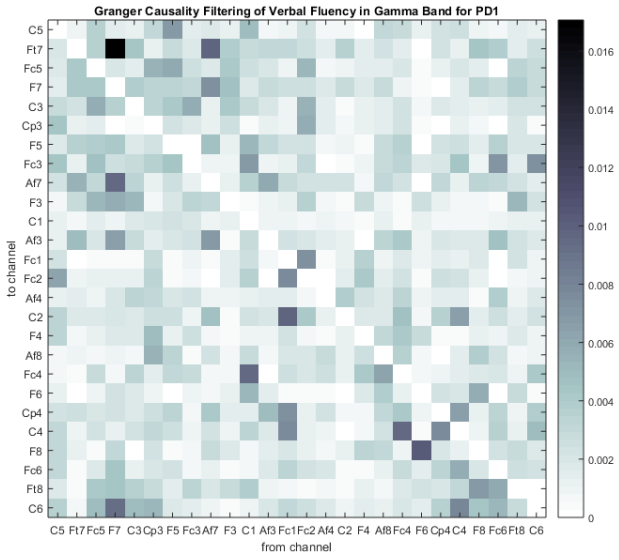


(a) for PD_1 (b) for HC_1

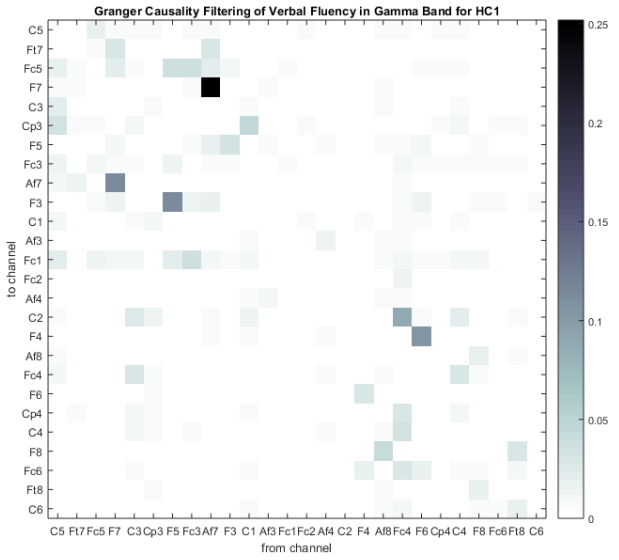
Figure 3.9: Head maps of Granger causality channel strength in beta band where the values are normalized for visual comparison of Granger causality channel strengths between subjects PD_1 and HC_1.

Gamma Band

The gamma band is the fastest frequency wave of EEG signals. It appears in short-term memory or recognized tasks, like sound or sensation. The band pass filter is designed from 33 Hz to 45 Hz where the computation of Granger causality of verbal fluency tasks is averaged for 45 trails with 300 ms before participants speak. The gamma frequency band has the most considerable Granger causality for verbal fluency channels, which are in the left Broca's area with other channels, shown in Figure 3.10a and Figure 3.10b. The Granger causality channel strength shows the highest causality channel strength are channels of left frontal areas, as in Figure 3.11a and Figure 3.11b. Also, the head maps of Granger causality channel strength can demonstrate activation in Broca's area for both PD_1 and HC_1 in the gamma band, as evident in Figure 3.12a and Figure 3.12b. As a result, the gamma band is the most effective frequency band for Granger causality analysis of verbal fluency tasks averaged for 45 trails.

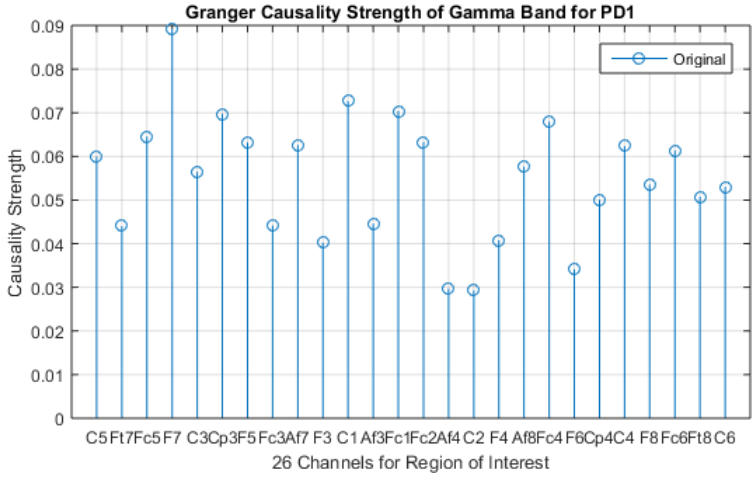


(a) for PD_1

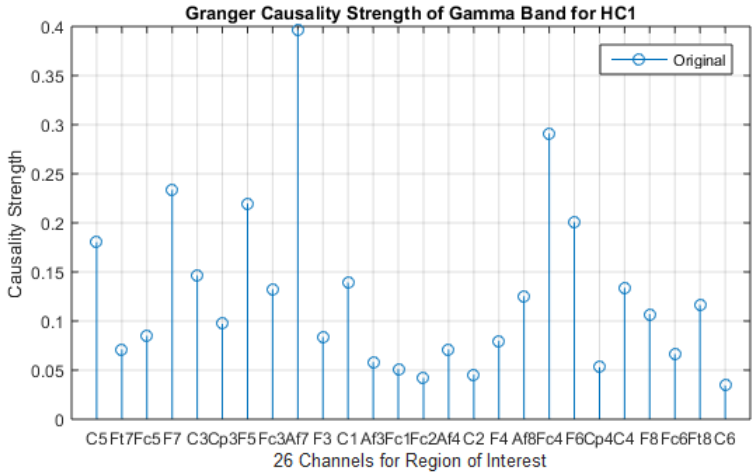


(b) for HC_1

Figure 3.10: Granger causality for 26 EEG channels in gamma band which is averaged for 45 trails and the time window is 300 ms before PD_1 and HC_1 subjects speak.

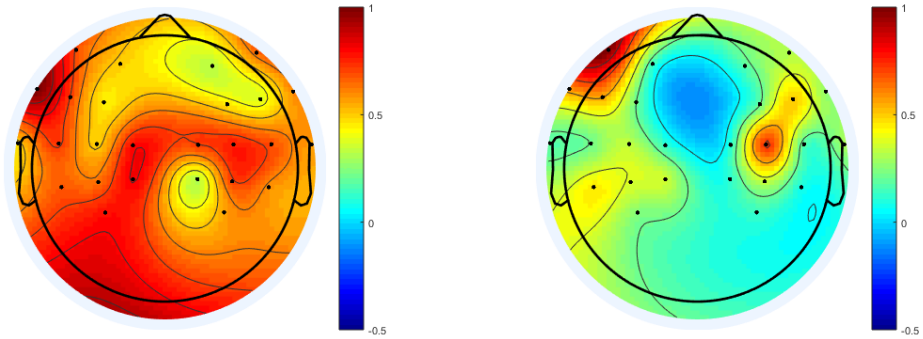


(a) for PD_1



(b) for HC_1

Figure 3.11: Channel strength causality for 26 EEG channels in gamma band for the Granger causality graph in the alpha band for subjects PD_1 and HC_1.



(a) for PD_1 (b) for HC_1

Figure 3.12: Head maps of Granger causality channel strength in gamma band where the values are normalized for visual comparison of Granger causality channel strengths between subjects PD_1 and HC_1.

3.2.3 Granger Causality of Verbal Fluency

Since the most active frequency band for Granger causality of verbal fluency tasks is the gamma band, the focus of analysis is only on this frequency band in order to investigate the relationships of different channels of verbal fluency tasks for all datasets PD_1, PD_2, HC_1 and HC_2. The objective is to define interactions within Broca's channels of each verbal fluency task where researchers expect the most causal channels to be Ft7 and F7. The computation for Granger causality begins with an average of five trails for each task, and the time window is 600 ms before the subjects speak. The computation of Granger causality is applied for the average of trials with a single fluency task. The total number of Granger causality analysis for each verbal fluency task is ten. Granger causality channel strength is found for each computation.

3.2.4 Overview of Verbal Fluency Analysis Model

Granger causality analysis has multiple procedure steps in order to achieve the proposed method. The method analyzes four datasets (PD_1, PD_2, HC_1 and HC_2)

for verbal fluency analysis such as phonemic fluency, semantic fluency, category semantic fluency and reading fluency. The first step of analyzing each dataset is down-sampling the dataset to 240 Hz. The second step is extracting the region of interest (ROI) for EEG channels which are twenty six channels in the BA 44, BA 45, BA 4 and BA 6. The third step is filtering the signal with band pass filter in the gamma band. The fourth step is averaging the signals of multivariate channels with five trials for each computation of Granger causality and the time window of is 600 ms before subject speaking. The fifth step is applying Granger causality multivariate time series regression which produces a Granger causality graph for the averaged signals trials. The sixth step is extracting the graph features which is denoted as the channel strength in order to visualize the causality of the averaged signal. The seventh step is illustrating the normalized values of channel strengths for the region of interest (ROI) by head maps. The eighth step is applying Granger causality statistical testing which indicates the significance of channel strengths. The computation of Granger causality is repeated ten times for a single verbal fluency task. The Granger causality method is applied for each dataset of PD_1, PD_2, HC_1 and HC_2. The ninth step is applying graph learning for support vector machine in order to classify the channels which are considered as the most causal channels for PD datasets and HC datasets. The tenth step is summarizing the SVM soft margin classification of the verbal fluency for PD vs. HC. The following flow chart demonstrates the procedure process for the Granger causality analysis of verbal fluency as in the Figure 3.13

Phonemic Fluency

The performance of phonemic fluency is based on patients generating as many words that start with a certain letter, such as F, A and S, as they can think of in 60s. The computation of Granger causality is based on the average of five trials and 600

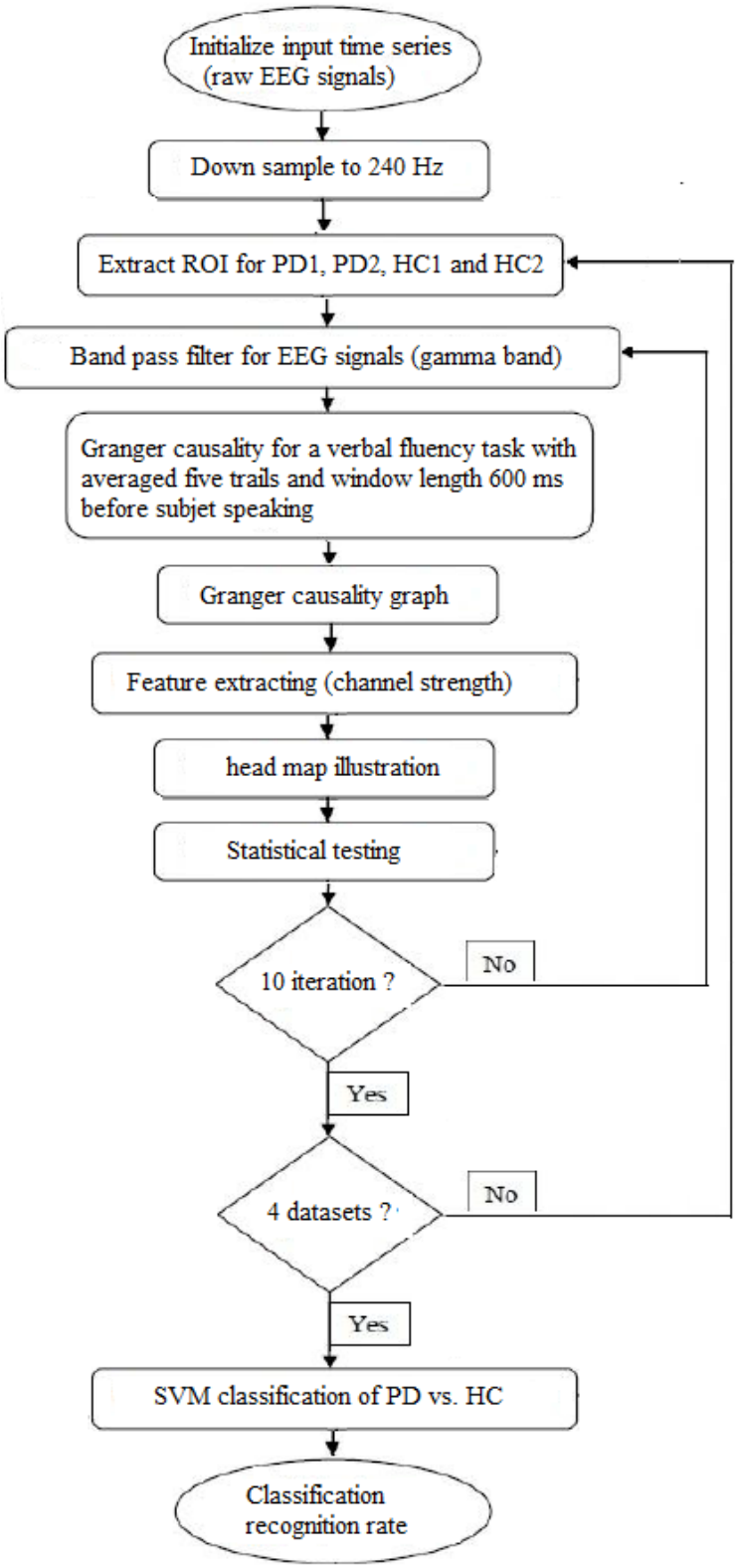
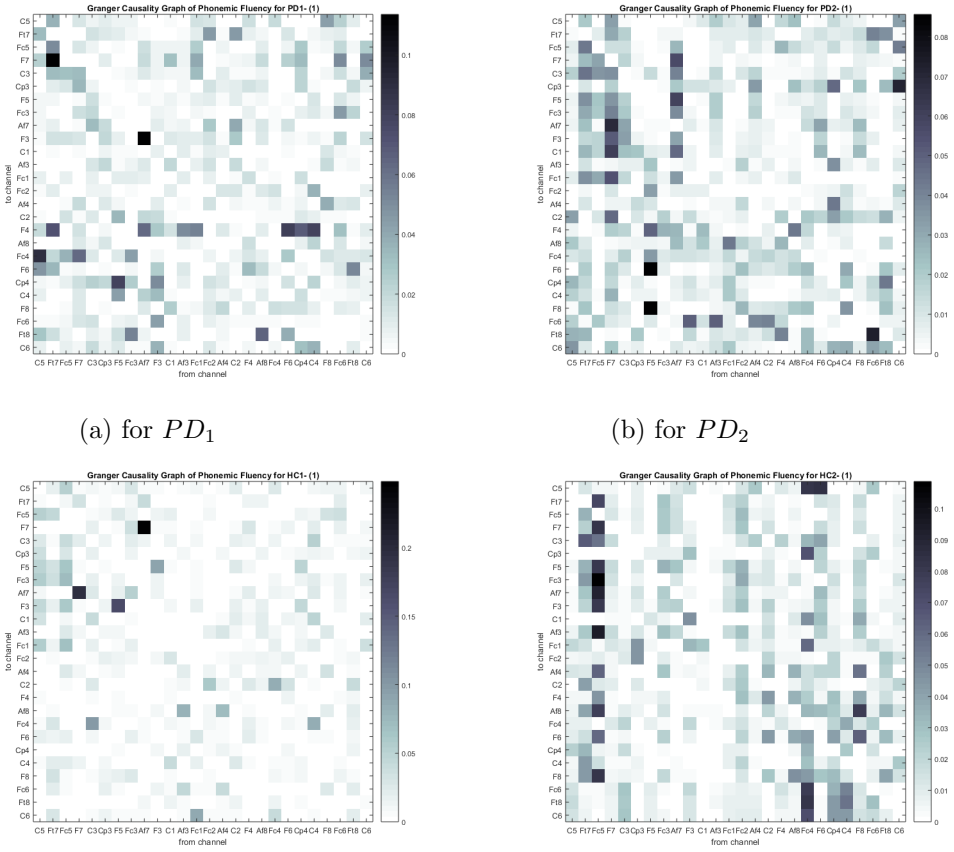


Figure 3.13: The flow chart of Granger causality analysis for verbal fluency tasks.

ms before speaking onset when the total number of iterations is ten computations. A sample of Granger causality computation is presented in Figure 3.14 for each dataset PD_1, PD_2, HC_1 and HC_2. The most causal channels in general for PD_1 and PD_2 is channel Ft7 for most of the Granger causality computations; however, the most causal channel for HC_1 and HC_2 is channel Fc5 for most of the computations as in Figure 3.14. The highest Granger causality channel strength for each iteration of PD patients is Ft7 and for healthy controls Fc5, as shown in Figure 3.15. Also, it is clear from the head maps that Ft7 is the most activated channel for PD patients and Fc5 channel is activated for healthy controls, as in Figure 3.16.

The second sample of computation for Granger causality graphs, channel strengths and head maps of PD_1, PD_2, HC_1 and HC_2 are illustrated in Figures 3.17, 3.18, and 3.19, respectively. The values of channel strength for the most causal channels in each Granger causality computation are listed in order to use them as feature classes of graphs for PD_1, PD_2, HC_1 and HC_2, as in the following Tables 3.1, 3.2, 3.3 and 3.4.



