## **Florida International University [FIU Digital Commons](https://digitalcommons.fiu.edu/?utm_source=digitalcommons.fiu.edu%2Fetd%2F971&utm_medium=PDF&utm_campaign=PDFCoverPages)**

[FIU Electronic Theses and Dissertations](https://digitalcommons.fiu.edu/etd?utm_source=digitalcommons.fiu.edu%2Fetd%2F971&utm_medium=PDF&utm_campaign=PDFCoverPages) **EXECUTE:** [University Graduate School](https://digitalcommons.fiu.edu/ugs?utm_source=digitalcommons.fiu.edu%2Fetd%2F971&utm_medium=PDF&utm_campaign=PDFCoverPages)

11-13-2013

# Network Construction and Graph Theoretical Analysis of Functional Language Networks in Pediatric Epilepsy

Anas Salah Eddin anas.salaheddin@fiu.edu

**DOI:** 10.25148/etd.FI13120407 Follow this and additional works at: [https://digitalcommons.fiu.edu/etd](https://digitalcommons.fiu.edu/etd?utm_source=digitalcommons.fiu.edu%2Fetd%2F971&utm_medium=PDF&utm_campaign=PDFCoverPages) Part of the [Bioelectrical and Neuroengineering Commons](http://network.bepress.com/hgg/discipline/231?utm_source=digitalcommons.fiu.edu%2Fetd%2F971&utm_medium=PDF&utm_campaign=PDFCoverPages), [Biomedical Commons](http://network.bepress.com/hgg/discipline/267?utm_source=digitalcommons.fiu.edu%2Fetd%2F971&utm_medium=PDF&utm_campaign=PDFCoverPages), [Biostatistics](http://network.bepress.com/hgg/discipline/210?utm_source=digitalcommons.fiu.edu%2Fetd%2F971&utm_medium=PDF&utm_campaign=PDFCoverPages) [Commons,](http://network.bepress.com/hgg/discipline/210?utm_source=digitalcommons.fiu.edu%2Fetd%2F971&utm_medium=PDF&utm_campaign=PDFCoverPages) and the [Other Computer Sciences Commons](http://network.bepress.com/hgg/discipline/152?utm_source=digitalcommons.fiu.edu%2Fetd%2F971&utm_medium=PDF&utm_campaign=PDFCoverPages)

#### Recommended Citation

Salah Eddin, Anas, "Network Construction and Graph Theoretical Analysis of Functional Language Networks in Pediatric Epilepsy" (2013). *FIU Electronic Theses and Dissertations*. 971. [https://digitalcommons.fiu.edu/etd/971](https://digitalcommons.fiu.edu/etd/971?utm_source=digitalcommons.fiu.edu%2Fetd%2F971&utm_medium=PDF&utm_campaign=PDFCoverPages)

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact [dcc@fiu.edu](mailto:dcc@fiu.edu).

### FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

## NETWORK CONSTRUCTION AND GRAPH THEORETICAL ANALYSIS OF FUNCTIONAL LANGUAGE NETWORKS IN PEDIATRIC EPILEPSY

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ELECTRICAL ENGINEERING

by

Anas Salah Eddin

2013

To: Dean Amir Mirmiran College of Engineering and Computing

This dissertation, written by Anas Salah Eddin, and entitled Network Construction and Graph Theoretical Analysis of Functional Language Networks in Pediatric Epilepsy, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

Armando Barreto

Jean H. Andrian

S. Masoud Sadjadi

Malek Adjouadi, Major Professor

\_

\_

\_

\_

Date of Defense: November 13, 2013

The dissertation of Anas Salah Eddin is approved.

Dean Amir Mirmiran College of Engineering and Computing

\_

\_

Dean Lakshmi N. Reddi University Graduate School

Florida International University, 2013

© Copyright 2013 by Anas Salah Eddin

All rights reserved.

## DEDICATION

To family and loved ones.

#### ACKNOWLEDGMENTS

I would like to thank my entire family and friends for all their support and patience during my doctoral studies. They gave me strength and motivated me to move forward, their advice helped me become the person I am now.

I would also like to express my sincere gratitude to my mentor and advisor, Dr. Malek Adjouadi, for all his support through two graduate degrees, for his patience, and for his guidance. I must also thank my committee members Dr. Armando Barreto, Dr. Jean Andrian and Dr. S. Masoud Sadjadi for their valuable input.

Several others have helped me during my graduate journey and I do not have enough words to thank them. I would like to especially thank Dr. Neda Ladbon-Bernasconi and Dr. Andrea Bernasconi who helped put me on a focused research path. The discussions and input from Daniel Tarquinio were particularly insightful and I would like to thank him for his patience and support. I would also like to thank my colleagues at the Center for Advanced Technology and Education for all their help and support, I am especially appreciative of the support provided by: Jin Wang, Saman Sargolzaie, Javier Delgado, Mouncef Lahlou, and Ana Guzman.

Finally, I would like to recognize the support of the National Science Foundation under grant Numbers: CNS-0959985, CNS-1042341, HRD-0833093, and IIP-1230661. Which facilitated the research conducted to complete this dissertation.

#### ABSTRACT OF THE DISSERTATION

## NETWORK CONSTRUCTION AND GRAPH THEORETICAL ANALYSIS OF FUNCTIONAL LANGUAGE NETWORKS IN PEDIATRIC EPILEPSY

by

Anas Salah Eddin

Florida International University, 2013

Miami, Florida

#### Professor Malek Adjouadi, Major Professor

This dissertation introduces a new approach for assessing the effects of pediatric epilepsy on the language connectome. Two novel data-driven network construction approaches are presented. These methods rely on connecting different brain regions using either extent or intensity of language related activations as identified by independent component analysis of fMRI data. An auditory description decision task (ADDT) paradigm was used to activate the language network for 29 patients and 30 controls recruited from three major pediatric hospitals. Empirical evaluations illustrated that pediatric epilepsy can cause, or is associated with, a network efficiency reduction. Patients showed a propensity to inefficiently employ the whole brain network to perform the ADDT language task; on the contrary, controls seemed to efficiently use smaller segregated network components to achieve the same task. To explain the causes of the decreased efficiency, graph theoretical analysis was carried out. The analysis revealed no substantial global network feature differences between the patient and control groups. It also showed that for both subject groups the language network exhibited small-world characteristics; however, the patient's extent of activation network showed a tendency towards more random networks.

It was also shown that the intensity of activation network displayed ipsilateral hub reorganization on the local level. The left hemispheric hubs displayed greater centrality values for patients, whereas the right hemispheric hubs displayed greater centrality values for controls. This hub hemispheric disparity was not correlated with a right atypical language laterality found in six patients. Finally it was shown that a multi-level unsupervised clustering scheme based on self-organizing maps, a type of artificial neural network, and k-means was able to fairly and blindly separate the subjects into their respective patient or control groups. The clustering was initiated using the local nodal centrality measurements only. Compared to the extent of activation network, the intensity of activation network clustering demonstrated better precision. This outcome supports the assertion that the local centrality differences presented by the intensity of activation network can be associated with focal epilepsy.

## TABLE OF CONTENTS

## **CHAPTER**





## LIST OF TABLES



### LIST OF FIGURES



weighted undirected graph, (b) and (d) Thresholded correlation values representing a binary undirected graph. ..37

Figure 13: The thresholding relationship between the absolute value of Pearson's correlation coefficient and its corresponding graph density applied to: (a) a control subject's adjacency matrix, (b) a patient's adjacency matrix. The figure clearly shows that higher correlation thresholds lead to a less dense graph...38

Figure 14: Histograms of the combined correlation values for the patients and controls groups. Controls: Solid Blue, Patients: Dashed Red. (a) Extent of Activation Network, (b) Intensity of Activation Network, (c) Traditional Network. .........................39

Figure 15: The normalized size of the largest connected component across a range of graph densities (10% - 70%) for each subject. Controls: Solid Blue, Patients: Dashed Red. The insets show box plots of the areas under the curve for the patients and the controls. (a) Extent of Activation Network, (b) Intensity of Activation Network.............42

Figure 16: The normalized size of the largest connected component across a range of graph densities (10% - 70%) for each subject. Controls: Solid Blue, Patients: Dashed Red. The insets show box plots of the areas under the curve for the patients and the controls. (c) Traditional Network, (d) Modified Traditional Network. .............................43

Figure 17: Schematic diagram of the dual-level classification system. The system consists of three self-organizing map classifiers followed by a single k-means classifier. The input feature vectors of each patient were: degree centrality, betweenness centrality, and eigenvector centrality..44

Figure 18: Global network features box plots. (a) Average clustering coefficient of the extent of activation network, (b) Average clustering coefficient of the intensity of activation network, (c) Characteristic path length of the extent of activation network, (d) Characteristic Path Length of the intensity of activation network...............................48

Figure 19: Global network features box plots. (a) Average clustering coefficient of the traditional network, (b) Average clustering coefficient of the modified traditional network, (c) Characteristic path length of the traditional network, (d) Characteristic Path Length of the modified traditional network...49

Figure 20: Extent of activation local network features at different False Discovery Rates (FDR). Red nodes indicate patients have a greater value, blue nodes indicate controls have a greater value, and gray nodes indicate no significant difference between the patients and controls groups. The 10% FDR showed no significant nodes...52

Figure 21: Intensity of activation local network features at different False Discovery Rates (FDR). Red nodes indicate patients have a greater value, blue nodes indicate controls have a greater value, and gray nodes indicate no significant difference between the patients and controls groups. The 10% FDR showed no significant nodes...53 Figure 22: Traditional local network features at different False Discovery Rates (FDR). Red nodes indicate patients have a greater value, blue nodes indicate controls have a greater value, and gray nodes indicate no significant difference between the patients and controls groups...54

Figure 23: Clustering results for (a) intensity of activation network and (b) extent of activation network. The first three lines of each subfigure illustrate the results of the self-organizing maps for the: degree centrality, betweenness centrality, and eigenvector centrality feature vectors in order. The last line shows the final and second level k-means clustering results. The subjects on the left side are patients and the subjects on the right side are controls. A red square represents a subject clustered as a patient; whereas, a blue square represents a subject clustered as a control................60

Figure 24: The first two principal components of the different centrality measures. (a) The degree centrality of the extent of activation network. (b) The degree centrality of the intensity of activation network. (c) The betweenness centrality of the extent of activation network. (d) The betweenness centrality of the intensity of activation network. (e) The eigenvector centrality of the extent of activation network. (f) The eigenvector centrality of the intensity of activation network...61

#### **CHAPTER 1**

#### Introduction

#### **1.1 General Statement of Problem Area**

The human brain is a complex functional structure that has confounded scientists for centuries. Numerous studies have tried to unravel its underlying mechanisms in both its normal and pathological states. In particular, language functionality has taken a special interest because of its importance that sets humans apart from other species. Typical functional studies have tried to identify brain regions associated with specific language tasks, i.e. Broca's (Brodmann Area 44 and Brodmann Area 45) and Wernicke's (Brodmann Area 22) (Brodmann, 1909). The advent of recent imaging modalities i.e. Computed Tomography (CT), Positron Emission Tomography (PET), Single Photon Emission Computed Tomography (SPECT), Magneto Encephalography (MEG), and Magnetic Resonance Imaging (MRI) in the last few decades has spurred many relevant brain functional studies (Wu et al., 2013; Wang et al., 2013; You et al., 2012; 2011).

Particularly, Functional Magnetic Resonance Imaging (fMRI) has been widely used because of its safety (no radioactive materials), its balanced spatial/temporal resolution, and its non-invasiveness. The fMRI recording modality relies on detecting Blood Oxygenation Level Dependent (BOLD) signal as an indirect hemodynamic measurement of neuronal activity. The accessibility of fMRI and its use in clinical practice has generated studies of brain functional changes and shifts associated with particular

neurological disorders, such as epilepsy. Moreover, recent studies (Kim et al., 2011) have confirmed that the brain is a highly integrated structure of functional units, as opposed to being composed of isolated compartments responsible for independent tasks. Therefore, it is crucial to investigate the brain as an intricate network of interwoven elements and to try to detect changes in this network caused by neurological disorders in a more meaningful way.

#### **1.2 Research Purpose**

The focus in this dissertation was on language networks in pediatric epilepsy and the goal of the research was threefold:

- 1) Constructing a functional brain network by connecting different spatially independent units generated by Independent Components Analysis (ICA) (Comon, 1994).
- 2) Investigating the functional network topology and characteristics using Graph Theory (Chartrand, 1985) by identifying and extracting relevant features for the task at hand.
- 3) Studying the network changes induced by pediatric epilepsy while taking into account confounding factors such as individual differences, epilepsy type, date of first seizure, number of seizures, and normal/abnormal brain development, among other relevant clinical information.

#### **1.3 Significance of the study**

Functional brain networks are traditionally constructed by correlating the temporal BOLD signal between several voxels in the gray matter (Zhang et al., 2011); alternatively, causal models can be used to investigate the connection between a limited number of predefined regions in the brain (Friston, 2011) to yield a partial depiction of the entire brain network. The functional network construction methods proposed in this dissertation are novel in that they define isolated individual brain units generated by a blind source separation method (ICA) (McKeown et al., 1998) and connect these independent units into an intricate network representing the entire functional brain network.