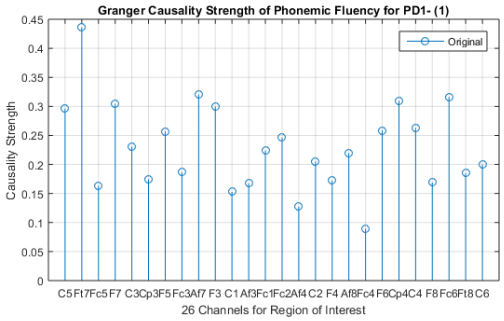
(a) for PD_1

(b) for PD_2

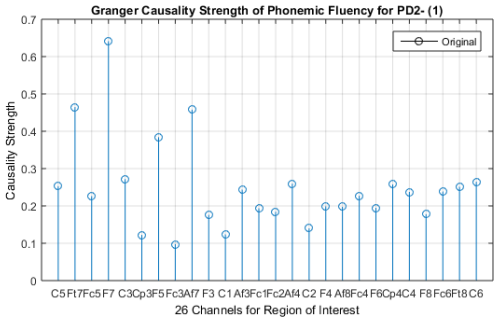
(c) for HC_1

(d) for HC_2

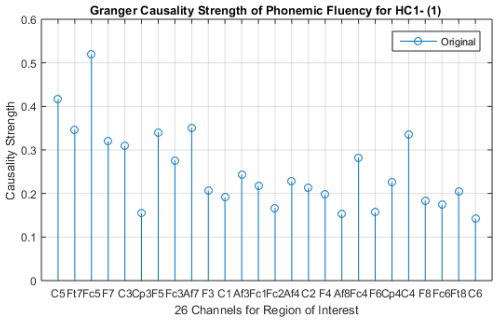
Figure 3.14: Sample Granger causality graphs for phonemic fluency shows a number of causalities in the Broca's area channels for PD_1 in channel Ft7 causing other channels, PD_2 in channel Ft7 and F7, HC_1 in channel Fc5 and HC_2 in channel Fc5.



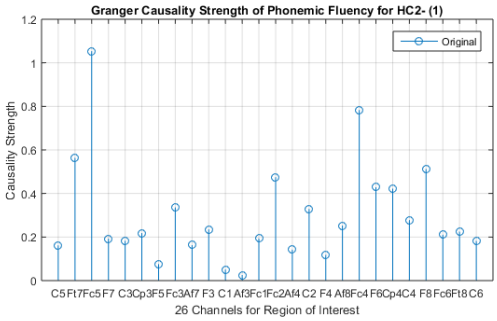
(a) for PD_1



(b) for PD_2

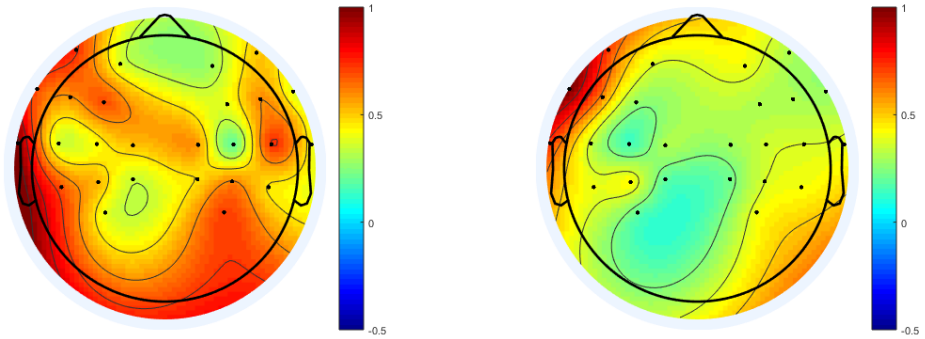


(c) for HC_1



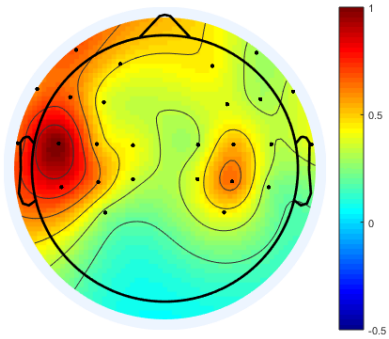
(d) for HC_2

Figure 3.15: Sample Granger causality channel strengths for phonemic fluency shows the causality strength of each channel in the Granger causality graph for each dataset.

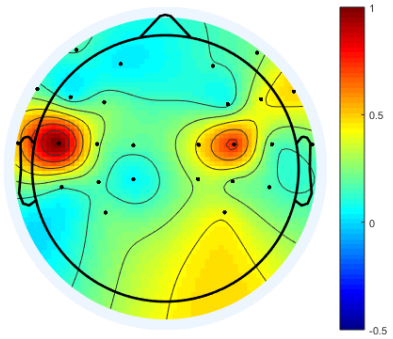


(a) for PD_1

(b) for PD_2



(c) for HC_1



(d) for HC_2

Figure 3.16: Sample Granger causality head maps for phonemic fluency shows the area of activation in the head map for each dataset during phonemic fluency.

No./Channel	Ft7	Fc5
1	0.449	0.16
2	0.54	0.21
3	0.325	0.24
4	0.385	0.32
5	0.445	0.266
6	0.395	0.36
7	0.285	0.24
8	0.125	0.24
9	0.21	0.16
10	0.21	0.17

Table 3.1: Channel strength values (Ft7 and Fc5) of phonemic fluency for PD_1.

No./Channel	Ft7	Fc5
1	0.46	0.23
2	0.63	0.21
3	0.65	0.285
4	0.18	0.15
5	0.3	0.49
6	0.28	0.6
7	0.4	0.33
8	0.15	0.39
9	0.35	0.1
10	0.198	0.245

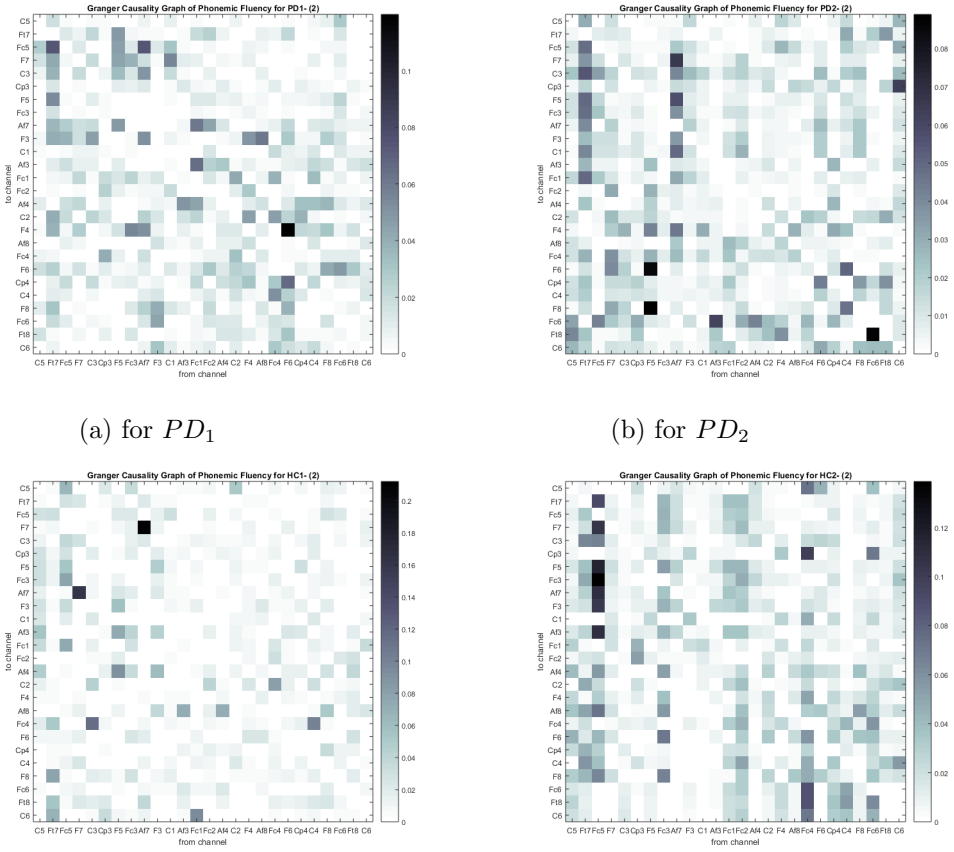
Table 3.2: Channel strength values (Ft7 and Fc5) of phonemic fluency for PD_2.

No./Channel	Ft7	Fc5
1	0.35	0.52
2	0.37	0.49
3	0.27	0.65
4	0.225	0.59
5	0.2	0.5
6	0.18	0.48
7	0.28	0.575
8	0.3	0.415
9	0.35	0.325
10	0.225	0.395

Table 3.3: Channel strength values (Ft7 and Fc5) of phonemic fluency for HC_1.

No./Channel	Ft7	Fc5
1	0.56	1.05
2	0.62	1.22
3	0.85	0.44
4	0.7	0.21
5	0.4	0.59
6	0.14	0.5
7	0.2	0.3
8	0.59	0.63
9	0.45	0.73
10	0.12	0.22

Table 3.4: Channel strength values (Ft7 and Fc5) of phonemic fluency for HC_2.

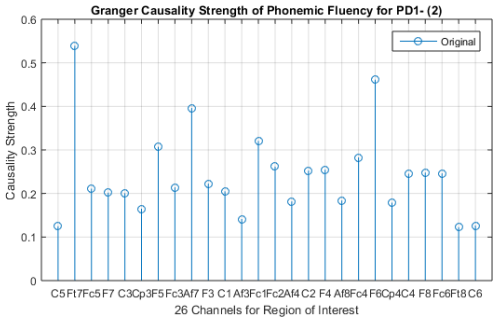


(a) for PD_1 (b) for PD_2

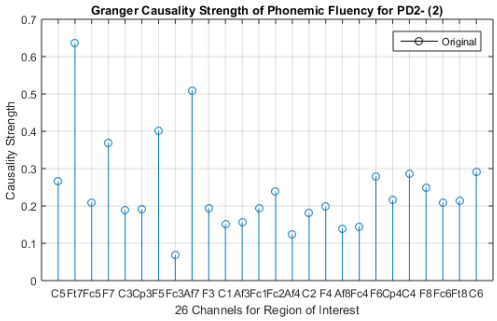
(c) for HC_1 (d) for HC_2

Figure 3.17: Sample Granger causality graphs for phonemic fluency shows a number of causalities in the Broca's area channels for PD_1 in channel Ft7 causing other channels, PD_2 in channel Ft7, HC_1 in channel Fc5 and HC_2 in channel Fc5.

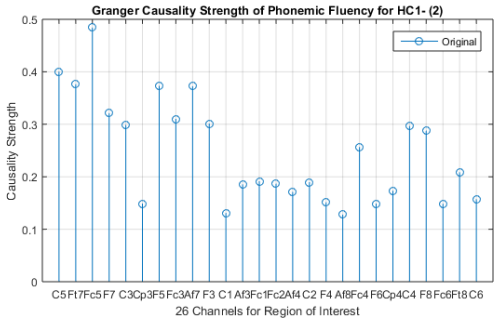
The statistical test for the models of Granger causality for phonemic fluency are given in the bootstrapping histograms for B=1000 times with replacing channel strength values of all the channels within the ten computations of a dataset, such as PD_1, PD_2, HC_1 and HC_2 as in Figure 3.20. The bootstrapping of the channel strengths shows the distribution of channel strength values for 1000 times. The distributions of the bootstrapping are assumed to be normally distributed for each dataset. The goal of the bootstrapping is to present the probability of each value of channel strength. The p-test threshold for the models of Granger causality is set as



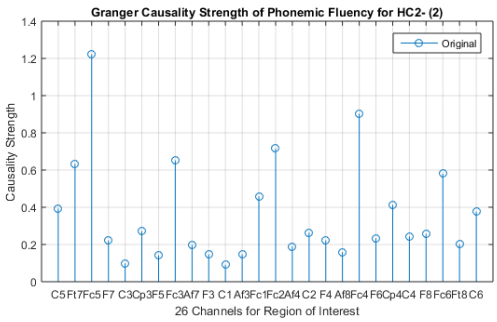
(a) for PD_1



(b) for PD_2

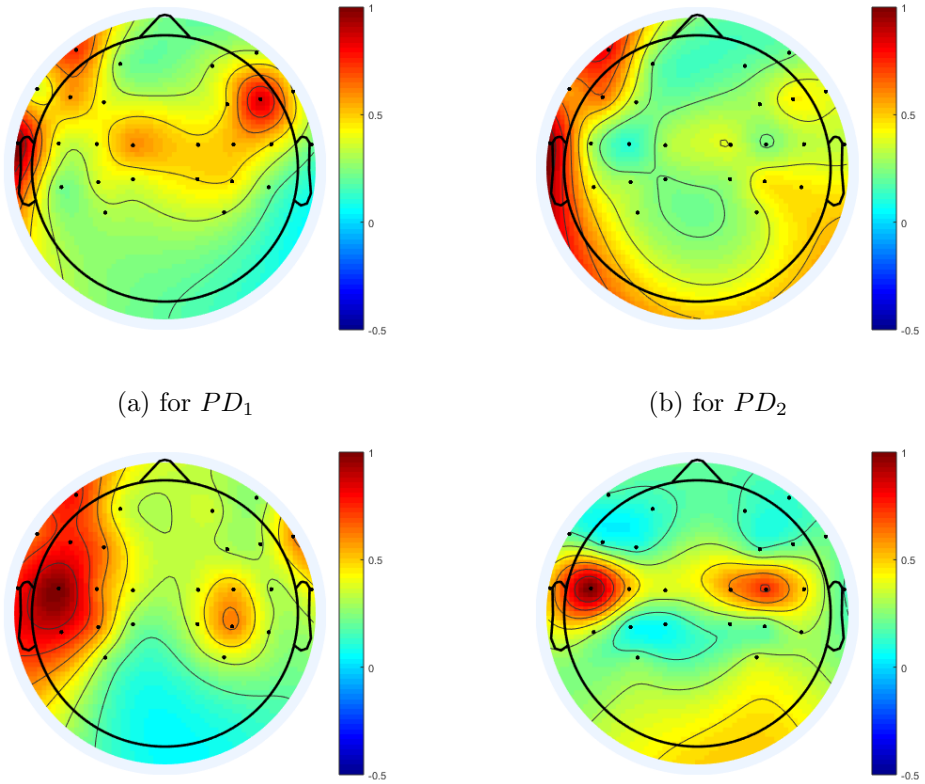


(c) for HC_1



(d) for HC_2

Figure 3.18: Sample Granger causality channel strength for phonemic fluency shows the causality strength of each channel in the Granger causality graph for each dataset.



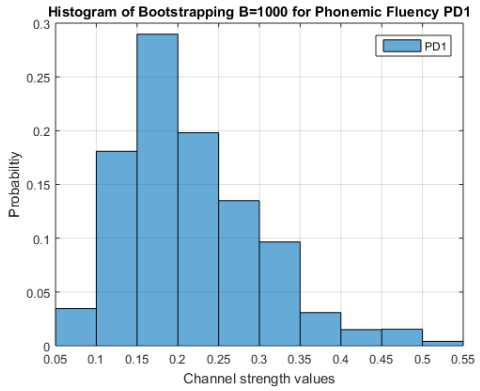
(a) for PD_1

(b) for PD_2

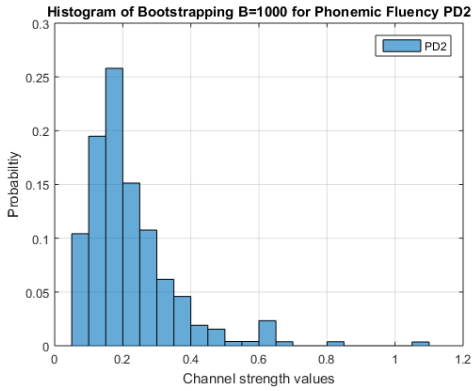
(c) for HC_1

(d) for HC_2

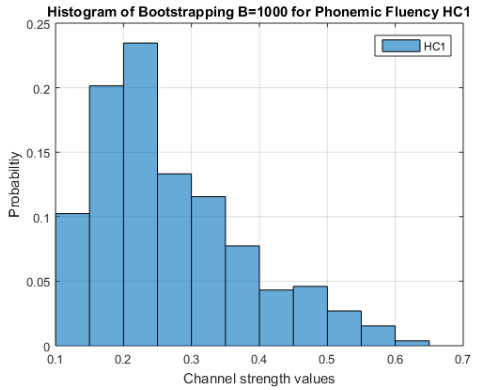
Figure 3.19: Sample Granger causality head maps for phonemic fluency shows the area of activation in the head map for each dataset during phonemic fluency.



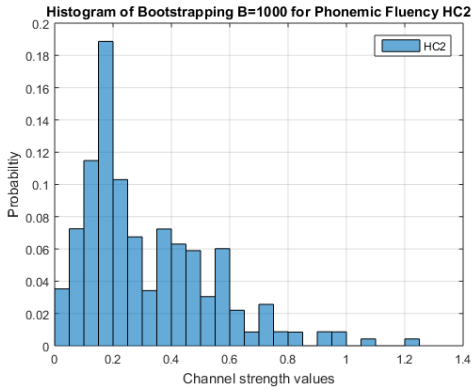
(a) for PD_1



(b) for PD_2



(c) for HC_1



(d) for HC_2

Figure 3.20: Histogram of bootstrapping for phonemic fluency of all the datasets. The distribution shows the probability of the channel strength values.

$\alpha = 0.05$, and the p-value should be less than or equal to the significance level α. The null hypothesis H_0 of the Granger causality statistical test is assumed as "no differences between channel strengths values under test and all the channel strength in each dataset". On other hand, the assumption of the alternative hypothesis H_1 is related to "differences between the channel strength values under test and all other values in the dataset". The statistical test is one sided of the bootstrapping for Granger causality analysis where the vertical coordinates are the probability values which are computed under the null hypothesis for the different outcomes of channel strengths in the horizontal coordinates as in Figures 3.20. The statistical significance p-value is the area which is less than or equal to the significance level α. The p-value of the channel Ft7 values for PD_1 and PD_2 is significant where the p-value ≤ 0.05. For instant, Ft7 channel strength value of PD_1 of the first sample is 0.449 where this value in the bootstrapping histogram of PD_1 has its p-value is $p \leq 0.05$ which is considered as statistical significant and the null hypothesis is rejected. Furthermore, the p-value of the channel Fc5 values for HC_1 and HC_2 is significant where the its p-value is $p \leq 0.05$. For example, Fc5 channel strength value for HC_1 of the first sample is 0.52 where this value in the bootstrapping histogram of HC_1 has its p-value is $p \leq 0.05$ which is considered as statistical significant and the null hypothesis is rejected.