More importantly, the identification of the independent units is completely data driven with no a priori assumptions. These units hypothetically perform different functions that are integrated to perform a given task. The resulting functional networks were used to provide a deeper understanding of the subtle changes in the language network that may have been caused by the neurological disorder, in this case epilepsy. Moreover, the network features and characteristics might provide a potential biomarker that can be used for early detection of epilepsy or for monitoring the effects of certain medication on the language network. Clinical assessments can also be made post-surgery to evaluate how such functional networks shift or are rearranged to overcome the effect of surgery.

#### **1.4 Research Questions and Hypotheses**

- Question #1: Do language networks show a specific graph topology? If yes, is it completely regular, completely random, or have the structure of small-world networks?
- Question #2: Does pediatric epilepsy induce (or is it associated with) any changes in the language network features or topology? If yes, what types of changes do occur and is the small-world networks architecture still the same as for the healthy controls?
- Question #3: Is there any central language hubs reorganization associated with pediatric epilepsy, if yes is it ipsilateral or contralateral?
- Hypothesis #1: Functional language networks, as constructed from fMRI, show a small-world networks architecture conforming to prior studies of fMRI functional networks.
- Hypothesis #2: Pediatric epilepsy alters the language network features and topology while preserving its small-world networks architecture.
- Hypothesis #3: The major network hubs will shift ipsilateraly causing suboptimal information transfer in the network.

#### **1.5 Structure of the Research**

This dissertation introduces novel methods for constructing functional brain networks and describes a methodology for their application to study the effects of pediatric epilepsy on language networks. The language networks were captured by functional magnetic resonance imaging (fMRI) while performing an auditory descriptive decision task. The dissertation is structured around six chapters starting with this chapter outlining the content of the dissertation.

Chapter 2 describes the necessary background, which lays the foundation for the technical and statistical methods used in the dissertation. It explains the fundamentals of the model and model-free fMRI analysis by introducing the general linear model, the independent components analysis, and the principal components analysis. Thereafter, it introduces basics of graph theory and the mathematical framework used in computing specific global and local graph measures such as the average clustering coefficient and the different centrality measures.

Chapter 3 features the methods used in constructing the language networks. It describes the data collection procedure, the subjects' demographics and how the data is arranged. Then, technical details are presented in carrying out the steps of preprocessing the fMRI datasets, and co-registering a brain atlas to the fMRI space of each subject. The two novel data-driven network construction methods are then presented along with two traditional methods to which they were compared. The chapter, then, introduces an objective thresholding scheme proven to simplify further analyses. Finally, the chapter ends with a description of a unique dual-level clustering system based on neural networks and the kmeans algorithm.

Chapter 4 uses the methods described in chapter 3 to study the effects of pediatric epilepsy on the functional networks constructed. It begins with an analysis of the efficiency of the patients' networks and comparing it to that of a control group. Then it ascertains the causes of the differences found using graph measures. Global network features are studied for all network construction methods, and then the topology of these networks is investigated and compared between the subjects' groups. Furthermore, local network features as captured by different node centrality measures are studied for indepth analysis. The chapter finally ends by using the dual-level clustering scheme, which was introduced in chapter 3, to separate the subject groups using the local network features. The chapter also shows plots of the first two principal components in each construction method to justify the use of the dual-level clustering system.

Chapter 5 highlights the main findings in this dissertation and compares them to other studies on the resulting effects of epilepsy on the different functional networks. It describes the effects of pediatric epilepsy on the language network efficiency. Thereafter, the chapter explores the causes for the differences found by looking at the global and local network features in addition to the network's topology. The chapter includes a description of the reorganizational effects introduced by epilepsy and verifies them in terms of a clustering system.

Finally chapter 6 concludes the dissertation by providing an overview of the main findings and by assessing their clinical relevance. The chapter also suggests other potential uses of the proposed methodology. Lastly, the chapter provides suggestions on future extension of the work towards investigating network attacks and structural/functional network coupling in relation to such neurological disorders.

#### **CHAPTER 2**

#### Literature Review and Background

#### **2.1 Literature Review**

Recent studies have focused on exploring the morphology as well the functional activities of the brain as intricate and often subtle networks of interconnected elements (Essen et al., 2013; Fornito et al., 2013; Hosseini and Kesler, 2013; Sporns, 2013; 2011a; 2011b; Varoquaux and Craddock, 2013). Such elements might be structural or functional depending on the study's objectives. Graph theory (Chartrand, 1985) has played an essential role in offering a formal framework to study such networks. For example, language activation patterns, as captured by fMRI, have been studied extensively in healthy subjects (Gaillard, 2004; K. Wu et al., 2013). Yet, new insights are gained when looking at these activation patterns from a graph perspective.

Additionally, other studies have highlighted the possible effects of different neurological disorders on different brain networks. For example, epilepsy (Liao et al., 2010a; 2010b; Z. Wang et al., 2012), Parkinson's disease (T. Wu et al., 2009), aphasia (Sonty et al., 2007), attention deficit hyperactivity disorder (Milham, 2013; Yu-Feng et al., 2007), Alzheimer's disease (Greicius et al., 2004), and depression (Zeng et al., 2013; 2012). Epilepsy is a 'network disease' (Bonilha et al., 2012), and its effects have consequently been investigated by functional connectivity networks (Vlooswijk et al., 2010; Z. Zhang et al., 2009; Waites et al., 2006), morphological or structural networks (Bernhardt et al., 2011), or a combination of them (Z. Zhang et al., 2011).

Most functional networks studies rely on resting state fMRI; however, in epilepsy, both resting state networks and language networks are disrupted. Therefore, the focus in this dissertation was put on language network connectivity in pediatric epilepsy, during task, from a graph theoretical perspective.

Four functional brain networks were constructed for each subject, a traditional network, a modified traditional network and two independent components analysis (ICA)-based networks. The fMRI independent components were used to construct the two distinct ICA-based networks. These ICA-based networks were constructed to capture both extent and intensity of language related activation separately. All of the large-scale whole brain networks were used to study the effects of epilepsy on the network hemispheric reorganization and topology. Afterwards, certain network features were used to blindly cluster the subjects into two groups representing the patients and controls. The unsupervised clustering was performed to ascertain if the disease did indeed induce, or is associated with, the changes observed in the network features.

#### **2.2 The General Linear Model (GLM)**

fMRI statistical data analysis translates the BOLD temporal signal into statistically meaningful activation maps. Typical data analysis is carried out as several steps in a pipeline. The main components of this pipeline are:

- 1) Preprocessing the data to remove noise and artifacts (Jezzard et al., 2002; Penny et al., 2006).
- 2) Setting up a statistical model employing a priori knowledge about the activation pattern while taking into account other confounding factors, i.e. head motion artifacts and other vital signals that might appear as noise (Jezzard et al., 2002; Penny et al., 2006).
- 3) Finding the best fit between the model and the preprocessed BOLD signal using the GLM.
- 4) Executing a post-statistical analysis while correcting for the multiple comparison problem using Gaussian Random Field Theory (Worsley and Friston, 1995).