Semantic Fluency

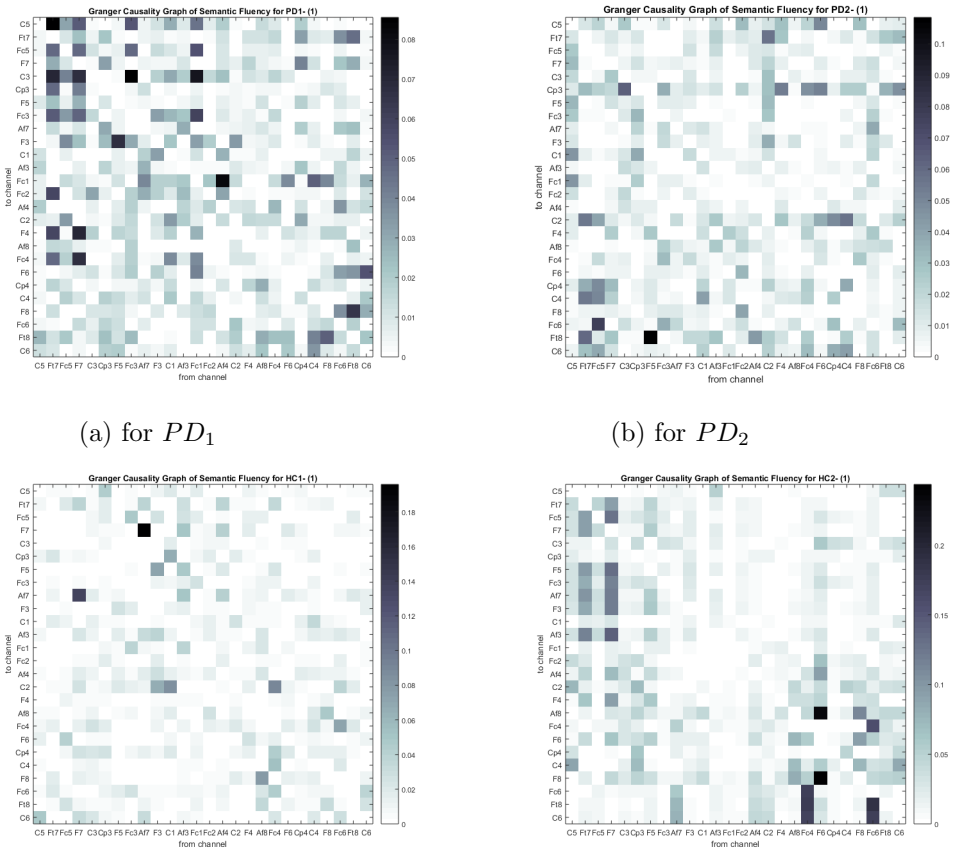
The performance of semantic fluency is based on participants generating as many words as as possible that are categorized with a given category (for instance, animals) in 60 s. The computation of Granger causality is based on the average of five trials and 600 ms before speaking onset when the total number of iterations is ten computations. A sample of Granger causality computation for each dataset PD_1, PD_2, HC_1 and HC_2 is presented Figure 3.21. The finding from the results is

No./Channel	Ft7	F7
1	0.61	0.575
2	0.62	0.485
3	0.68	0.71
4	0.43	0.38
5	0.35	0.45
6	0.55	0.43
7	0.28	0.425
8	0.25	0.35
9	0.3	0.45
10	0.27	0.49

Table 3.5: Channel strength values (Ft7 and F7) of semantic fluency for PD_1.

that the most causal channel causing other channels for PD_1 and PD_2 is channel Ft7 for most of the computations; however, the most causal channel for HC_1 and HC_2 is channel F7 for most of the computations, as observed in Figure 3.21. As well, the highest Granger causality channel strength for each computation of PD patients is Ft7 and for healthy controls F7, as shown in Figure 3.22. From the head maps, it is evident that Ft7 is the most activated channel for PD patients, and the F7 channel is activated for healthy controls, as in Figure 3.23.

The second sample of computation for Granger causality graphs, channel strengths, and head maps of PD_1, PD_2, HC_1 and HC_2 are illustrated in Figures 3.24, 3.25 and 3.26, respectively. The values of channel strength for the most causal channels are listed in order to use them as feature classes of graphs for PD_1, PD_2, HC_1 and HC_2 in the following Tables 3.5, 3.6, 3.7 and 3.8.



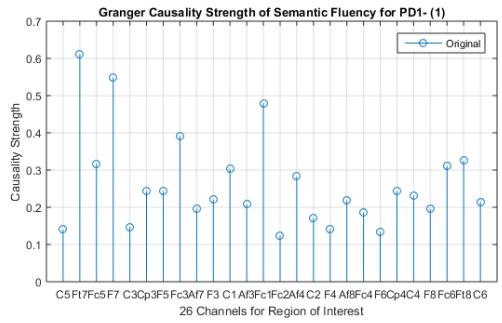
(a) for PD_1

(b) for PD_2

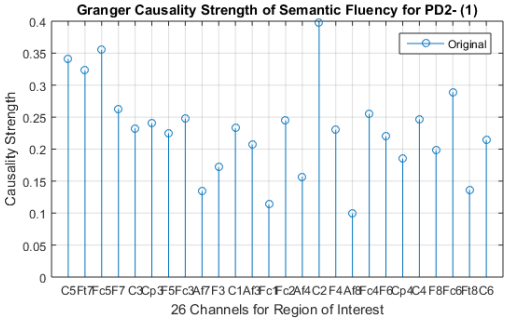
(c) for HC_1

(d) for HC_2

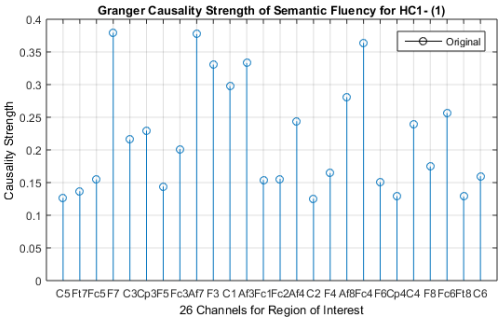
Figure 3.21: Sample Granger causality graphs for semantic fluency shows a number of causalities in the Broca's area channels for PD_1 in channel Ft7 causing other channels, PD_2 in channel Ft7, HC_1 in channel F7 and HC_2 in channel F7.



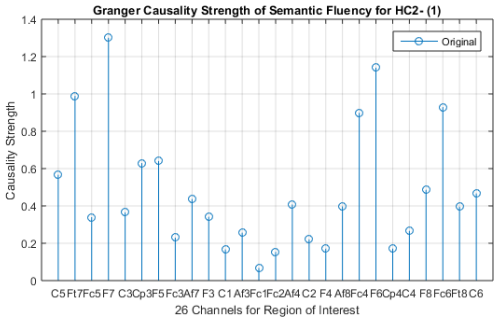
(a) for PD_1



(b) for PD_2



(c) for HC_1

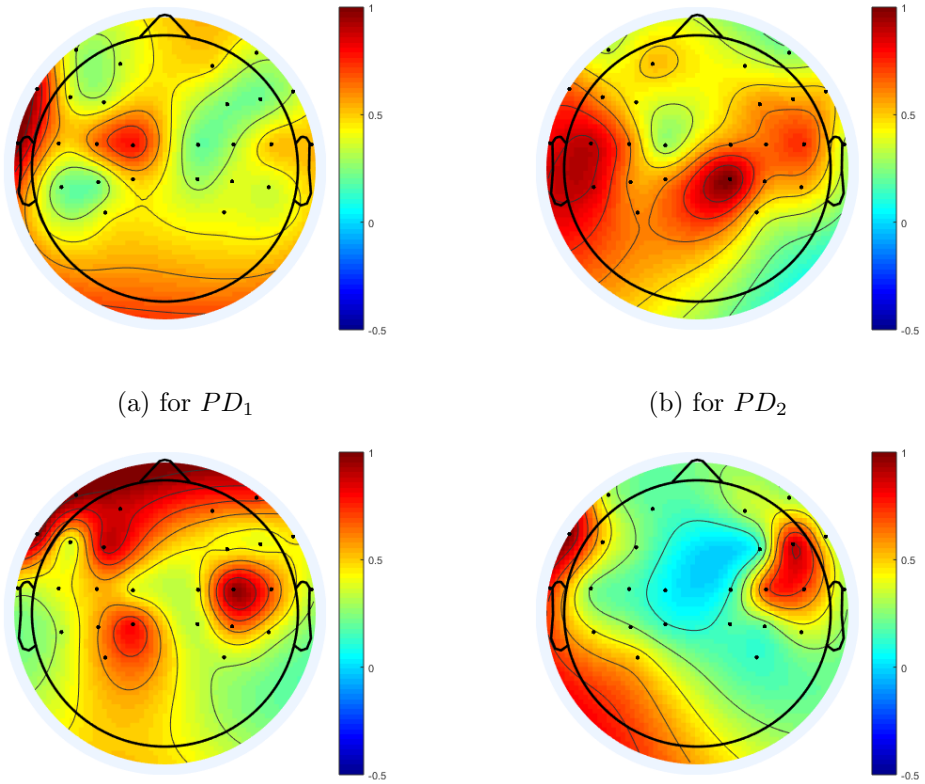


(d) for HC_2

Figure 3.22: Sample Granger causality channel strengths for semantic fluency shows the causality strength of each channel in the Granger causality graph for each dataset.

No./Channel	Ft7	F7
1	0.35	0.26
2	0.348	0.155
3	0.13	0.76
4	0.225	0.29
5	0.15	0.25
6	0.9	1.1
7	0.11	0.1
8	0.2	0.095
9	0.11	0.1
10	0.35	0.16

Table 3.6: Channel strength values (Ft7 and F7) of semantic fluency for PD_2.



(a) for PD_1

(b) for PD_2

(c) for HC_1

(d) for HC_2

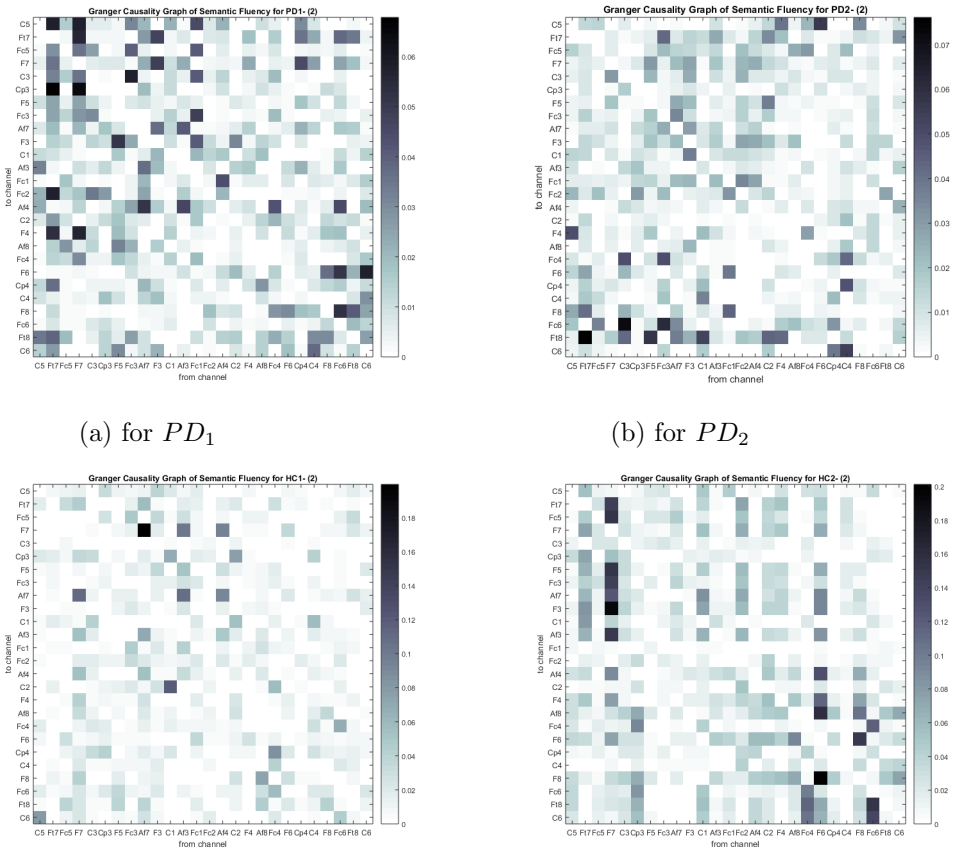
Figure 3.23: Sample Granger causality head maps for semantic fluency shows the area of activation in the head map for each dataset during semantic fluency.

No./Channel	Ft7	F7
1	0.14	0.36
2	0.12	0.6
3	0.46	0.36
4	0.2	0.4
5	0.205	0.22
6	0.25	0.425
7	0.245	0.252
8	0.24	0.3
9	0.17	0.225
10	0.29	0.395

Table 3.7: Channel strength values (Ft7 and F7) of semantic fluency for HC_1.

No./Channel	Ft7	F7
1	0.3	0.6
2	0.1	1.3
3	0.88	0.98
4	1.2	0.48
5	1.78	0.9
6	1.45	1.35
7	0.81	1.59
8	0.61	0.9
9	0.46	0.95
10	0.12	0.35

Table 3.8: Channel strength values (Ft7 and F7) of semantic fluency for HC_2.



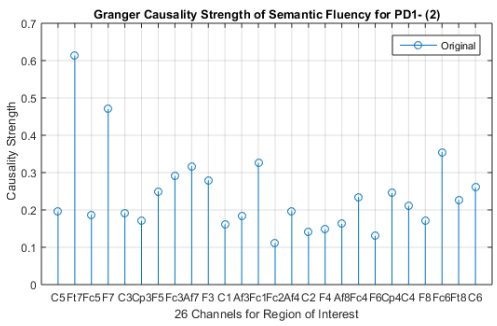
(a) for PD_1

(b) for PD_2

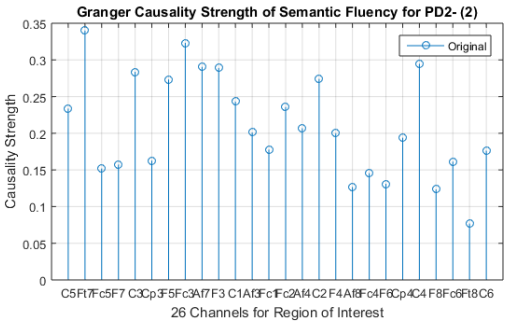
(c) for HC_1

(d) for HC_2

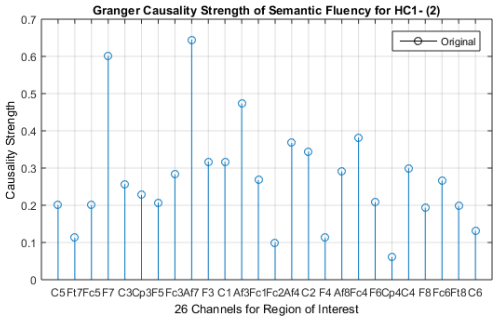
Figure 3.24: Sample Granger causality graphs for semantic fluency shows a number of causalities in the Broca's area channels for PD_1 in channels Ft7 and F7 causing other channels, PD_2 in channel Ft7, HC_1 in channel F7 and HC_2 in channel F7.



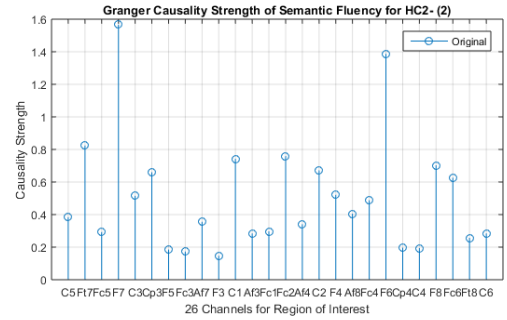
(a) for PD_1



(b) for PD_2

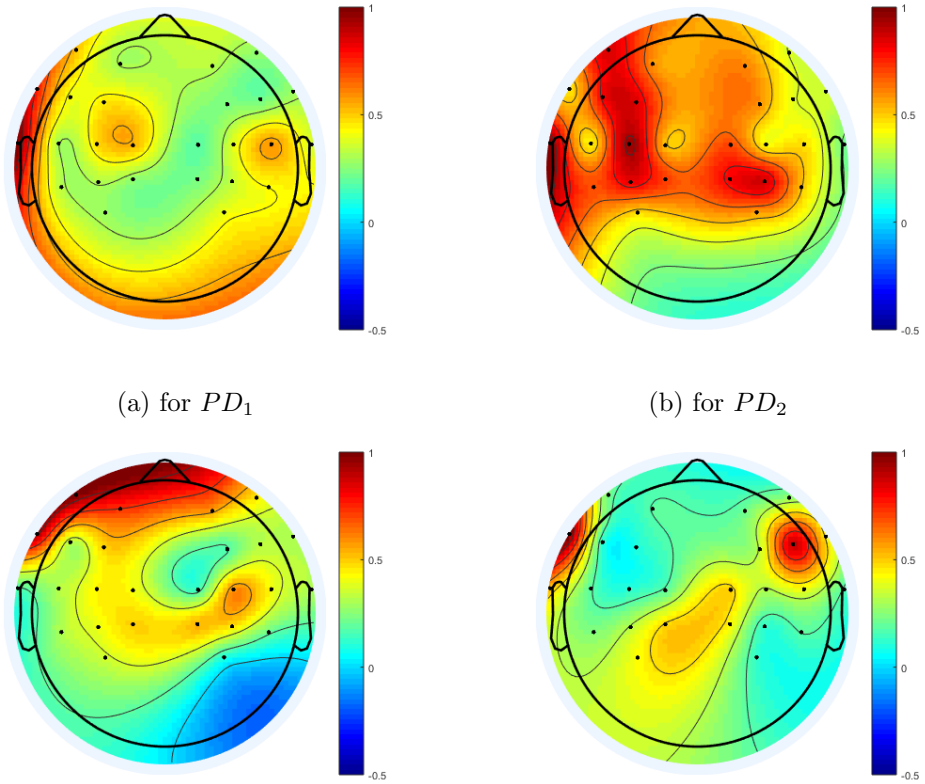


(c) for HC_1



(d) for HC_2

Figure 3.25: Sample Granger causality channel strength for semantic fluency shows the causality strength of each channel in the Granger causality graph for each dataset.



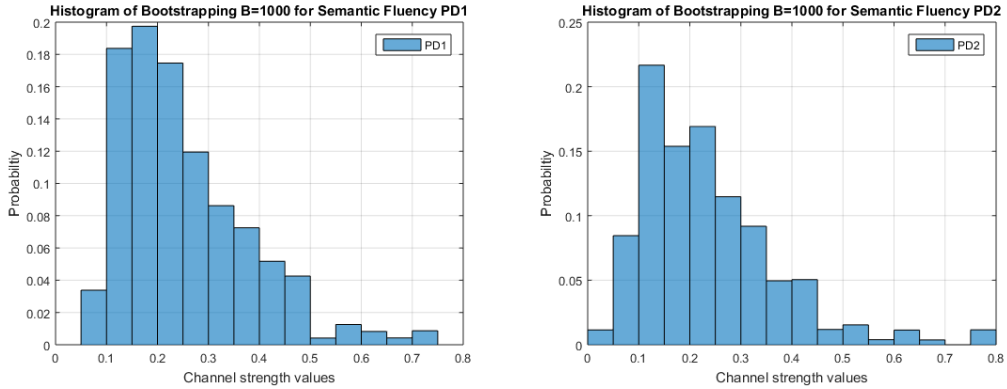
(a) for PD_1

(b) for PD_2

(c) for HC_1

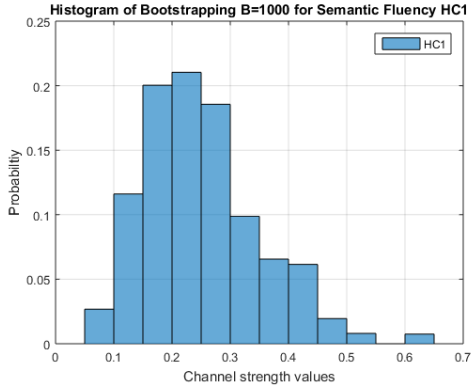
(d) for HC_2

Figure 3.26: Sample Granger causality head maps for semantic fluency shows the area of activation in the head map for each dataset during semantic fluency.

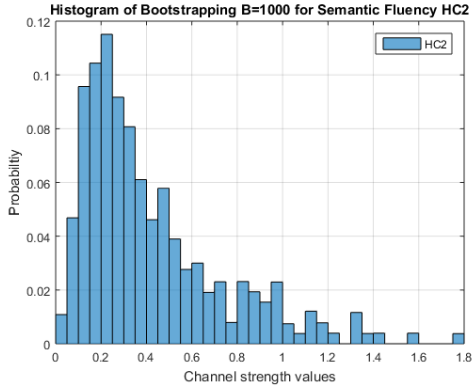


(a) for PD_1

(b) for PD_2



(c) for HC_1



(d) for HC_2

Figure 3.27: Histogram of bootstrapping for semantic fluency of all the datasets. The distribution shows the probability of the channel strength values.

The statistical test for the models of Granger causality for semantic fluency are given in the bootstrapping histograms for B=1000 times with replacing channel strength values of all the channels within the ten computations of a dataset, such as PD_1, PD_2, HC_1 and HC_2 as in Figure 3.27. The bootstrapping of the channel strengths shows the distribution of channel strength values for 1000 times. The distributions of the bootstrapping are assumed to be normally distributed for each dataset. The goal of the bootstrapping is to present the probability of each value of channel strength. The p-test threshold for the models of Granger causality for semantic fluency is set as $\alpha = 0.05$, and the p-value should be less than or equal to

the significance level α. The null hypothesis H_0 of the Granger causality statistical test is assumed as "no differences between channel strengths values under test and all the channel strength in each dataset". On other hand, the assumption of the alternative hypothesis H_1 is related to "differences between the channel strength values under test and all other values in the dataset". The statistical test is one sided of the bootstrapping for Granger causality analysis where the vertical coordinates are the probability values which are computed under the null hypothesis for the different outcomes of channel strengths in the horizontal coordinates as in Figures 3.27. The statistical significance p-value is the area which is less than or equal to the significance level α. The p-value of the channel Ft7 values for PD_1 and PD_2 is significant with the p-value ≤ 0.05. For example, Ft7 channel strength value for PD_1 of the second sample is 0.62 where this value in the bootstrapping histogram of HC_1 has its p-value ≤ 0.05 which is considered as statistical significant and the null hypothesis is rejected. Furthermore, the p-value of the channel F7 values for HC_1 and HC_2 is significant with the p-value ≤ 0.05. For example, F7 channel strength value for HC_1 of the second sample is 0.6 where this value in the bootstrapping histogram of HC_1 has its p-value ≤ 0.05 which is considered as statistical significant and the null hypothesis is rejected.

Category Semantic Fluency

The second type of semantic fluency is category semantic fluency where the performance of category semantic fluency is based on participants abilities to generate words that are categorized into different given areas and then switched between them (example: animals to jungle animals); subjects state as many words as they can in 60 s. The computation of Granger causality is based on the average of five trials and 600 ms before speaking onset when the total number of iterations is ten computations. A sample of Granger causality computation for each dataset PD_1,

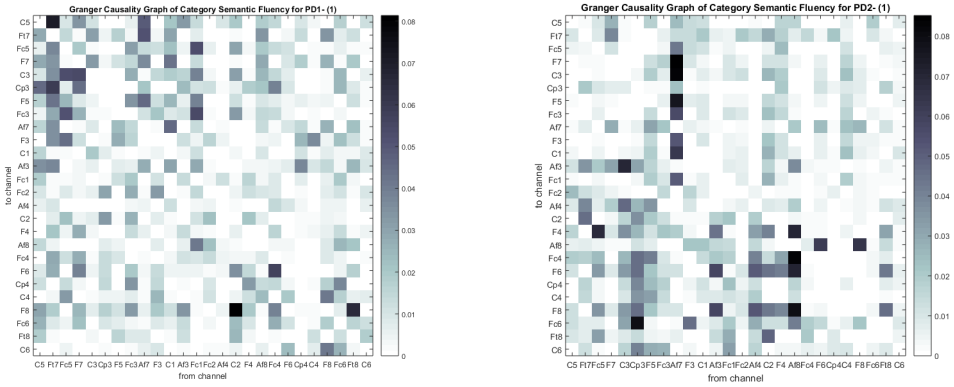
No./Channel	Ft7	F7
1	0.55	0.36
2	0.435	0.25
3	1.1	0.68
4	0.68	0.33
5	0.45	0.43
6	0.43	0.45
7	0.26	0.51
8	1.05	0.6
9	0.93	0.45
10	0.8	0.54

Table 3.9: Channel strength values (Ft7 and F7) of category semantic fluency for PD_1.

PD_2, HC_1 and HC_2 is presented in Figure 3.28, establishing that the most causal channel with other channels for PD_1 and PD_2 is channel Ft7 for most of the iterations; however, the most causal channel for HC_1 and HC_2 is channel F7 for most of the iterations, as in sample Figure 3.28. Additionally, the highest Granger causality channel strength for each iteration of PD patients are Ft7 and for healthy controls F7, as demonstrated in Figure 3.29. It is clear from the head maps that Ft7 is the channel that is more often activated for PD patients, and the F7 channel is more often activated for healthy controls, as in Figure 3.30.

The second sample of computation for Granger causality graphs, channel strengths, and head maps of PD_1, PD_2, HC_1 and HC_2 are illustrated in Figures 3.31, 3.32 and 3.33, respectively. The values of channel strength for the most causal channels are listed in order to use them as feature classes of graphs for PD_1, PD_2, HC_1 and HC_2 as in the following Tables 3.9, 3.10, 3.11 and 3.12.

The statistical test for the models of Granger causality for category semantic fluency are given in the bootstrapping histograms for B=1000 times with replacing



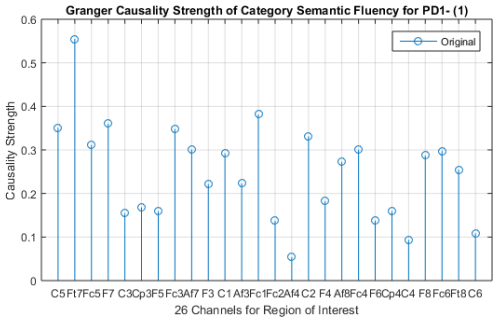
(a) for PD_1

(b) for PD_2

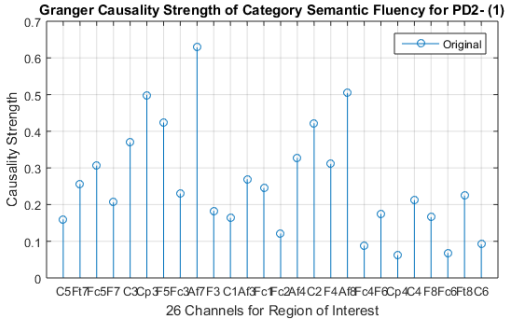
(c) for HC_1

(d) for HC_2

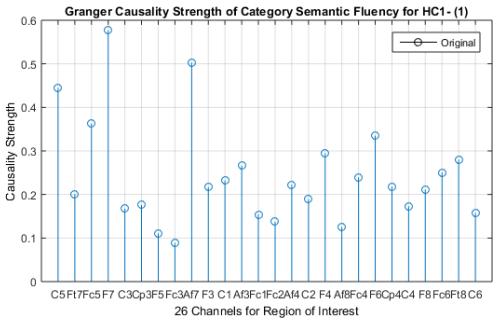
Figure 3.28: Sample Granger causality graphs for category semantic fluency shows a number of causalities in the Broca's area channels for PD_1 in channel Ft7 causing other channels, PD_2 in channel Ft7 and Af7, HC_1 in channel F7 and HC_2 in channel Ft7 and F7.



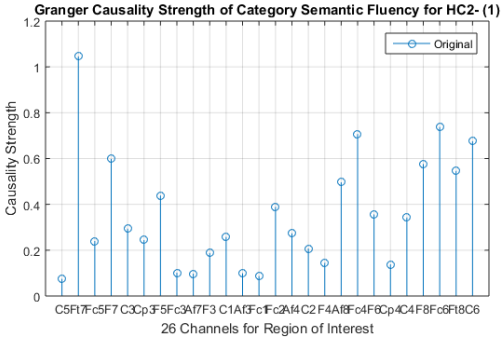
(a) for PD_1



(b) for PD_2



(c) for HC_1

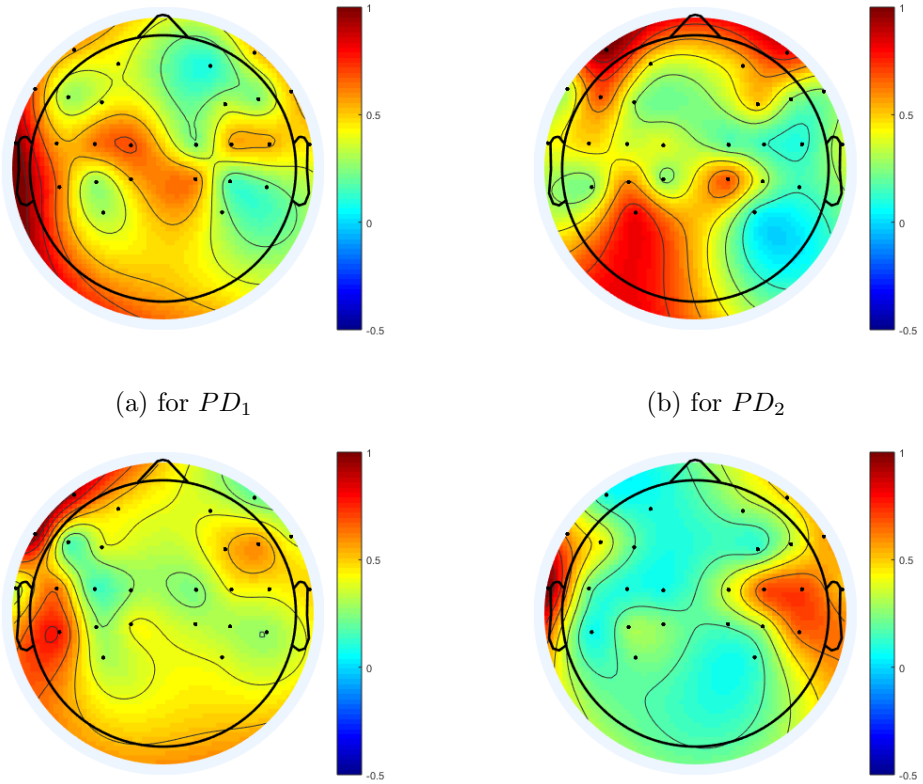


(d) for HC_2

Figure 3.29: Sample Granger causality channel strengths for category semantic fluency shows the causality strength of each channel in the Granger causality graph for each dataset.

No./Channel	Ft7	F7
1	0.26	0.2
2	0.59	0.28
3	0.44	0.19
4	0.1	0.22
5	0.2	0.275
6	0.3	0.13
7	0.205	0.395
8	0.108	0.44
9	0.16	0.54
10	0.26	0.65

Table 3.10: Channel strength values (Ft7 and F7) of category semantic fluency for PD_2.



(a) for PD_1

(b) for PD_2

(c) for HC_1

(d) for HC_2

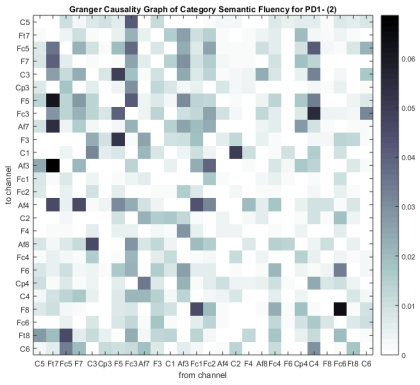
Figure 3.30: Sample Granger causality head maps for category semantic fluency shows the area of activation in the head map for each dataset during category semantic fluency.

No./Channel	Ft7	F7
1	0.2	0.58
2	0.22	0.57
3	0.1	0.75
4	0.18	0.53
5	0.1	0.6
6	0.19	0.495
7	0.2	0.42
8	0.14	0.34
9	0.175	0.4
10	0.1	0.53

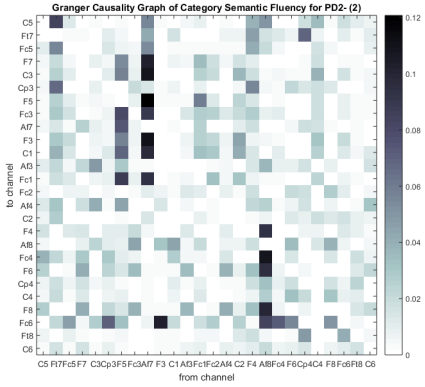
Table 3.11: Channel strength values (Ft7 and F7) of category semantic fluency for HC_1.

No./Channel	Ft7	F7
1	1.05	0.6
2	0.34	0.54
3	0.2	0.35
4	0.9	0.7
5	0.5	0.4
6	0.9	0.41
7	0.26	0.73
8	0.63	0.3
9	0.2	0.5
10	0.7	0.8

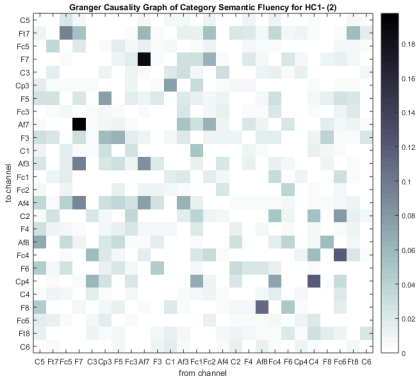
Table 3.12: Channel strength values (Ft7 and F7) of category semantic fluency for HC_2.



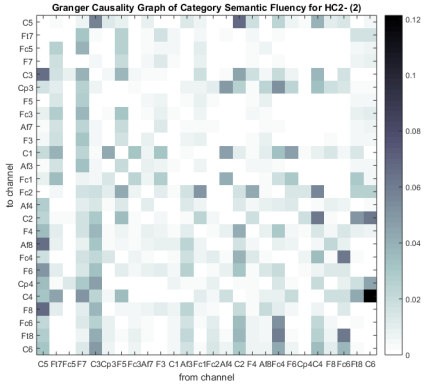
(a) for PD_1



(b) for PD_2

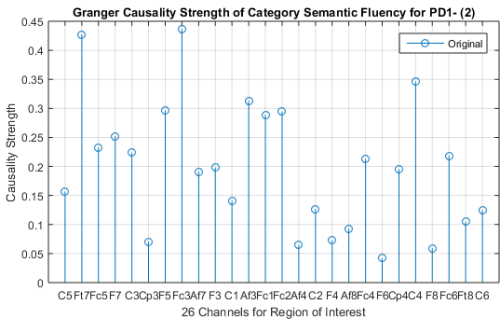


(c) for HC_1

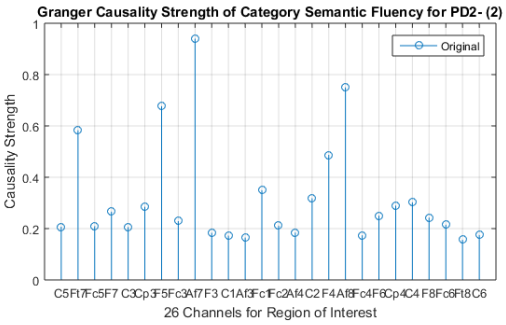


(d) for HC_2

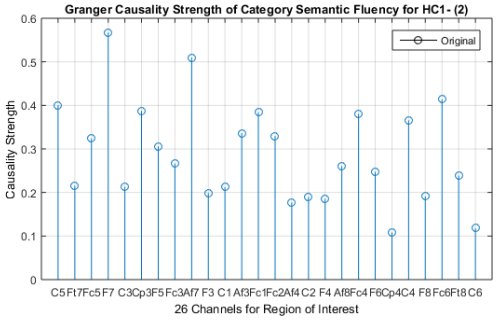
Figure 3.31: Sample Granger causality graphs for category semantic fluency shows a number of causalities in the Broca's area channels for PD_1 in channel Ft7 causing other channels, PD_2 in channel Ft7, HC_1 in channel F7 and HC_2 in channel F7.