Moreover, most fMRI studies include several subjects or groups of subjects (Calhoun & Adali, 2012) that oftentimes make it necessary to register each subject to a brain template (Fonov et al., 2011) in order to compare the groups or subjects. The order of these steps can slightly differ depending on the research protocol or the software used; yet, the outcome of any of the pipelines is a single statistical map representing the activation associated with the used task paradigm. Figure 1 illustrates an example of such an activation statistical map as generated by GLM. This method is highly constrained by the statistical model used, which in turn depends on the task paradigm and the hemodynamic response function selected.



Figure 1: GLM's statistical activation map of a control subject while performing an auditory descriptive decision task.

#### **2.3 Independent Components Analysis (ICA)**

ICA is a blind source separation algorithm that splits a set of signals into their associated latent sources; Figure 2 shows an example of source and mixture signals. ICA does not assume any a priori knowledge of these sources. The only constraint imposed on these sources is that they are statistically independent and at most one of them is Gaussian. Figure 3 illustrates a normal distribution and a uniform distribution histograms. Figure 4 shows an example of the geometry of unmixing a uniformly distributed set of mixtures and compares it to that of unmixing a set of normally distributed signals, clearly the normally distributed signals are not easily unmixed.



Figure 2: Example of source signals and a set of corresponding mixture signals. ICA is a blind source separation algorithm that solves the problem of extracting the sources from their mixtures.



Figure 3: Statistical distributions. (a) Normal distribution. (b) Uniform Distribution



Figure 4: The geometry of unmixing two uniformly distributed sources and normally distributed sources.

ICA appeared in 1991 (Jutten and Herault, 1991); yet, the term was first coined by (Comon, 1994). After its inception, most research groups used the Maximum Likelihood Estimation algorithm (MLE) (Pham et al., 1992) to perform ICA. (Bell and Sejnowski, 1995) used a more complex iterative algorithm based on maximizing mutual information (infomax). Subsequently (Cardoso, 1997) proved that MLE and infomax are congruent algorithms. In 1999 (Hyvarinen, 1999) introduced an elegant fixed-point algorithm (FastICA) where he used negentropy, a normalized differential entropy or a measure of non-gaussianity, as a cost function. He also proved that non-gaussianity is equivalent to independence.

ICA in fMRI appeared in 1998 as the first model-free method to extract activation maps as opposed to the GLM analysis explained in section 2.2, which heavily depends on an a priori model. The results of this first fMRI ICA study, which are astounding, were several spatially independent maps associated with unconstrained temporal signals. Indeed (McKeown et al., 1998) showed that one of the maps highly correlates with the output of the GLM. Nevertheless, there were many other maps that were harder to interpret. Figure 5 illustrates an example of an independent component that is highly correlated with the performed task as highlighted in Figure 1.

In 2004 (Beckmann and Smith, 2004) introduced Probabilistic ICA (PICA). This particular ICA approach used FastICA as its backend algorithm and included a noise term in the model. This noise model resulted in statistically valid ICA maps as opposed to the arbitrary maps in (McKeown et al., 1998). Additionally, PICA had a model order estimation step (Roberts and Everson, 2001) that estimated the optimal number of independent components. The output of PICA is optimal in that each spatial map is optimally integrated and represents one activation unit as opposed to being separated into several maps that are harder to interpret. Figure 6 illustrates a set of independent components for a control subject while performing the ADDT language task.



Figure 5: An independent component of a control subject, which is highly correlated with the task paradigm, along with its temporal profile. The control subject is the same one used to generate the activation map in Figure 1.



Independent Component 9 Figure 6: Example of several independent components.



Independent Component 11

Independent Component 12

Figure 6 (continued): Example of several independent components for a control subject while performing a language task. The components were randomly selected from a set of 32 independent components.

#### **2.4 Principal Components Analysis (PCA)**

Principal components analysis is generally used as a reliable and fast dimensionality reduction algorithm. PCA works by separating a set of mixture signals into a set of orthogonal and uncorrelated signals. The set of orthogonal signals can be used to reduce the dimensions of the original mixture signals by keeping signals with high variance (or Eigenvalues). Unlike ICA, PCA is not a blind source separation algorithm because it uses correlation as opposed to independence as a cost function. PCA is the core algorithm used in constructing the modified traditional networks introduced later in section 3.7.4.

#### **2.5 Graph Measures**

Graph theoretical analysis was carried out on the functional networks. Graph theory provides many global and local quantitative measures to analyze the brain network dynamics. The following succinctly summarizes the measures used in this dissertation. Most of these measures were implemented using functions from the Python NetworkX library (Hagberg et al., 2008).

#### **2.5.1 Clustering Coefficient**

The clustering coefficient  $c_i$  of a node  $n_i$  captures the cliquishness within its neighborhood  $k_i$ . The clustering coefficient of a node in a binary undirected graph is alculated as the fraction of the number of edges  $E_i$  to the total number of possible edges within the node's neighborhood (Watts and Strogatz, 1998). This coefficient is formulated as follows:

$$
c_i = \frac{E_i}{\frac{k_i(k_i - 1)}{2}}\tag{1}
$$

Thereafter, the clustering coefficient of the whole network was calculated as the average clustering coefficient of all its constituent nodes *n*:

$$
C = \frac{1}{n} \sum_{i=1}^{n} c_i
$$
 (2)

#### **2.5.2 Characteristic Path Length**

The shortest path length  $d(n_i, n_j)$  between all  $n_i$  and  $n_j$  where  $i \neq j$  was calculated. Thereafter, these path lengths were averaged to get the characteristic path length as expressed in equation 3:

$$
L = \frac{1}{n(n-1)} \sum_{i \neq j} d(n_i, n_j)
$$
 (3)

#### **2.5.3 Small-World Index**

Large-scale brain networks are typically studied using Graph Theory (Sporns, 2011a). In (Watts and Strogatz, 1998), different possible network topologies were discussed and regular networks were defined as having a high clustering factor and relatively longer path lengths, whereas, random networks have a low clustering factor with very short path lengths (Bollabas, 1985). They also demonstrated that small-world networks are the middle ground between the two extremes with high clustering factor and short characteristic path length. Small-world networks topology has higher information transfer efficiency, better synchronization, and is usually associated with sparse graphs, which have more vertices than edges. Several studies have showed that typical and atypical brain networks have a small-world architecture and that neurological disorders introduce changes to the network topology while maintaining the overall small-world networks architecture (Bernhardt et al., 2011; Zhang et al., 2011, Sporns, 2011b).