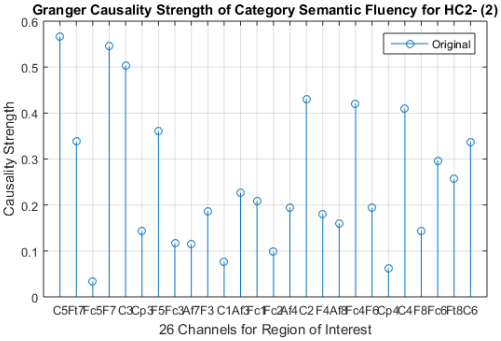
(a) for PD_1



(b) for PD_2

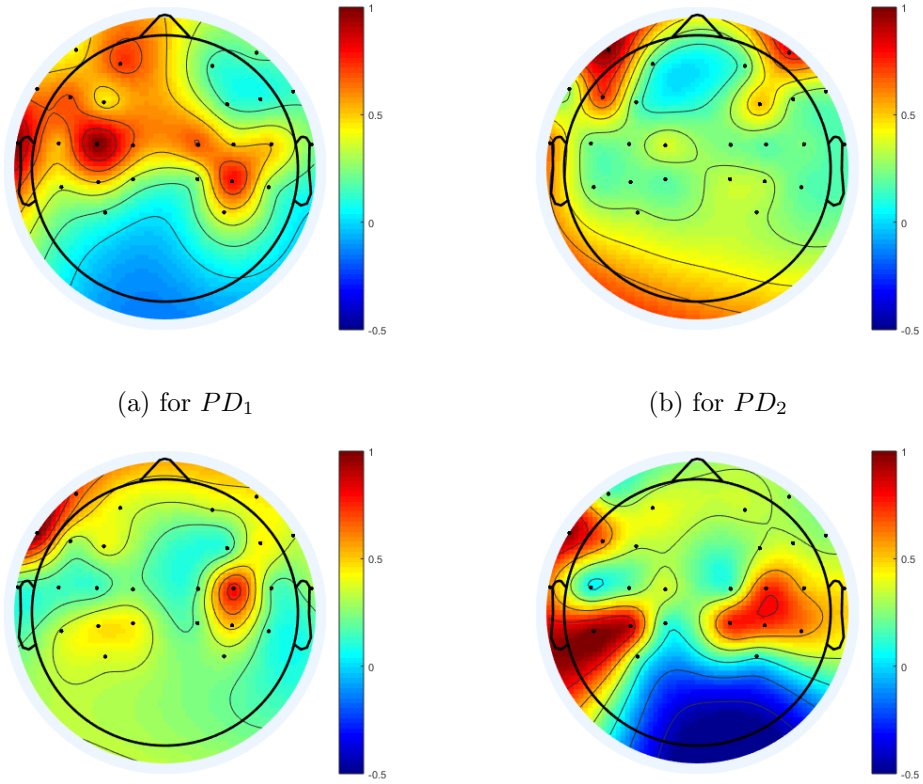


(c) for HC_1



(d) for HC_2

Figure 3.32: Sample Granger causality channel strength for category semantic fluency shows the causality strength of each channel in the Granger causality graph for each dataset.

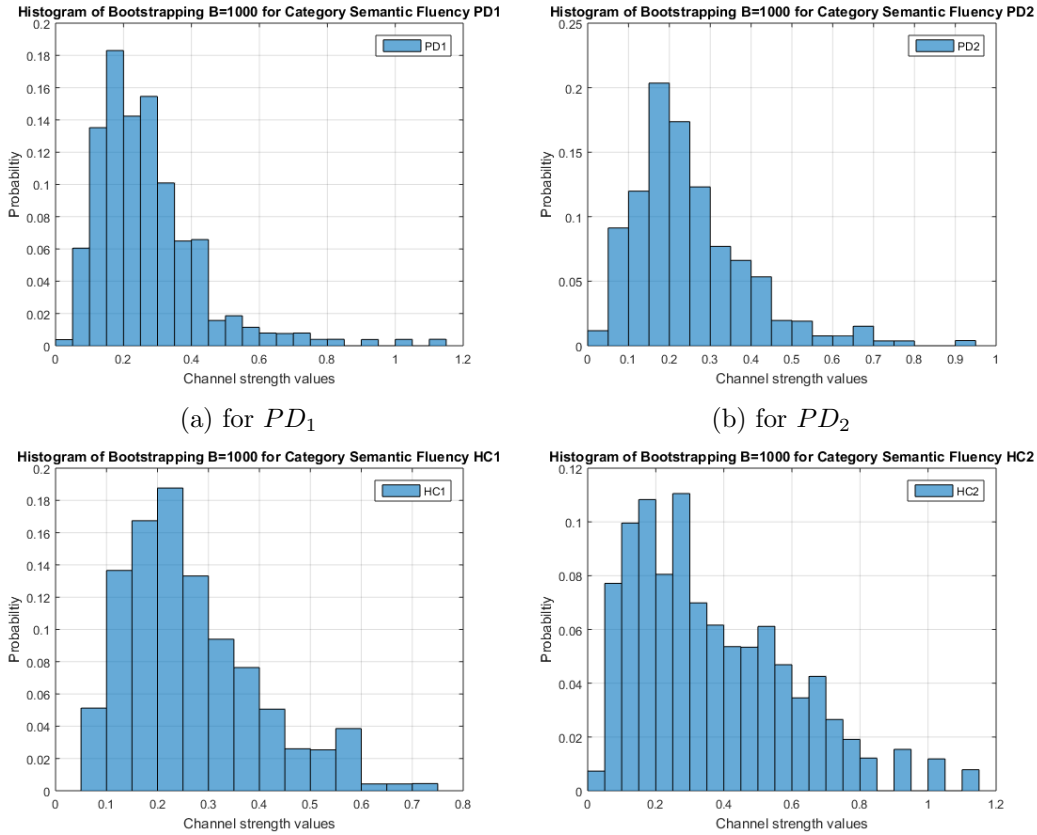


(a) for PD_1 (b) for PD_2

(c) for HC_1 (d) for HC_2

Figure 3.33: Sample Granger causality head maps for category semantic fluency shows the area of activation in the head map for each dataset during category semantic fluency.

channel strength values of all the channels within the ten computations of a dataset such as, PD_1, PD_2, HC_1 and HC_2 as in Figure 3.34. The bootstrapping of the channel strengths shows the distribution of channel strength values for 1000 times. The distributions of the bootstrapping are assumed to be normally distributed for each dataset. The goal of the bootstrapping is to present the probability of each value of channel strength. The p-test threshold for the models of Granger causality is set as $\alpha = 0.05$, and the p-value should be less than or equal to the significance level α. The null hypothesis H_0 of the Granger causality statistical test is assumed as "no differences between channel strengths values under test all the channel strength



(a) for PD_1

(b) for PD_2

(c) for HC_1

(d) for HC_2

Figure 3.34: Histogram of bootstrapping for category semantic fluency of all the datasets. The distribution shows the probability of the channel strength values.

in each dataset". On other hand, the assumption of the alternative hypothesis H_1 is related to "differences between the channel strength values under test and of all other values in the dataset". The statistical test is one sided of the bootstrapping for Granger causality analysis where the vertical coordinates are the probability values which are computed under the null hypothesis for the different outcomes of channel strengths in the horizontal coordinates as in Figures 3.34. The statistical significance p-value is the area which is less than or equal to the significance level α. The p-value of the channel Ft7 values for PD_1 and PD_2 is significant with the p-value is $p \leq 0.05$. For instant, Ft7 channel strength value for PD_1 of the first sample is 0.55 where this value in the bootstrapping histogram of PD_1 has its p-value≤ 0.05 which is considered as statistical significant and the null hypothesis is rejected. Furthermore, the p-value of the channel F7 values for HC_1 and HC_2 is significant with the p-value is $p \leq 0.05$. For example, F7 channel strength value for HC_1 of the first sample is 0.58 where this value in the bootstrapping histogram of HC_1 has its p-value ≤ 0.05 which is considered as statistical significant and the null hypothesis is rejected.

Reading Fluency

The last type of verbal fluency tests is reading fluency where its performance is based on subjects reading as much of a given text as possible in 60s. The computation of Granger causality is based on the average of five trials and 600 ms before speaking onset when the total number of iterations is ten computations. A sample of Granger causality computation for each dataset PD_1, PD_2, HC_1 and HC_2 is presented in Figure 3.35. The most causal channel causing other channels for PD_1 and PD_2 is channel Ft7 for most of the computations; however, the most causal channel for HC_1 and HC_2 is channel F7 for most of the iterations, as in Figure 3.35. Also, the highest Granger causality channel strength for each iteration of PD patients is

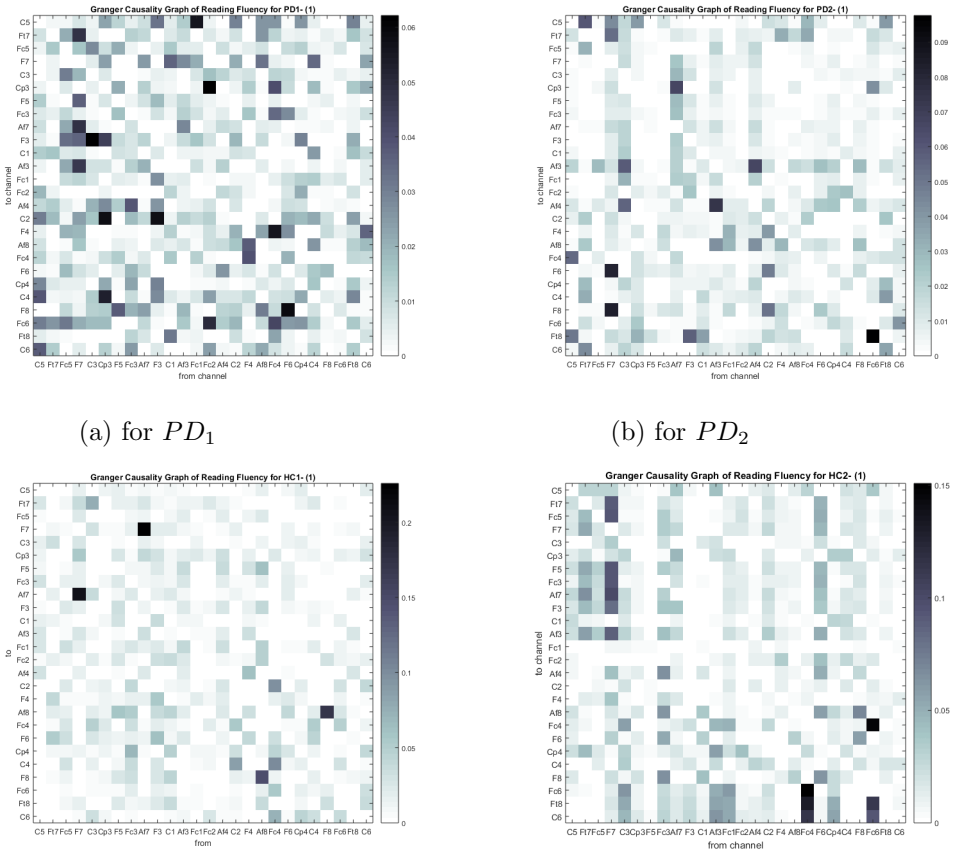
No./Channel	Ft7	F7
1	0.14	0.39
2	0.2	0.6
3	0.2	0.63
4	0.18	0.22
5	0.2	0.18
6	0.13	0.2
7	0.26	0.325
8	0.16	0.26
9	0.125	0.34
10	0.11	0.22

Table 3.13: Channel strength values (Ft7 and F7) of reading fluency for PD_1.

Ft7 and for healthy controls, F7, as shown in Figure 3.36. It is also evident from the head maps that Ft7 is a more frequently activated channel for PD patients, and F7 channel is more often activated for healthy controls, as evident in Figure 3.37.

The second sample of computation for Granger causality graphs, channel strengths, and head maps of PD_1, PD_2, HC_1 and HC_2 are illustrated in Figures 3.38, 3.39, and 3.40 respectively. In regards to the rest of the results, the values of channel strength for the most causal channels are presented in order to use them as feature classes of graphs for PD_1, PD_2, HC_1 and HC_2 in the following Tables 3.13, 3.14, 3.15 and 3.16.

The statistical test for the models of Granger causality for reading fluency are given in the bootstrapping histograms for B=1000 times with replacing channel strength values of all the channels within the ten computations of a dataset such as PD_1, PD_2, HC_1 and HC_2 as in Figure 3.41. The bootstrapping of the channel strengths shows the distribution of channel strength values for 1000 times. The distributions of the bootstrapping are assumed to be normally distributed for each dataset. The goal of the bootstrapping is to present the probability of each value of



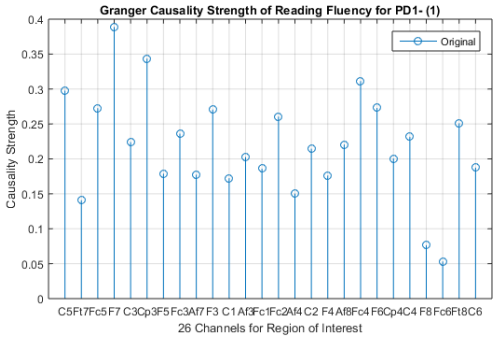
(a) for PD_1

(b) for PD_2

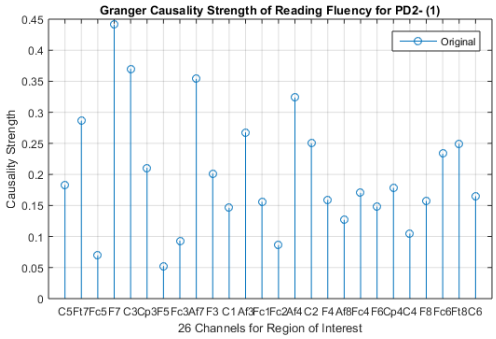
(c) for HC_1

(d) for HC_2

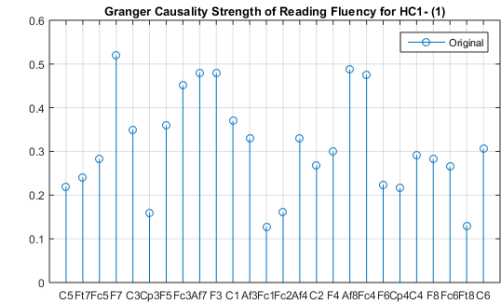
Figure 3.35: Sample Granger causality graphs for reading fluency shows a number of causalities in the Broca's area channels for PD_1 in channel F7 causing other channels, PD_2 in channel Ft7 and F7, HC_1 in channel F7 and HC_2 in channel Ft7 and F7.



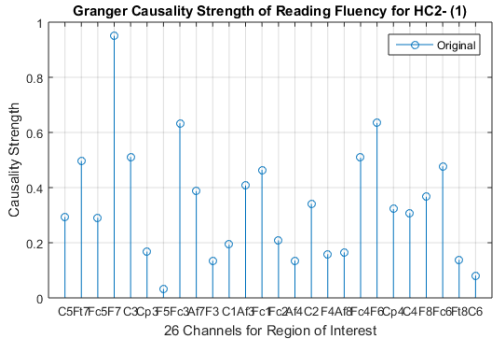
(a) for PD_1



(b) for PD_2



(c) for HC_1

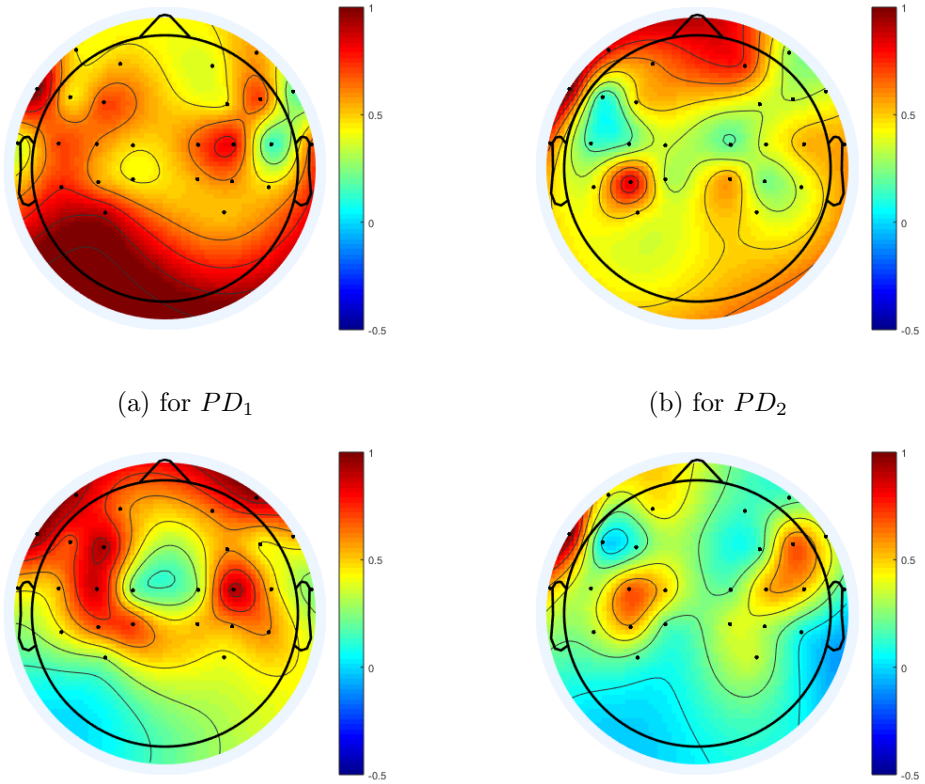


(d) for HC_2

Figure 3.36: Sample Granger causality channel strengths for reading fluency shows the causality strength of each channel in the Granger causality graph for each dataset.

No./Channel	Ft7	F7
1	0.285	0.449
2	0.09	0.47
3	0.195	0.448
4	0.12	0.37
5	0.225	0.39
6	0.22	0.3
7	0.137	0.29
8	0.27	0.2
9	0.34	0.195
10	0.272	0.454

Table 3.14: Channel strength values (Ft7 and F7) of reading fluency for PD_2.



(a) for PD_1

(b) for PD_2

(c) for HC_1

(d) for HC_2

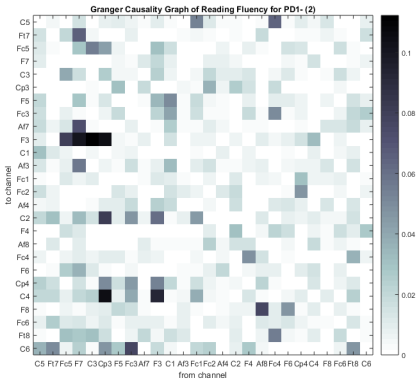
Figure 3.37: Sample Granger causality head maps for reading fluency shows the area of activation in the head map for each dataset during reading fluency.

No./Channel	Ft7	F7
1	0.24	0.52
2	0.125	0.495
3	0.145	0.41
4	0.2	0.45
5	0.25	0.7
6	0.4	0.83
7	0.56	0.48
8	0.22	0.33
9	0.35	0.55
10	0.23	0.46

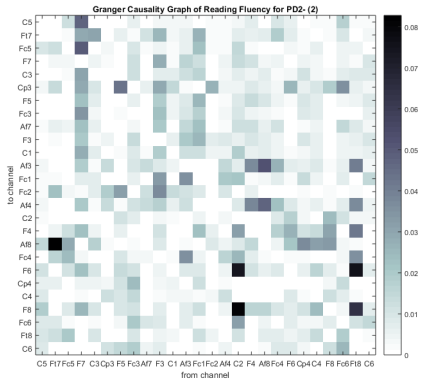
Table 3.15: Channel strength values (Ft7 and F7) of reading fluency for HC_1.

No./Channel	Ft7	F7
1	0.5	0.95
2	0.55	0.76
3	0.8	0.62
4	0.71	0.68
5	1.1	0.82
6	0.72	0.4
7	0.69	0.595
8	0.19	0.49
9	0.4	1.2
10	0.39	1.55

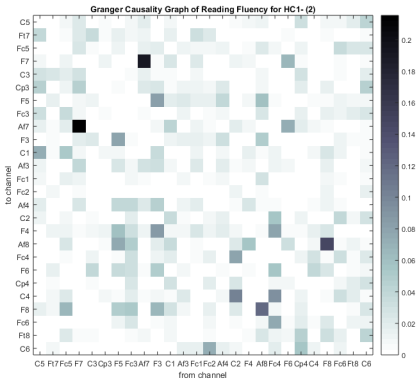
Table 3.16: Channel strength values (Ft7 and F7) of reading fluency for HC_2.



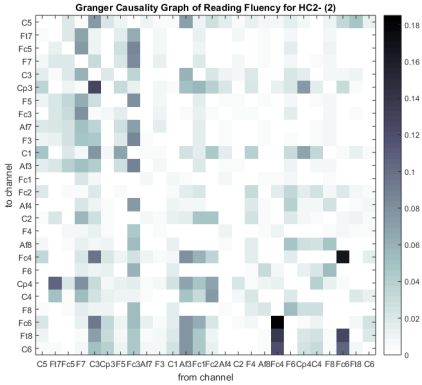
(a) for PD_1



(b) for PD_2

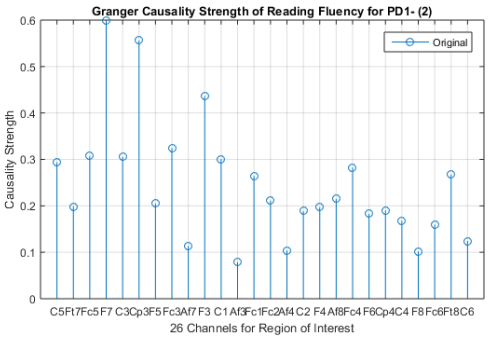


(c) for HC_1

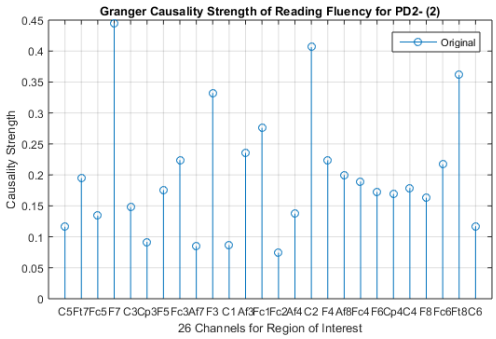


(d) for HC_2

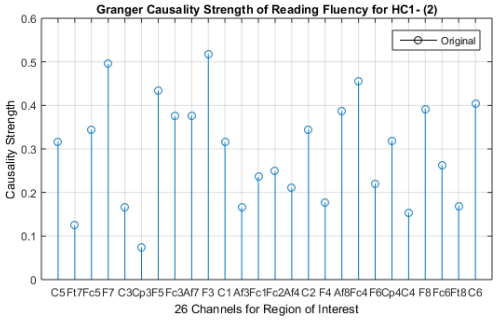
Figure 3.38: Sample Granger causality graphs for reading fluency shows a number of causalities in the Broca's area channels for PD_1 in channel Ft7 causing other channels, PD_2 in channel Ft7 and F7, HC_1 in channel F7 and HC_2 in channel F7.



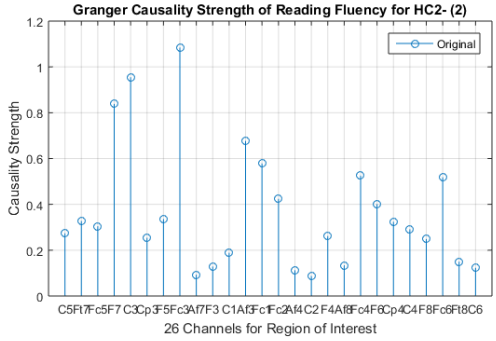
(a) for PD_1



(b) for PD_2

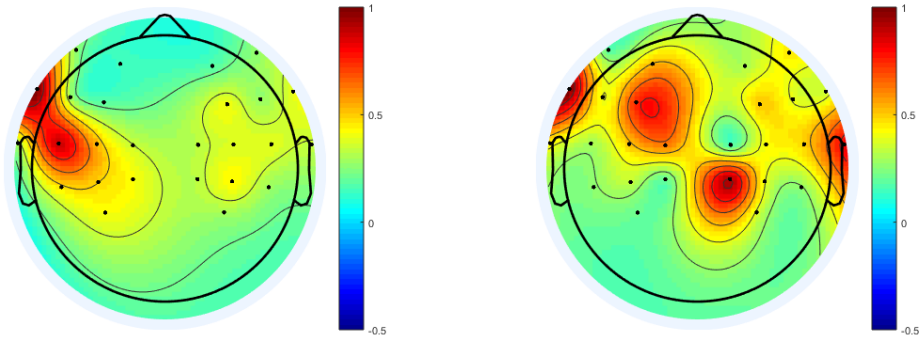


(c) for HC_1



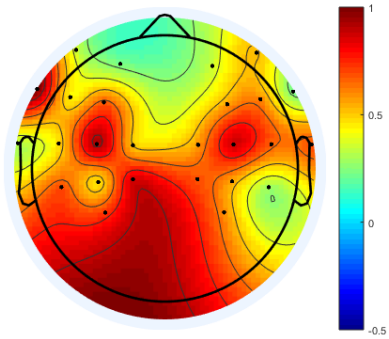
(d) for HC_2

Figure 3.39: Sample Granger causality channel strength for reading fluency shows the causality strength of each channel in the Granger causality graph for each dataset.

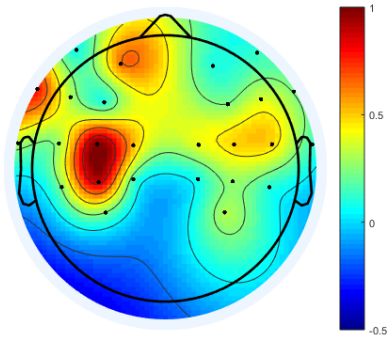


(a) for PD_1

(b) for PD_2

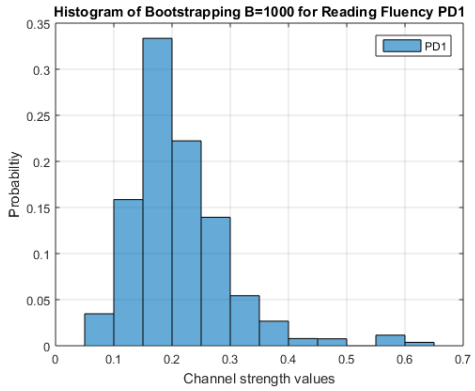


(c) for HC_1

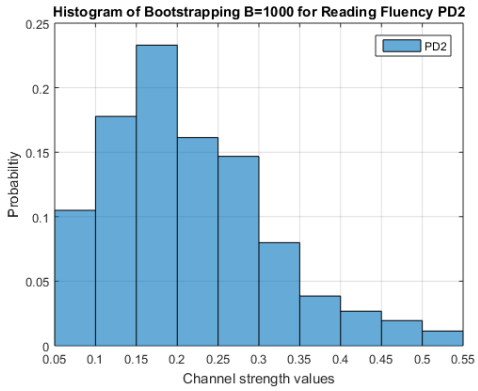


(d) for HC_2

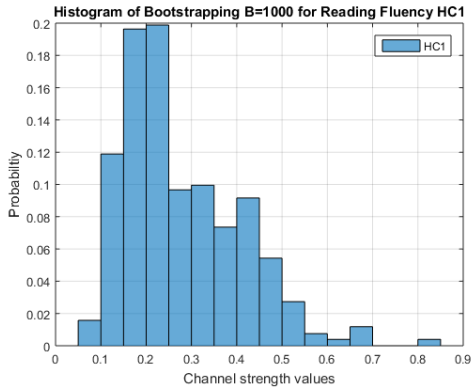
Figure 3.40: Sample Granger causality head maps for reading fluency show the area of activation in the head map for each dataset during reading fluency.



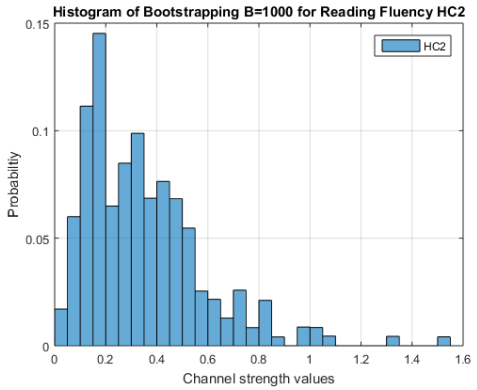
(a) for PD_1



(b) for PD_2



(c) for HC_1



(d) for HC_2

Figure 3.41: Histogram of bootstrapping for reading fluency of all the datasets. The distribution shows the probability of the channel strength values.

channel strength. The p-test threshold for the models of Granger causality is set as $\alpha = 0.05$, and the p-value should be less than or equal to the significance level α. The null hypothesis H_0 of the Granger causality statistical test is assumed as "no differences between channel strengths values under test and all the channel strength in each dataset". On other hand, the assumption of the alternative hypothesis H_1 is related to "differences between the channel strength values under test and all other values in the dataset". The statistical test is one sided of the bootstrapping for Granger causality analysis where the vertical coordinates are the probability values which are computed under the null hypothesis for the different outcomes of channel strengths in the horizontal coordinates as in Figures 3.41. The statistical significance p-value is the area which is less than or equal to the significance level α. The p-value of the channel F7 values for PD_1 and PD_2 is significant with the probability being $p \leq 0.05$. For example, F7 channel strength value for PD_1 of the first sample is 0.39 where this value in the bootstrapping histogram of PD_1 has its p-value ≤ 0.05 which is considered as statistical significant and the null hypothesis is rejected. Furthermore, the probability of the channel F7 values for HC_1 and HC_2 is significant with the probability being $p \leq 0.05$. For example, F7 channel strength value for HC_1 of the first sample is 0.52 where this value in the bootstrapping histogram of HC_1 has its p-value≤ 0.05 which is considered as statistical significant and the null hypothesis is rejected.

Chapter 4

Graph Learning

4.1 Support Vector Machine

Linear support vector machine is considered one of the easiest machine learning techniques for classification, as it maps the feature classifications of linear data space into linear classification with linear hyperplane. It can be achieved by finding the best possible hyperplane that separates the feature classes of the linear data. The hyperplane of linear separation of SVM is placed at the maximum separation region (margin) when the theory is to minimize the normal vector of the hyperplane. The technique of SVM soft margin is applied when a perfect hyperplane does not exist. This method selects the best hyperplane that can separate the data points as clean as possible by maximizing the distance from the margin to the data points in each class. It is also based on the trade-off between the maximum margin and small error penalty of the slack variables. The optimization problem of SVM soft margin is subject to the slackness degree, which indicates whether the points fall into the correct region or are misclassified. The data points in this thesis are the two maximum values of Granger causality channel strengths for each verbal fluency task (phonemic fluency, semantic fluency, category semantic fluency, and reading fluency). These data points fall into two regions of PD=-1 and Control= 1 where

PD is for (PD_1 and PD_2) and Control is for (HC_1 and HC_2). The total number of data points vector for each PD or Control class is twenty because the total number of computation iterations for Granger causality is ten times for each verbal fluency task of each dataset. The dataset of the data points is divided into two subsets for training and testing data points. The ratio of dividing the training points and testing points is considered to be 90% of the total data points for training and 10% of the total data points for testing. The training points are achieved by eighteen points for PD and eighteen points for HC where the testing points are two points for each class. Therefore, the total number of data points for training are thirty six points and for testing are four points. The testing points are selected as one data point from each dataset of PD_1, PD_2, HC_1 and HC_2 in order to examine the classification procedure for the two classes.

Phonemic Fluency

The set of data points for the phonemic fluency is collected from the results of the computation of Granger causality channel strength, as in Figure 3.15. We noticed from the ten iterations of Granger causality channel strength that PD_1 and PD_2 have the channel Ft7 as the highest or considerably highest causal value for most of the iterations. Also, we noticed HC_1 and HC_2 have the channel Fc5 as the highest or considerably highest causal for other channels for most of the ten iterations. The total number of points in the SVM dataset for phonemic fluency is forty belong to the channel strength of Ft7 and Fc5 for the two classes denoted as PD and HC. The dataset is subdivided into training and testing datasets where the ratio is 90% and 10% respectively. We used one data point for classification testing for each dataset PD_1, PD_2, HC_1 and HC_2 with total of four testing data points and thirty six training data points. The SVM soft margin is given in Figure 4.1.

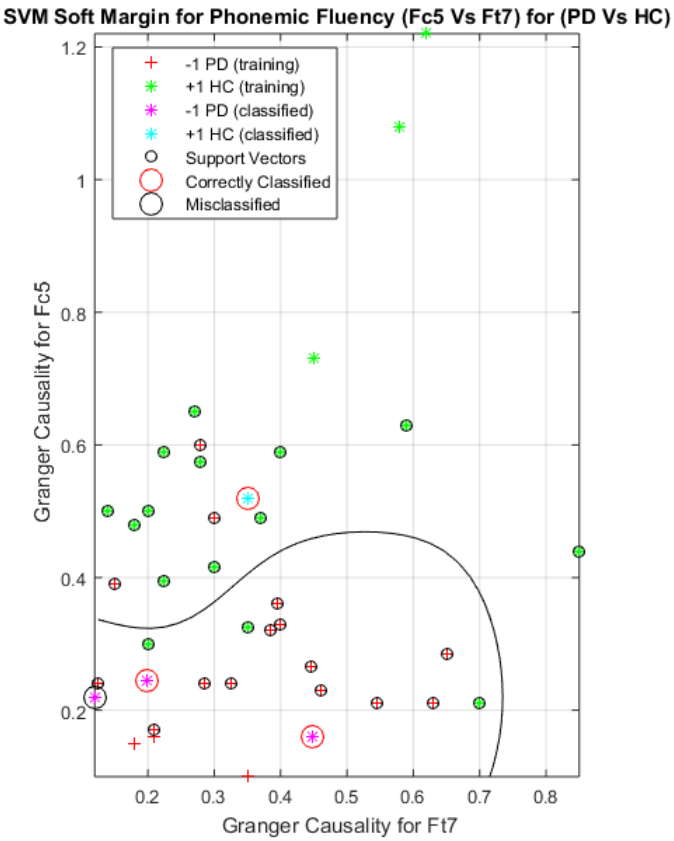


Figure 4.1: SVM soft margin of PD Vs Control of phonemic fluency shows the classification of training and testing points where the ratio is 90:10 respectively.

	HC (Fc5)	PD (Ft7)
HC (Fc5)	16	4
PD (Ft7)	3	17

Table 4.1: Confusion matrix of SVM soft margin of phonemic fluency summarizes the classification of phonemic fluency.

The confusion matrix of the phonemic fluency shows the information of the data points which is subdivided to the training points and the testing points. These data points which are correctly classified and misclassified fall into each class, as in Table

4.1. The true negative rate is for PD patients that have the most causal channel during phonemic verbal in channel Ft7 is 85%. Furthermore, the true positive rate for healthy control subjects is 80% for channel Fc5, which is most causal when they perform phonemic fluency. The accuracy rate of the confusion matrix which indicates the correctly classified data points for the two classes of PD and HC is 82%.

Semantic Fluency

The set of data points for semantic fluency is collected from the results of the computation of Granger causality channel strength, as in Figure 3.22. We noticed from the ten iterations of Granger causality channel strength that PD_1 and PD_2 have the channel Ft7 as the highest or considerably highest causal value for most of the iterations. Also, we noticed HC_1 and HC_2 have the channel F7 as the highest or considerably highest causal for other channels for most of the ten iterations. The total number of points in the SVM dataset of semantic fluency is forty belong to the channel strength of Ft7 and F7 for the two classes PD and HC. The dataset is subdivided into training and testing datasets where the ratio is 90% and 10% respectively. We used one data point for classification testing for each dataset PD_1, PD_2, HC_1 and HC_2 with total of four data points and thirty six training data points. The SVM soft margin is given in Figure 4.2.

	HC (F7)	PD (Ft7)
HC (F7)	14	6
PD (Ft7)	4	16

Table 4.2: Confusion matrix of SVM soft margin of semantic fluency summarizes the classification of semantic fluency.