Each subjects' functional brain network was compared to 100 Erdős Rényi random graphs (Erdős and Rényi, 1959); these random graphs were generated with a connection probability that equals the brain graph density. The averages of the 100 clustering coefficients *Crandom* and the characteristic path lengths *Lrandom* were compared to the subject's brain clustering coefficient *Cnet* and characteristic path length *Lnet*. Compared to random networks, a small-world network has greater clustering coefficient  $y =$  $C_{net}/C_{random} > 1$  and similar characteristic path length  $\lambda = L_{net}/L_{random} \approx 1$  (Watts and Strogatz, 1998). The two small-world features can be combined into a single scalar index  $\sigma = \gamma/\lambda$  which is typically greater than 1 for small-world networks.

#### **2.5.4 Degree Centrality**

The normalized degree centrality  $C_D$  of a node  $n_i$  in a graph with  $n$  nodes is simply the fraction of the number of connection  $k_i$  the node has over the total number of possible connections *(n-1)*:

$$
C_D(n_i) = \frac{k_i}{(n-1)}\tag{4}
$$

This degree centrality measure provides an indication of a nodes importance in a network in terms of the number of nodes it connects to. The measure was used in this dissertation to study the local network characteristics and as a feature vector for an unsupervised clustering system.

#### **2.5.5 Betweenness Centrality**

The betweenness centrality  $C_B$  of a node  $n_i$  in a graph measures its participation in the formation of the graph's shortest paths. It can be calculated by the ratio of the number of shortest path lengths which  $n_i$  is part of  $\delta(n_k, nj | n_i)$ ,  $\forall k \neq j$ , to the total number of shortest path lengths in the graph  $\delta(n_k, n_i)$ ,  $\forall k \neq j$  (Brandes, 2001):

$$
C_B(n_i) = \sum_{k \neq j} \frac{\delta(n_k, n_j | n_i)}{\delta(n_k, n_j)}
$$
(5)

This measure was also used to aid in exploring the local network characteristics and as one of the input feature vectors for a data-driven clustering system.

#### **2.5.6 Eigenvector Centrality**

Eigenvector centrality measures the importance of a node in a graph; it is a referential measure that gives higher values to nodes that connect to higher value nodes. In other words, it assigns high values to nodes communicating with central nodes in the network (Bonacich, 1987).

#### **CHAPTER 3**

#### Functional Network Construction

#### **3.1 Participants and Data Collection**

The data used in this dissertation was collected by a multisite consortium and repository for pediatric epilepsy. This repository is established by the Center for Advanced Technology and Education (CATE), which host the repository at (mri-cate.fiu.edu). The repository, which was funded by the American Epilepsy Society in collaboration with the Children's National Medical Center in Washington D.C., includes 13 leading pediatric hospitals here in the US and overseas. The consortium was formed with the intent to study the effects of pediatric epilepsy on the brain structure and function (Lahlou et al., 2006; You et al., 2013; 2011; Wang et al., 2013a; Sun et al., 2013).

The datasets selected for this dissertation came from three leading pediatric hospitals within the consortium: British Columbia Children's Hospital (BCCH), Children's Healthcare of Atlanta (CHOA), and Children's National Medical Center (CNMC). All three locations used a 3 Tesla Siemens Trio MRI scanner, and were selected mainly in seeking high resolution MRIs, as we have demonstrated in earlier studies that statistical analysis proved independence of scanner, whether a 1.5T or 3T was used. IRB requirements were followed where the parents gave written informed consent and children gave assent. The datasets were de-identified to insure confidentiality.
All subjects underwent fMRI data acquisition while performing an Auditory Descriptive Decision Task (ADDT) devised to stimulate the temporal and the inferior frontal cortex as described in (Berl et al., 2012; Gaillard et al., 2007). The subjects were shown an object and then were subjected to an auditory stimulus describing the object. If the description matched the object the subjects were instructed to press the 'True' button, otherwise they were instructed to press the 'False' button. The description was repeated every three-second period where a 'True' pair appeared pseudo-randomly with a 70% chance. At rest the subjects listened to description in reverse speech and were coached to press a button on beeps generated following the audio. Furthermore, the difficulty of the paradigm was adjusted appropriately to match the subject's age group.

A total of 29 pediatric epilepsy patients and 30 age and sex matched control subjects were recruited. The patients were between 9.5 and 18.5 years with an average age of 13.5±2.45 years, 14 females, 21 right handed, 17 with a remote symptomatic seizure etiology, 5.5±4.70 years average age of first seizure, 24 left hemispheric focus, 4 right hemispheric focus, 1 bilateral focus; all patients went through pre-surgical evaluation, and 13 patients underwent epilepsy surgery. The controls were between 10 and 20 years of age with an average of 13.5±2.98 years, 14 females, 18 right handed, and 12 with unknown handedness. All controls were native English speakers free of any current or past neurological or psychiatric disease. Furthermore, there were six patients with right language laterality whereas all controls had left language laterality. The appendices list the detailed clinical and demographical information of each subject and Table 1 summarizes the demographics.

	Age (years)			Females	<b>Handedness</b>			Language Laterality		
	Min	Max	Average	<b>STD</b>		<b>Right</b>	Left	Unknown	Left	<b>Right</b>
<b>Patients</b> (29)	9.5	18.5	13.5	2.45	14	21	6	C.	23	
<b>Controls</b> (30)	10	20	13.5	2.98	14	18	0	12	30	

Table 1: Summary of the Demographics for Patients and Controls

In addition, a high resolution isotropic structural T1 MRI scan was acquired for each subject. These T1 scans aid in registering each subject's fMRI space to a common stereotaxic space, which is defined by a brain template. The T1 scan parameters were: BCCH= (211x288x288) voxels, (1x1x1) mm; CHOA= (176x240x256) voxels, (1x1x1) mm; and CNMC= (256x256x30) voxels, (0.8x0.8x4) mm. The fMRI scan parameters were: BCCH= (128x128x36) voxels, (1.87x1.87x3.97) mm, TR=2 sec, 150 time points; CHOA=  $(72x72x49)$  voxels,  $(3x3x3.3)$  mm, TR=3 sec, 100 time points; and CNMC= (64x64x30) voxels, (3.44x3.44x4) mm, TR=2 sec, 150 time points. Table 2 summarizes each hospital's contribution to the datasets used in this dissertation.