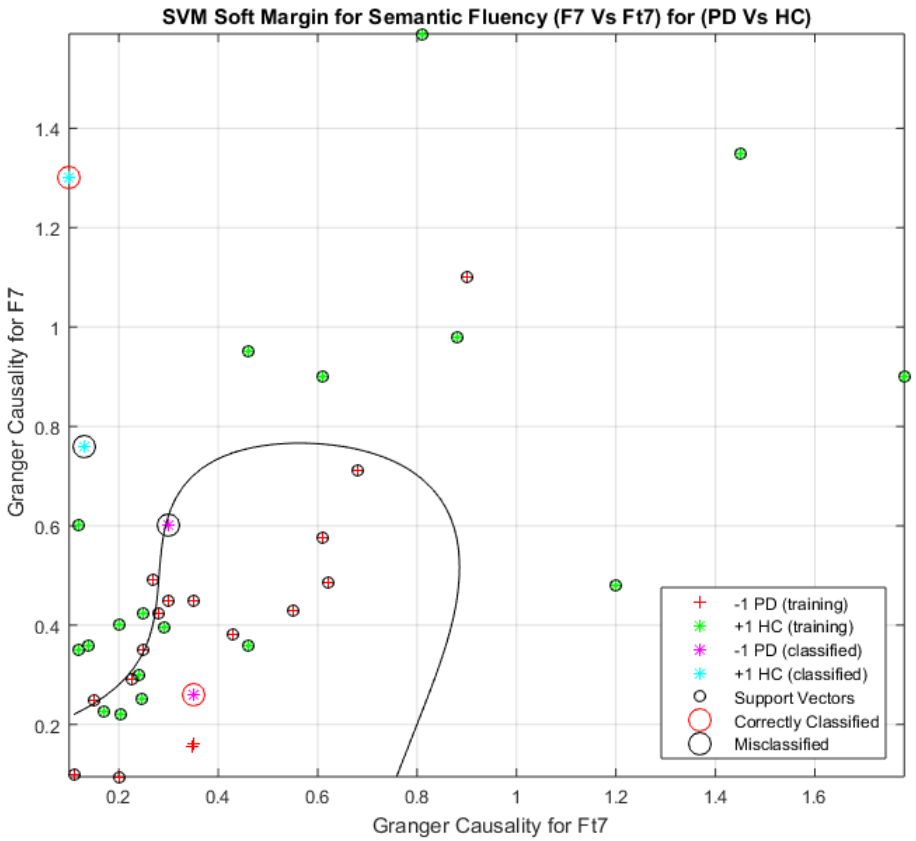


Figure 4.2: SVM soft margin of PD Vs Control of semantic fluency shows the classification of training and testing points where the ratio is 90:10 respectively.

The confusion matrix for semantic fluency demonstrates the information of the data points which is subdivided to the training points and the testing points. These data points which are correctly classified and misclassified that fall into each class, as in Table 4.2. The true negative rate is for PD patients that have the most causal channel during semantic verbal in channel Ft7 is 80%. Furthermore, the true

positive rate for healthy control subjects is 70% for channel F7, which is most causal when they perform phonemic fluency. The accuracy rate of the confusion matrix for the semantic which indicates the correctly classified data points for the two classes of PD and HC is 75%.

Category Semantic Fluency

The set of data points for the category semantic fluency are collected from the results of the computation of Granger causality channel strength, as in Figure 3.29. We noticed from the ten iterations of Granger causality channel strength that PD_1 and PD_2 have the channel Ft7 as the highest or considerably highest causal value for most of the iterations. Also, we noticed HC_1 and HC_2 have the channel F7 as the highest or considerably highest causal for other channels for most of the ten iterations. The total number of points in the SVM dataset of category semantic fluency is forty belong to the channel strength of Ft7 and F7 for the two classes PD and HC. The dataset is subdivided into training and testing datasets where the ratio is 90% and 10% respectively. We used one data point for classification testing for each dataset PD_1, PD_2, HC_1 and HC_2 with total of four testing data points and thirty six training data points. The SVM soft margin is given in Figure 4.3.

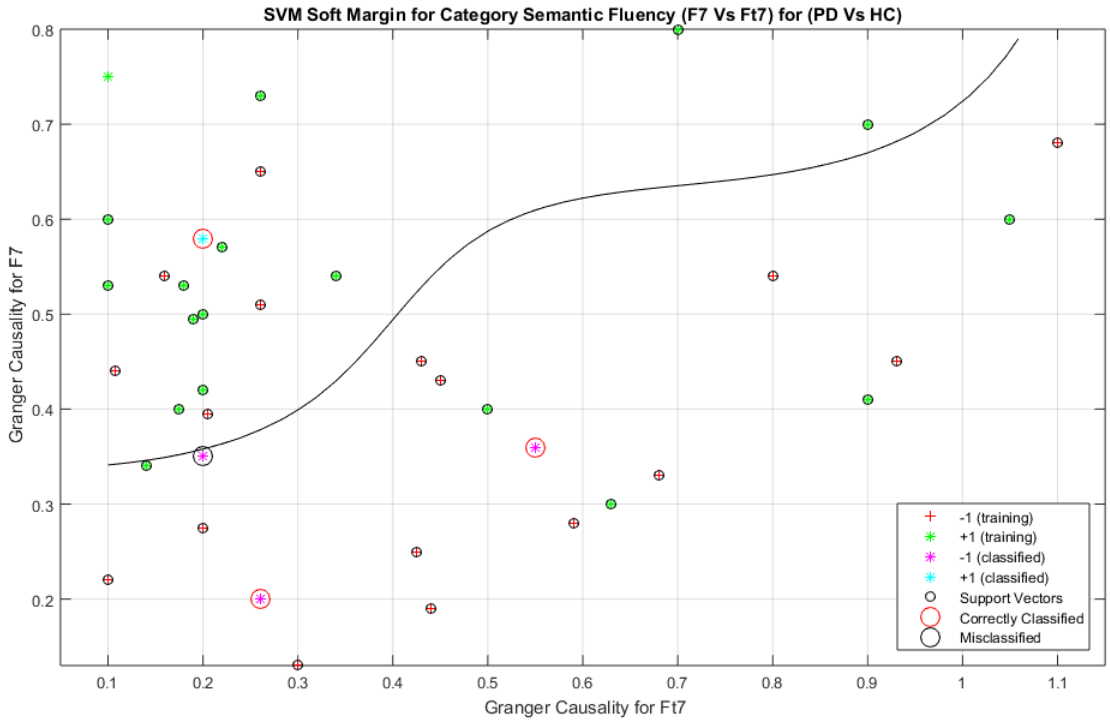


Figure 4.3: SVM soft margin of PD Vs Control of category semantic fluency shows the classification of training and testing points where the ratio is 90:10 respectively.

	HC (F7)	PD (Ft7)
HC (F7)	14	6
PD (Ft7)	5	15

Table 4.3: Confusion matrix of SVM soft margin of category semantic fluency PD Vs Control summarizes the classification of category semantic fluency.

The confusion matrix of the category semantic fluency illustrates the information of the data points which is subdivided to the training points and the testing points. These data points which are correctly classified and misclassified fall into each class, as in Table 4.3.The true negative rate for PD patients that have the most causal channel during semantic verbal in channel Ft7 is 75%. Furthermore, the true positive rate for healthy control subjects is 70% for channel F7, which is most causal when

they perform phonemic fluency. The accuracy rate of the confusion matrix for the semantic which indicates the correctly classified data points for the two classes of PD and HC is 72.5%.

Reading Fluency

The set of data points of the reading fluency test are collected from the results of the computation of Granger causality channel strength, as in Figure 3.36. We noticed from the ten iterations of Granger causality channel strength that PD_1 and PD_2 have the channel Ft7 as the highest or considerably highest causal value for most of the iterations. Also, we noticed HC_1 and HC_2 have the channel F7 as the highest or considerably highest causal for other channels for most of the ten iterations. The total number of points in the SVM dataset of reading fluency is forty belong to the channel strength of Ft7 and F7 for the two classes PD and HC. The dataset is subdivided into training and testing datasets where the ratio is 90% and 10% respectively. We used one data point for classification testing for each dataset PD_1, PD_2, HC_1 and HC_2 with total of four testing data points and thirty six training data points. The SVM soft margin is given in Figure 4.4.

	HC (F7)	PD (Ft7)
HC (F7)	14	6
PD (Ft7)	3	17

Table 4.4: Confusion matrix of SVM soft margin of reading fluency PD Vs Control summarizes the classification of reading fluency.

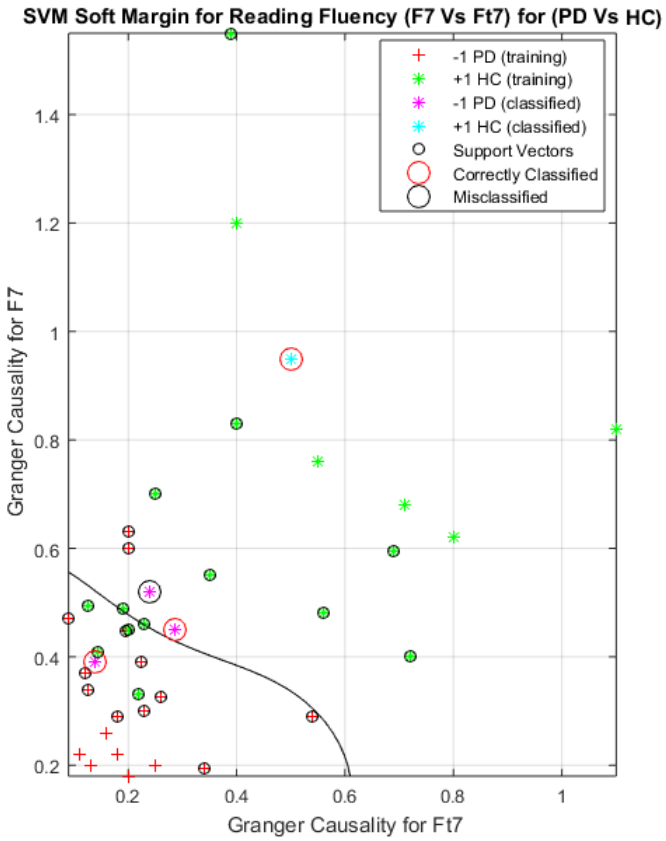


Figure 4.4: SVM soft margin of PD Vs Control of reading fluency shows the classification of training and testing points where the ratio is 90:10 respectively.

The confusion matrix of the reading fluency illustrates the information of the data points which is subdivided to the training points and the testing points. These data points which are correctly classified and misclassified fall into each class, as in Table 4.4. The true negative rate for PD patients that have the most causal channel during semantic verbal in channel Ft7 is 85%. Furthermore, the true positive

rate for healthy control subjects is 70% for channel F7, which is most causal when they perform phonemic fluency. The accuracy rate of the confusion matrix for the semantic which indicates the correctly classified data points for the two classes of PD and HC is 77.5%.

Chapter 5

Discussion and Future Work

5.1 Conclusion

Electroencephalographic (EEG) signals consist of valuable information about brain activity for investigating interaction and causality of different brain areas during the performance of behavior tasks. Signal processing tools play an important role in analyzing EEG signals that are recorded with a number of electrodes on the brain scalps of subjects at the Colorado Neurology Institute (CNI). The method of the study includes processing tools, which classify the brain activity frequency bands during performance of behavior tasks. Multivariate Granger causality method is used to identify the causality components of an EEG signal in regions of interest for participants of Parkinson's Disease (PD) and healthy controls. The channel strength of the Granger causality graph determines the most causal channel in relation to others during different behavior task performance. When channel strength is significant for a number of computations, the support vector machine (SVM) helps to classify the statistical chances of appearance for PD patients and healthy controls.

5.1.1 Current Work

The study applies the method of Granger causality to two main behavior tasks, which are finger movement behavior and verbal fluency for PD patients and healthy controls. PD patients have difficulties when they start movements and switch between movements; they may have difficulties with mental activity when they switch from one thought to another.

Finger Movement Behavior

The main purpose of this research is to identify pathological connectivity in motor cortical areas in Parkinson's Disease patients. The task for this study investigates brain connectivity in relation to tapping (with the left or right hand) when subjects hear an audio cue. The study reveals increased connectivity in the left and right motor planning areas (F3 and F4) when the recognition of the left and right motor planning areas had a rate 83.3% for F3 and 91.7% for F4. As well, there is increased connectivity in the left and right sensorimotor integration areas (C3 and C4) when the recognition rate is 91.7% for C3 and 91.7% for C4. These promising results of research have already been accepted to Asilomar conference 2015 [56].

Verbal Fluency

The main purpose of this thesis is to identify pathological connectivity in Broca's area within Brodmann's area for verbal fluency in Parkinson's Disease subjects. This study investigates brain connectivity for verbal fluency tasks, such as phonemic fluency, semantic fluency, category semantic fluency, and reading fluency for subjects when they generate as many words as they can think of in 60 seconds. The study concludes that there is connectivity in Broca's area of Brodmann's area of BA44 and B45 EEG electrodes (Ft7 and Fc5) for phonemic fluency. And the recognition rate of PD vs. HC in phonemic fluency is 85% for PD participants and 80% for healthy

controls. Also, there is increased connectivity in Broca's area of Brodmann's area of BA44 and B45 (Ft7 and F7) for semantic fluency when the recognition rate is 80% for PD participants and 70% for healthy controls. As well, there is also increased connectivity in Broca's area of Brodmann's area of BA44 and B45 (Ft7 and F7) for category semantic fluency when the recognition rate is 75% for PD participants and 70% for for healthy controls. Finally, there is connectivity in Broca's area of Brodmann's area of BA44 and B45 (Ft7 and F7) for reading fluency when the recognition rate is 85% for PD participants and 70% for for healthy controls.

5.1.2 Future Work

From the two behavior studies on motor movement and verbal fluency, the promising results show great potential for extended applications in the medical field; these applications can be used to identify the cortical information flow patterns for different behavior tasks or to identify information patterns of patients who are on or off medication, or have deep brain stimulation, etc.

The further work includes the investigation of brain connectivity for multi-behavior performances, such as verbal fluency and tapping together, Furthermore, applying the developed approach for more participants in order to further validate the conclusions we drew in the thesis.

Bibliography

[1] M. Teplan, "Fundementals of EEG measurment," *Measurement Science Review*, vol. 2, p. Section 2, 2002.

[2] S. L. Bressler and A. K. Seth, "Wiener Granger causality: A well established methodology," *NeuroImage*, vol. 58, no. 2, pp. 323–329, 2011.

[3] M. Z. Koubeissi, "Niedermeyers electroencephalography, basic principles, clinical applications, and related fields, 6th ed," *Archives of Neurology*, vol. 68, no. 11, p. 1481, 2011.

[4] H. H. Jasper, "The ten-twenty electrode system of the international federation," *Electroencephalography and Clinical Neurophysiology*, vol. 10, pp. 371–375, 1958.

[5] J. Chiang, Z. J. Wang, and M. J. McKeown, "EEG source extraction by autoregressive source separation reveals abnormal synchronization in Parkinson's disease," *IEEE Engineering in Medicine and Biology Society*, pp. 1868 – 1872, 2009.

[6] S. Haufe, V. V. Nikulin, K.-R. Mller, and G. Nolte, "A critical assessment of connectivity measures for EEG data: A simulation study," *NeuroImage*, vol. 64, pp. 120–133, 2013.

[7] I. Daly, F. Pichiorri, J. Faller, V. Kaiser, A. Kreilinger, R. Scherer, and G. Muller-Putz, "What does clean EEG look like?," in *IEEE Engineering in Medicine and Biology Society*, pp. 3963–3966, September 2012.

[8] I. Daly, M. Billinger, R. Scherer, and G. Muller-Putz, "On the automated removal of artifacts related to head movement from the EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 3, pp. 427–434, 2013.

[9] M. A. Jamal, *Analysis and Classification of EEG signals using Mixture of Features and Committee Neural Network*. PhD thesis, National Institute of Technology, Rourkela-769008, Odisha, India, 2012.

[10] L. Zhang, G. Zhong, Y. Wu, M. G. Vangel, B. Jiang, and J. Kong, "Using Granger Geweke causality model to evaluate the effective connectivity of primary motor cortex, supplementary motor area and cerebellum," *Journal of Biomedical Science and Engineering*, vol. 03, pp. 848–860, September 2010.

[11] J. Orgogozo and B. Larsen, "Activation of the supplementary motor area during voluntary movement in man suggests it works as a supramotor area," *Science*, vol. 206, no. 4420, pp. 847–850, 1979.

[12] M. Catalan, "The functional neuroanatomy of simple and complex sequential finger movements: a PET study," *Brain*, vol. 121, no. 2, pp. 253–264, 1998.

[13] M.-P. Deiber, R. Passingham, J. Colebatch, K. Friston, P. Nixon, and R. Frackowiak, "Cortical areas and the selection of movement: a study with positron emission tomography," *Experimental Brain Research*, vol. 84, no. 2, 1991.

[14] C. Andrew and G. Pfurtscheller, "Event-related coherence as a tool for studying dynamic interaction of brain regions," *Electroencephalography and Clinical Neurophysiology*, vol. 98, no. 2, pp. 144–148, 1996.

[15] P. Manganotti, C. Gerloff, C. Toro, H. Katsuta, N. Sadato, P. Zhuang, L. Leocani, and M. Hallett, "Task-related coherence and task-related spectral power changes during sequential finger movements," *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control*, vol. 109, no. 1, pp. 50–62, 1998.

[16] K. J. Friston, "Functional and effective connectivity in neuroimaging: A synthesis," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 56–78, 1994.

[17] Q. Gao, H. Chen, and Q. Gong, "Evaluation of the effective connectivity of the dominant primary motor cortex during bimanual movement using Granger causality," *Neuroscience Letters*, vol. 443, no. 1, pp. 1–6, 2008.

[18] M. Boenstrup, J. Feldheim, K. Heise, C. Gerloff, and F. C. Hummel, "The control of complex finger movements by directional information flow between mesial frontocentral areas and the primary motor cortex," *European Journal of Neuroscience*, vol. 40, pp. 2888–2897, September 2014.

[19] K. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273–1302, 2003.

[20] M. J. Kaminski and K. J. Blinowska, "A new method of the description of the information flow in the brain structures," *Biological Cybernetics*, vol. 65, pp. 203–210, July 1991.