**Institution Scanner Number of Subjects Number of Controls fMRI Field of View [mm] T1 Field of View [mm] BCCH** Siemens<br>Trio 3T 24 0 239.36x239.36x142.92 211x288x288 CHOA Siemens Trio 3T | 12 | 12 | 216x216x161.7 | 176x240x256 CNMC Siemens Trio 3T 23 18 220.16x220.16x120 204.8x204.8x120

Table 2: Subjects distribution across hospitals

## **3.2 Language Laterality in Patients and Controls**

The language laterality index was calculated by combining a bootstrap procedure with a histogram analysis (Wang et al., 2013a; Wilke and Schmithorst, 2006). The masked areas were thresholded, and then data was obtained for the left and right sides. The data was converted into a vector whose elements were used in a bootstrapped resampling method to compute the lateralization index as given in equation (6):

$$
LI = \frac{\sum activation_{Left} - \sum activation_{Right}}{\sum activation_{Left} + \sum activation_{Right}} \tag{6}
$$

All the possible LI values were plotted in a histogram, from which the mean of the 50% central values was used as the selected LI value. The threshold was obtained by the mean intensity of the voxels in the image. The toolbox used for these calculations is described in (Wilke and Lidzba, 2007).

In this work, the laterality index was computed for Broca's and Wernicke's areas. If the index was greater than or equal to 0.2 the area's laterality was considered left dominant, if the index was less than or equal to -0.2 the area's laterality was considered right dominant, otherwise the dominance was considered bilateral. Finally, the general language laterality was considered right dominant if both areas showed right dominance or one was right dominant and the other was bilateral. Similarly, the general language laterality was considered left dominant if both areas demonstrated left dominance or one was bilateral and the other was left dominant. If both areas were bilaterally dominant or one was left and the other was right dominant the general language laterality was considered bilaterally dominant. Table 3 lists the detailed results for each subject.

	Broca's	Wernicke's			
	Laterality	Laterality	Broca's	Wernicke's	Language
	<b>Index</b>	<b>Index</b>	Laterality	Laterality	Laterality
BCCH 11	$-0.0015$	$-0.45$	<b>Bilateral</b>	Right	Right
BCCH 12	0.37	0.018	Left	<b>Bilateral</b>	Left
BCCH 13	0.55	0.72	Left	Left	Left
BCCH 14	0.68	0.77	Left	Left	Left
BCCH 20	0.26	0.48	Left	Left	Left
BCCH <sub>22</sub>	0.83	0.77	Left	Left	Left
BCCH <sub>23</sub>	0.81	0.58	Left	Left	Left
BCCH 29	0.62	0.65	Left	Left	Left
BCCH 30	0.65	0.43	Left	Left	Left
BCCH 31	0.5	0.64	Left	Left	Left
BCCH 33	0.68	0.57	Left	Left	Left
BCCH 34	$-0.36$	0.1	Right	Bilateral	Right
BCCH 35	$-0.52$	$-0.54$	Right	Right	Right
BCCH 36	$-0.71$	$-0.71$	Right	Right	Right
BCCH 37	0.34	0.31	Left	Left	Left
BCCH 38	$-0.57$	$-0.13$	Right	<b>Bilateral</b>	Right
BCCH 39	0.71	0.79	Left	Left	Left
BCCH 40	0.53	0.66	Left	Left	Left
BCCH 41	0.64	0.73	Left	Left	Left
BCCH 43	0.41	0.53	Left	Left	Left
BCCH 44	0.29	$-0.054$	Left	<b>Bilateral</b>	Left
BCCH 45	$-0.83$	$-0.81$	Right	Right	Right
BCCH 6	0.75	0.56	Left	Left	Left
BCCH <sub>8</sub>	0.34	0.51	Left	Left	Left
<b>CNMC 320</b>	0.58	0.74	Left	Left	Left
<b>CNMC 331</b>	0.57	0.6	Left	Left	Left
<b>CNMC 333</b>	0.19	0.67	<b>Bilateral</b>	Left	Left
CNMC 82	0.32	0.59	Left	Left	Left
CNMC 84	0.5	0.69	Left	Left	Left
<b>CNMC</b> 105	0.51	0.14	Left	<b>Bilateral</b>	Left
<b>CNMC</b> 106	0.74	0.61	Left	Left	Left

Table 3: Detailed laterality index results for each subject.



## **3.3 Data Arrangement**

As the data is collected from several institutions, the data was not uniform and needed manual intervention to prepare it for the next processing steps. First each dataset was

thoroughly inspected to insure image quality; datasets with low resolution, high artifact, or high motion errors were excluded from the study. Thereafter, all datasets were renamed to match a predetermined string pattern. The pattern consisted of the institution's acronym followed by an underscore then a unique subject identifier provided by the hospital, for example subject 12 from British Columbia Children's Hospital was coded as BCCH\_12. This renaming arrangement was necessary for later bulk processing of the data. All datasets were then converted to the Neuroimaging Informatics Technology Initiative (NIFTI) format. Finally, each dataset was reoriented to a neurological coordinate system matching that of the Montreal Neurological Institute 152 (MNI152) brain atlas. The MNI152 neurological convention assumes that the *x* axis is perpendicular to the sagittal plane of the brain with the positive side going to the right side of the brain, the *y* axis is perpendicular to the coronal plane with the positive side going to the anterior side of the brain, and the *z* axis perpendicular to the axial plane with the positive side going to the superior side of the brain. It is noted that only subjects with high resolution fMRI and T1 datasets were selected for this dissertation, such high resolution helps reduce image registration errors as explained in section 3.5.

## **3.4 fMRI Preprocessing**

Each subject's fMRI dataset was preprocessed using the FMRIB Software Library (FSL) (Jenkinson et al., 2012; S. Smith et al., 2004) as follows: temporally high pass filtered with a cutoff frequency of 0.01 Hz to remove the MRI scanner's baseline wandering effect, head motion corrected using MCFLIRT (Jenkinson et al., 2002), slice time corrected, deskulled to remove non-brain tissues using Brain Extraction Tool (BET) (S.

M. Smith, 2002), and spatially smoothed with a 5 mm full width at half maximum (FWHM) to increase the signal to noise ratio. Datasets were inspected to ensure that head motion did not exceed 1 mm in any of the major axes directions. Afterwards, each of the preprocessed datasets was passed to the probabilistic independent component analysis (PICA) algorithm to get its spatially independent latent sources as implemented in MELODIC (Beckmann and S. Smith, 2004; Beckmann, 2012). The PICA algorithm provides spatially independent maps with their corresponding temporal profiles as explained earlier in 2.3.

## **3.5 Co-registering the AAL90 atlas to the fMRI Space**

The Automated Anatomical Labeling (AAL90) atlas (Tzourio-Mazoyer et al., 2002) includes 90 cortical and sub-cortical regions in the MNI152 space; these regions are symmetrically distributed on each hemisphere (45 regions on each hemisphere). The regions were manually drawn every 2 mm on the axial planes using the single subject T1 atlas provided by the MNI (Collins et al., 1998). Figure 7 show the AAL90 atlas overlaid on top of the MNI152 atlas. The AAL90 atlas was registered to each subject's fMRI space; afterwards, it was superimposed over each of the subject's spatially independent components.



Figure 7: The Automated Anatomical Labeling (AAL90) atlas overlaid on the MNI152 atlas.

For each subject the registration was performed as follows: deskulling of the fMRI dataset, deskulling of its corresponding T1 both using the Brain Extraction Tool (BET) (S. M. Smith, 2002), registering the deskulled fMRI to the deskulled T1 using an affine transformation (12 degrees of freedom), registering the deskulled T1 to the MNI152 brain also using an affine transformation. The two aforementioned registration steps were concatenated into a single transformation matrix. This single matrix was then inverted and applied to register the AAL90 atlas to the subject's fMRI space. Each registration step was performed using the FMRIB's Linear Image Registration Tool (FLIRT) (Jenkinson et al., 2002).

## **3.6 The Processing Environment**

All data processing in this dissertation was performed on a computing environment that consisted of three computers connected in a grid. The grid was setup using the Grid Engine environment also known as Sun Grid Engine (Gentzsch, 2001). This grid contained two iMac computers running Mac OS X version 10.7.5 each has 4GB of memory and a 3.06 GHz Intel Core 2 Duo, the grid also had a Mac Server running Mac OS X version 10.7.5 server with 8 GB of memory and 2.26 GHz Intel Xeon Quad-Core. This computing environment was optimized to process all datasets in a parallel fashion, for example all subjects were processed simultaneously using FSL optimized commands.

## **3.7 Functional Network Construction**

Networks or graphs consist of a group of nodes connected by edges. The goal in this implementation step was to determine a functional brain network, which is essentially defining the nodes and connecting them by valid edges. The nodes were defined as the different 90 regions of the AAL90 atlas. Thus, these nodes will be comparable across subjects. Figure 8 illustrates an example of such a brain network. In this dissertation, two novel methods to connect these nodes based on the extent and intensity of the language related activation were proposed; thereafter, these two novel methods were compared and contrasted against a traditional method and a modified version of it.



Figure 8: Example of a functional brain network where the nodes are based on the AAL90 atlas.

## **3.7.1 Extent of Activation Network**

After co-registering the AAL90 atlas to the subject's fMRI space, the atlas is superimposed over each spatial independent component (IC). Hence, each of the 90 atlas regions will have several accompanying independent components. For each of these regions the number of activated voxels is counted at each IC. The temporal profile of the IC with the highest activated voxel count is associated with the particular atlas region. Eventually, every region in the atlas will be associated with a single temporal signal. A

90x90 correlation matrix is thus constructed using Pearson's correlation between the temporal signals for each subject. The resulting matrix is a graph adjacency matrix representing the functional brain network of the subject at hand while capturing its extent of activation. Figure 9 illustrates examples of such matrices for a control subject and a patient along with their associated binary undirected versions.



## Extent of Activation Network

Figure 9: Example of an adjacency matrix of an extent of activation network for a control subject and a patient: (a) and (c) Raw correlation values representing a weighted undirected graph, (b) and (d) Thresholded correlation values representing a binary undirected graph.

## **3.7.2 Intensity of Activation Network**

Similarly, an adjacency matrix capturing the intensity of activation can be constructed by repeating the same steps as in section 3.7.1, but instead of counting the number of activated voxel, the average absolute *z* values of the activated voxels is computed within each region at each IC. The resulting graph adjacency matrix represents the functional brain network of the subject at hand while capturing its intensity of activation. Figure 10

Intensity of Activation Network



Figure 10: Example of an adjacency matrix of an intensity of activation network for a control subject and a patient: (a) and (c) Raw correlation values representing a weighted undirected graph, (b) and (d) Thresholded correlation values representing a binary undirected graph.

illustrates an adjacency matrix example of an intensity of activation network for a control subject and a patient along with their binary undirected matrix.

## **3.7.3 Traditional Network**

For each subject the AAL90 atlas is registered to the fMRI space and then superimposed over the preprocessed fMRI as described in section 3.5. Then, for each of the 90 regions



Traditional Network

Figure 11: Example of an adjacency matrix of an traditional network for a control subject and a patient: (a) and (c) Raw correlation values representing a weighted undirected graph, (b) and (d) Thresholded correlation values representing a binary undirected graph.

the average of all BOLD signals is computed within that region; the averaged signal represents that region. A 90x90 correlation matrix is constructed using Pearson's correlation. The resulting adjacency matrix represents a traditional functional brain network. Figure 11 illustrates an example adjacency matrix of this traditional network for a control subject and a patient along with their binary undirected matrix.

## **3.7.4 Modified Traditional Network**

Likewise, instead of averaging the BOLD signals within each region in the traditional method as explained in section 3.7.3, the first principal component of the BOLD signals within a region, which captures the largest variance, can instead be used to represent that region. The resulting adjacency matrix represents a modified traditional functional brain network. Figure 12 illustrates an example of the adjacency matrix of a modified traditional network for a control subject and a patient along with their binary undirected matrix.

## **3.8 Thresholding the Adjacency Matrix**

All graphs constructed using either one of the connectivity methods, including the ICAbased methods, are undirected weighted graphs. In this dissertation the focus was placed on connectivity whether two nodes were connected or not; therefore, the absolute value of the adjacency matrix was thresholded to an unweighted form. In this section, a thresholding scheme based on graph density was established to facilitate the selection of an objective threshold comparable across all subjects.

# **Modified Traditional Network**



Figure 12: Example of an adjacency matrix of a modified traditional network for a control subject and a patient: (a) and (c) Raw correlation values representing a weighted undirected graph, (b) and (d) Thresholded correlation values representing a binary undirected graph.

A graph density can be defined as:

$$
D = \frac{Number\ of\ edges}{Number\ of\ all\ possible\ edges}
$$
 (7)

For a fully connected graph, where all nodes are connected directly to all other nodes,  $D = 1$ . On the other extreme, a disconnected graph will yield  $D = 0$ . Density is thus a

measure of a network's wiring cost. Therefore, thresholding using this measure would facilitate comparison across networks. For example, two brain networks with the same density will evidently have the same number of nodes and the same number of edges (same wiring cost). However, the edges are shifted reflecting the state of the subject's network.



Figure 13: The thresholding relationship between the absolute value of Pearson's correlation coefficient and its corresponding graph density applied to: (a) a control subject's adjacency matrix, (b) a patient's adjacency matrix. The figure clearly shows that higher correlation thresholds lead to a less dense graph.