[21] G. Deshpande, S. LaConte, G. A. James, S. Peltier, and X. Hu, "Multivariate Granger causality analysis of fMRI data," *Human Brain Mapping*, vol. 30, no. 4, pp. 1361–1373, 2009.

[22] A. Brovelli, M. Ding, A. Ledberg, Y. Chen, R. Nakamura, and S. L. Bressler, "Beta oscillations in a large-scale sensorimotor cortical network: Directional

influences revealed by Granger causality," *Proceedings of the National Academy of Sciences*, vol. 101, no. 26, pp. 9849–9854, 2004.

[23] C. Granger, "Economic processes involving feedback," *Information and Control*, vol. 6, no. 1, pp. 28–48, 1963.

[24] J. B. Rowe, "Connectivity analysis is essential to understand neurological disorders," *Frontiers in Systems Neuroscience*, vol. 4, 2010.

[25] A. Roebroeck, E. Formisano, and R. Goebel, "Mapping directed influence over the brain using Granger causality and fMRI," *NeuroImage*, vol. 25, no. 1, pp. 230–242, 2005.

[26] A. Schlgl and G. Supp, "Analyzing event-related EEG data with multivariate autoregressive parameters," *Progress in Brain Research*, vol. 159, p. 135147, 2006.

[27] F. Babiloni, F. Cincotti, C. Babiloni, F. Carducci, D. Mattia, L. Astolfi, A. Basilisco, P. Rossini, L. Ding, and Y. e. a. Ni, "Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function," *NeuroImage*, vol. 24, no. 1, pp. 118–131, 2005.

[28] M. Baker, K. Kapse, A. McMahon, and M. OBoyle, "Connectivity in math-gifted adolescents: Comparing structural equation modeling, Granger causality, and dynamic causal modeling," in *Image Analysis and Interpretation (SSIAI), 2012 Southwest Symposium Conference of the IEEE*, pp. 93 – 96, 2012.

[29] N. Wang, Y. Wang, Y. Li, Y. Tang, and J. Wang, "Gamma oscillation in brain connectivity in emotion recognition by granger causality," in *Biomedical Engineering and Informatics (BMEI), 2011 4th International Conference on IEEE*, pp. 762 – 766, October 2011.

[30] C. Dongwei, W. Fang, W. Zhen, L. Haifang, and C. Junjie, "EEG based emotion recognition with brain network using independent components analysis and Granger causality," in *International Conference on Computer Medical Applications (ICCMA)*, pp. 1 – 6, 2013.

[31] M. Ghasemi and A. Mahloojifar, "Disorganization of equilibrium directional interactions in the brain motor network of Parkinson's disease: New insight of resting state analysis using Granger causality and graphical approach," *Journal of Medical Signals and Sensors*, vol. 3, no. 2, p. 6978, 2015.

[32] C. Hammond, H. Bergman, and P. Brown, "Pathological synchronization in Parkinson's disease: networks, models and treatments," *Trends in Neurosciences*, vol. 30, no. 7, pp. 357–364, 2007.

[33] N. Fogelson, "Different functional loops between cerebral cortex and the subthalmic area in Parkinson's disease," *Cerebral Cortex*, vol. 16, no. 1, pp. 64–75, 2005.

[34] D. Williams, "Dopamine-dependent changes in the functional connectivity between basal ganglia and cerebral cortex in humans," *Brain*, vol. 125, no. 7, pp. 1558–1569, 2002.

[35] J. Hirschmann, T. zkurt, M. Butz, M. Homburger, S. Elben, C. Hartmann, J. Vesper, L. Wojtecki, and A. Schnitzler, "Differential modulation of STN -cortical and cortico-muscular coherence by movement and levodopa in Parkinson's disease," *NeuroImage*, vol. 68, pp. 203–213, 2013.

[36] E. Lalo, S. Thobois, A. Sharott, G. Polo, P. Mertens, A. Pogosyan, and P. Brown, "Patterns of bidirectional communication between cortex and basal ganglia during movement in patients with Parkinson disease," *Journal of Neuroscience*, vol. 28, no. 12, pp. 3008–3016, 2008.

[37] B. N. Jvor-Duray, M. Vinck, M. van der Roest, A. B. Mulder, C. J. Stam, H. W. Berendse, and P. Voorn, "Early-onset cortico-cortical synchronization in the hemiparkinsonian rat model," *Journal of Neurophysiology*, vol. 113, no. 3, pp. 925–936, 2014.

[38] R. Soikkeli, J. Partanen, H. Soininen, A. Pkknen, and P. Riekkinen, "Slowing of EEG in Parkinson's disease," *Electroencephalography and Clinical Neurophysiology*, vol. 79, no. 3, pp. 159–165, 1991.

[39] J. Weyhenmeyer, M. E. Hernandez, C. Lainscsek, T. J. Sejnowski, and H. Poizner, "Muscle artifacts in single trial EEG data distinguish patients with Parkinson's disease from healthy individuals," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pp. 3292 – 3295, 2014.

[40] J. S. George, J. Strunk, R. Mak-McCully, M. Houser, H. Poizner, and A. R. Aron, "Dopaminergic therapy in Parkinson's disease decreases cortical beta band coherence in the resting state and increases cortical beta band power during executive control," *NeuroImage: Clinical*, vol. 3, pp. 261–270, 2013.

[41] N. Crone, "Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. I. alpha and beta event- related desynchronization," *Brain*, vol. 121, no. 12, pp. 2271–2299, 1998.

[42] P. Brown, "Oscillatory nature of human basal ganglia activity: Relationship to the pathophysiology of Parkinson's disease," *Movement Disorders*, vol. 18, no. 4, pp. 357–363, 2003.

[43] M. Moazami-Goudarzi, J. Sarnthein, L. Michels, R. Moukhtieva, and D. Jeanmonod, "Enhanced frontal low and high frequency power and synchronization in

the resting EEG of parkinsonian patients," *NeuroImage*, vol. 41, no. 3, pp. 985–997, 2008.

[44] S. J. Palmer, P. Wen-Hsin Lee, Z. J. Wang, W.-L. Au, and M. J. McKeown, "Theta, beta but not alpha-band EEG connectivity has implications for dual task performance in Parkinsons disease," *Parkinsonism & Related Disorders*, vol. 16, no. 6, pp. 393–397, 2010.

[45] M. Sharman, R. Valabregue, V. Perlbarg, L. Marrakchi-Kacem, M. Vidailhet, H. Benali, A. Brice, and S. Lehricy, "Parkinson's disease patients show reduced cortical-subcortical sensorimotor connectivity," *Movement Disorders*, vol. 28, no. 4, pp. 447–454, 2012.

[46] J.-M. Melgari, "Alpha and beta EEG power reflects L-dopa acute administration in parkinsonian patients," *Frontiers in Aging Neuroscience*, vol. 6, 2014.

[47] M. Neufeld, S. Blumen, I. Aitkin, Y. Parmet, and A. Korczyn, "EEG frequency analysis in demented and nondemented parkinsonian patients," *Dementia*, vol. 5, no. 1, pp. 23–28, 1994.

[48] B. Pasquereau and R. S. Turner, "Primary motor cortex of the parkinsonian monkey: Differential effects on the spontaneous activity of pyramidal tract-type neurons," *Cerebral Cortex*, vol. 21, no. 6, pp. 1362–1378, 2010.

[49] E. Brazhnik, A. V. Cruz, I. Avila, M. I. Wahba, N. Novikov, N. M. Ilieva, A. J. McCoy, C. Gerber, and J. R. Walters, "State-dependent spike and local field synchronization between motor cortex and substantia nigra in Hemiparkinsonian rats," *Journal of Neuroscience*, vol. 32, no. 23, pp. 7869–7880, 2012.

[50] J.-S. Brittain and P. Brown, "Oscillations and the basal ganglia: Motor control and beyond," *NeuroImage*, vol. 85, pp. 637–647, 2014.

[51] P. Brown, "Abnormal oscillatory synchronisation in the motor system leads to impaired movement," *Current Opinion in Neurobiology*, vol. 17, no. 6, pp. 656–664, 2007.

[52] A. N. Vardy, E. E. van Wegen, G. Kwakkel, H. W. Berendse, P. J. Beek, and A. Daffertshofer, "Slowing of M1 activity in Parkinsons disease during rest and movement : An MEG study," *Clinical Neurophysiology*, vol. 122, no. 4, pp. 789–795, 2011.

[53] E. Herrera, F. Cuetos, and R. Ribacoba, "Verbal fluency in Parkinsons disease patients on/off dopamine medication," *Neuropsychologia*, vol. 50, no. 14, pp. 3636–3640, 2012.

[54] J. B. Pereira, C. Junqu, D. Bartrs-Faz, M. J. Mart, R. Sala-Llonch, Y. Compta, C. Falcn, P. Vendrell, l. Pascual-Leone, and J. Valls-Sol, "Modulation of verbal fluency networks by transcranial direct current stimulation (tDCS) in Parkinsons disease," *Brain Stimulation*, vol. 6, no. 1, pp. 16–24, 2013.

[55] A. Almalaq, X. Dai, J. Zhang, S. Hanrahan, J. Nedrud, and A. Hebb, "Causality graph learning on cortical information flow in Parkinsons disease patients during behaviour tests," in *IEEE Signal Processing Society, 49th Asilomar Conference on Singals, Systems and Computers*, 2015. Unpublished.

[56] [online] "Asilomar Conference on Signals Systems and Computers" Available: http://asilomarsscconf.org/. [Accessed: May 11 2015].

[57] J. B. Pereira, C. Junqu, M. J. Mart, B. Ramirez-Ruiz, D. Bartrs-Faz, and E. Tolosa, "Structural brain correlates of verbal fluency in Parkinson's disease," *NeuroReport*, vol. 20, no. 8, pp. 741–744, 2009.

[58] C. Randolph, A. R. Braun, T. E. Goldberg, and T. N. Chase, "Semantic fluency in Alzheimer's, Parkinson's, and Huntington's disease: Dissociation of storage and retrieval failures.," *Neuropsychology*, vol. 7, no. 1, pp. 82–88, 1993.

[59] D. M. Jacobs, K. Marder, L. J. Cote, M. Sano, Y. Stern, and R. Mayeux, "Neuropsychological characteristics of preclinical dementia in Parkinson's disease," *Neurology*, vol. 45, no. 9, pp. 1691–1696, 1995.

[60] J. D. Henry and J. R. Crawford, "A meta-analytic review of verbal fluency performance following focal cortical lesions.," *Neuropsychology*, vol. 18, no. 2, pp. 284–295, 2004.

[61] S. Abrahams, L. H. Goldstein, A. Simmons, M. J. Brammer, S. C. Williams, V. P. Giampietro, C. M. Andrew, and P. N. Leigh, "Functional magnetic resonance imaging of verbal fluency and confrontation naming using compressed image acquisition to permit overt responses," *Human Brain Mapping*, vol. 20, no. 1, pp. 29–40, 2003.

[62] M. Pihlajamaki, H. Tanila, T. Hanninen, M. Kononen, M. P. Laakso, K. Partanen, H. Soininen, and H. J. Aronen, "Verbal fluency activates the left medial temporal lobe: A functional magnetic resonance imaging study," *Neurobiology of Aging*, vol. 21, p. 106, 2000.

[63] C. D. Frith, K. J. Friston, S. Herold, D. Silbersweig, P. Fletcher, C. Cahill, R. J. Dolan, R. S. Frackowiak, and P. F. Liddle, "Regional brain activity in chronic schizophrenic patients during the performance of a verbal fluency task," *The British Journal of Psychiatry*, vol. 167, no. 3, pp. 343–349, 1995.

[64] J. L. Cummings, A. Darkins, M. Mendez, M. A. Hill, and D. F. Benson, "Alzheimer's disease and Parkinson's disease: Comparison of speech and language alterations," *Neurology*, vol. 38, no. 5, pp. 680–680, 1988.

[65] A. K. Troyer, M. Moscovitch, G. Winocur, L. Leach, Freedman, and Morris, "Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease," *Journal of the International Neuropsychological Society*, vol. 4, no. 2, pp. 137–143, 1998.

[66] D. Muslimovic, B. Post, J. D. Speelman, and B. Schmand, "Cognitive profile of patients with newly diagnosed Parkinson disease," *Neurology*, vol. 65, no. 8, pp. 1239–1245, 2005.

[67] U. Ellfolk, J. Joutsa, J. O. Rinne, R. Parkkola, P. Jokinen, and M. Karrasch, "Striatal volume is related to phonemic verbal fluency but not to semantic or alternating verbal fluency in early Parkinsons disease," *Journal of Neural Transmission*, vol. 121, no. 1, pp. 33–40, 2013.

[68] E. A. Hirshorn and S. L. Thompson-Schill, "Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency," *Neuropsychologia*, vol. 44, no. 12, pp. 2547–2557, 2006.

[69] S. G. Costafreda, C. H. Fu, L. Lee, B. Everitt, M. J. Brammer, and A. S. David, "A systematic review and quantitative appraisal of fMRI studies of verbal fluency: Role of the left inferior frontal gyrus," *Human Brain Mapping*, vol. 27, no. 10, pp. 799–810, 2006.

[70] S. Wagner, A. Sebastian, K. Lieb, O. Tscher, and A. Tadic, "A coordinate-based ALE functional MRI meta-analysis of brain activation during verbal fluency tasks in healthy control subjects," *BMC Neuroscience*, vol. 15, no. 1, p. 19, 2014.

[71] S. L. Thompson-Schill, M. D'Esposito, G. K. Aguirre, and M. J. Farah, "Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reeval-

uation," *Proceedings of the National Academy of Sciences*, vol. 94, no. 26, pp. 14792–14797, 1997.

[72] S. Schwartz and J. Baldo, "Distinct patterns of word retrieval in right and left frontal lobe patients: a multidimensional perspective," *Neuropsychologia*, vol. 39, no. 11, pp. 1209–1217, 2001.

[73] J. V. Baldo, A. P. Shaimamura, D. C. Delis, J. Kramer, and E. Kaplan, "Verbal and design fluency in patients with frontal lobe lesions," *Journal of the International Neuropsychological Society*, vol. 7, no. 5, pp. 586–596, 2001.

[74] B. T. Gold and R. L. Buckner, "Common prefrontal regions coactivate with dissociable posterior regions during controlled semantic and phonological tasks," *Neuron*, vol. 35, no. 4, pp. 803–812, 2002.

[75] S. Heim, S. B. Eickhoff, and K. Amunts, "Specialisation in Broca's region for semantic, phonological, and syntactic fluency?," *NeuroImage*, vol. 40, no. 3, pp. 1362–1368, 2008.

[76] [online]. " Brainm.com Electrode Postions " Available: http://www.brainm.com/software/pubs. [Accessed: Jun 17 2015].

[77] A. Juphard, J. R. Vidal, M. Perrone-Bertolotti, L. Minotti, P. Kahane, J.-P. Lachaux, and M. Baciu, "Direct evidence for two different neural mechanisms for reading familiar and unfamiliar words: An intra-cerebral EEG study," *Frontiers in Human Neuroscience*, vol. 5, 2011.

[78] M. C. Cervenka, "Electrocorticographic functional mapping identifies human cortex critical for auditory and visual naming," *NeuroImage, suppl. C*, pp. 267–276, April 1 2013.

[79] X. Chen, Z. Syed, and A. Hero, "EEG spatial decoding with shrinkage optimized directed information assessment," in *Acoustics, Speech and Signal Pro-*

cessing (ICASSP), 2012 International Conference on IEEE, pp. 577 – 580, 2012.

[80] *Handbook of Time Series Analysis; Recent Theoretical Development and Applications*. Wille-VCH, 1 ed., 2015.

[81] W. Hesse, E. Mller, M. Arnold, and B. Schack, "The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies," *Journal of Neuroscience Methods*, vol. 124, no. 1, pp. 27–44, 2003.

[82] K. J. Blinowska, R. Kus, and M. Kaminski, "Granger causality and information flow in multivariate processes," *Physical Review E*, vol. 70, no. 5, 2004.

[83] L. Barnett and A. K. Seth, "The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference," *Journal of Neuroscience Methods*, vol. 223, pp. 50–68, 2014.

[84] D. Posada and T. Buckley, "Model selection and model averaging in phylogenetics: Advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests," *Systematic Biology*, vol. 53, no. 5, pp. 793–808, 2004.

[85] C. J. Keylock, "Constrained surrogate time series with preservation of the mean and variance structure," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 73(2), no. 3, 2006.

[86] S. Hu, J. Cao, Yu. Zhang, K. Kong, Wanzeng. Yang, Y. Zhang, and X. Li, "More discussions for granger causality and new causality measures," *Cognitive Neurodynamics*, vol. 6, p. 3342, Sep 27 2011.

[87] B. Efron and R. Tibshirani, *An introduction to the bootstrap*. New York: Chapman & Hall, 1994.

[88] C. Diks and J. DeGoede, *Global analysis of dynamical systems*, ch. A general nonparametric bootstrap test for Granger causality., pp. 391–403. 2001.

[89] J. Richiardi, S. Achard, H. Bunke, and D. Van De Ville, "Machine learning with brain graphs: Predictive modeling approaches for functional imaging in systems neuroscience," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 58–70, 2013.

[90] F. Fallani, L. Costa, F. Rodriguez, L. Astolfi, G. Vecchiato, J. Toppi, G. Borghini, F. Cincotti, D. Mattia, and S. e. a. Salinari, "A graph-theoretical approach in brain functional networks. possible implications in EEG studies," *Nonlinear Biomed Phys*, vol. 4, no. Suppl 1, p. S8, 2010.

[91] L. d. F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas, "Characterization of complex networks: A survey of measurements," *Advances in Physics*, vol. 56, no. 1, pp. 167–242, 2007.

[92] S. Yin, X. Gao, H. R. Karimi, and X. Zhu, "Study on support vector machine-based fault detection in Tennessee eastman process," *Abstract and Applied Analysis*, vol. 2014, pp. 1–8, 2014.

[93] X. Xu, I. W. Tsang, and D. Xu, "Soft margin multiple kernel learning," *IEEE Trans. Neural Netw. Learning Syst.*, vol. 24, no. 5, pp. 749–761, 2013.