Figure 13 shows the relation between thresholding using Pearson's correlation coefficient and the graph density (for a control subject and for a patient). Clearly, higher correlation threshold values lead to lower density graphs. A bidisectional algorithm is thus used to determine the corresponding correlation threshold value given a selected density. For example, to determine the Pearson's correlation threshold value  $\overline{PC}$  that is associated with a certain density value  $D$ , a Pearson's correlation value range is defined by a low limit  $L_L$ , which is initially 0, and a high limit  $L_H$ , which is initially 1. Thereafter, the mid range value is computed as  $L_M = (L_L + L_H)/2$  and the corresponding density  $D_i$  is

calculated. If  $D_i$  is less than D then  $L_H$  is set to  $L_M$  otherwise  $L_L$  is set to  $L_M$ . This process is then iteratively repeated till  $D_i$  approaches the selected density  $D$  with an infinitesimal error  $\varepsilon$  at which point  $L_M$  will be the Pearson's correlation threshold PC that corresponds to  $D$ .

To justify using the graph density as opposed to the correlation values for thresholding, raw histogram correlation values of the patients and the controls are used to gauge both the extent and intensity of activation distributions in terms of Pearson's correlation versus the number of edges. Figure 14 shows these histograms for the ICA-based methods in contrast to the traditional network construction method, which uses the BOLD average within each region of the atlas as the representative temporal profile.

Figure 14 illustrates a slight distinction between the patients and controls for the traditional method, which agrees with (Fornito et al., 2013) in schizophrenia; whereas, the ICA-based methods produced correlation values that are comparable between the two subject groups. This makes the density threshold selection more neutral towards the



Figure 14: Histograms of the combined correlation values for the patients and controls groups. Controls: Solid Blue, Patients: Dashed Red. (a) Extent of Activation Network, (b) Intensity of Activation Network, (c) Traditional Network.

correlation value. In other words, selecting a specific density threshold for all subjects will produce correlation thresholds that are similar between the groups. Whereas, the traditional method produces different correlation values for the same density value in each subject group. Consequently, by using the proposed approach, the ambiguity as to which threshold is to be selected (Pearson's correlation or density) is resolved. The binary undirected graphs in Figure 9, Figure 10, Figure 11, and Figure 12 show examples of such thresholded adjacency matrices.

It is important to note that a density threshold that guarantees a connected graph with the cheapest wiring cost was the one adopted in this dissertation. A connected graph is henceforth a network where every node will have a path to any other node in the network, directly or indirectly. The normalized size of the largest connected component was used to measure a graph connectedness. A connected component in a graph is a group of nodes that can reach each other within the component; i.e. in a connected graph all nodes can reach each other and the largest connected component contains all nodes, and hence it's normalized size is 1. Similarly, if there are several connected components within the network, which do not have any connecting links, then the size of the largest connected component will be less than 1.

For each network construction method, the normalized size of the largest connected component was computed for every subject across a density range (from 10% to 70%) assumed on the basis of the results observed in Figure 15 and Figure 16, which show that after 65% all the subjects had a connected graph (normalized size of the largest connected  $component = 1$ , at which point the area under the resulting curve was determined. Figure 15 and Figure 16 illustrate these typical curves. The figures also show that density thresholds of 65%, 55%, 65%, and 70% guarantee a connected graph for all subjects in the extent of activation network, intensity of activation network, traditional network, and the modified traditional network, respectively. Additionally, these thresholds are all in the range suggested in (Reus and van den Heuvel, 2013) for structural networks and are very close to the 60% recommended value. In this dissertation, all subjects' networks were thresholded using these values and the resulting binary graphs were used for all further analyses.



Figure 15: The normalized size of the largest connected component across a range of graph densities (10% - 70%) for each subject. Controls: Solid Blue, Patients: Dashed Red. The insets show box plots of the areas under the curve for the patients and the controls. (a) Extent of Activation Network, (b) Intensity of Activation Network.



Figure 16: The normalized size of the largest connected component across a range of graph densities (10% - 70%) for each subject. Controls: Solid Blue, Patients: Dashed Red. The insets show box plots of the areas under the curve for the patients and the controls. (c) Traditional Network, (d) Modified Traditional Network.

## **3.9 Unsupervised Clustering**

This step aims at classifying the subjects into distinct groups in an unsupervised and datadriven approach. A dual-level clustering scheme was used, the first level consisted of three self-organizing map (SOM) classifiers; whereas, the second level, which aggregates and clusters the outputs of the first level, consisted of a single k-means classifier with  $k =$ *2*. Figure 17 describes the structure of this classification system. Three feature vectors were computed for each subject based on the degree, betweenness, and eigenvector centrality; each of these feature vectors were set up as 1x90-dimentional vector on the basis of the AAL90 atlas. Next, each feature vector was used as an input for one of the SOM classifiers. SOM is a type of artificial neural network.



Figure 17: Schematic diagram of the dual-level classification system. The system consists of three self-organizing map classifiers followed by a single k-means classifier. The input feature vectors of each patient were: degree centrality, betweenness centrality, and eigenvector centrality.

In this dissertation, 500 training steps were used and an initial neighborhood size of 3 was assumed. The performance of this classification system was then evaluated by computing its precision, sensitivity, specificity, and accuracy. Formulae (8) through (11) define these metrics in terms of the:

- 1) True Positive (TP): A patient is clustered as a patient.
- 2) False Positive (FP): A control is clustered as a patient.
- 3) True Negative (TN): A control is clustered as a control.
- 4) False Negative (FN): A patient is clustered as a control.

$$
Precision = \frac{TP}{TP + FP}
$$
 (8)

$$
Sensitivity = \frac{TP}{TP+FN}
$$
 (9)

$$
Specificity = \frac{TN}{TN + FP}
$$
 (10)

$$
Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{11}
$$

### **CHAPTER 4**

The Effects of Pediatric Epilepsy on the Language Networks

## **4.1 Density and Thresholding Analyses**

Figure 15 and Figure 16 shown earlier provided the normalized size of the largest connected component as a function of graph density. Figure 15 demonstrated that both the extent of activation networks and the intensity of activation networks were able to partially delineate the patients from the controls. Whereas, Figure 16 has clearly proven that the traditional networks and the modified traditional networks were not able to delineate the patients from the controls.

Furthermore, for the ICA-based networks, most patients had a very large connected component with relatively small density values. This observation implies that patients' brain tend to employ more regions to perform the language task; whereas, the controls tend to compartmentalize the brain into separate smaller connected components when performing the ADDT task. This is especially apparent in the extent of activation networks. To confirm these observations statistically, the area under the curve was calculated for each subject then the patient population was compared to the control population using a boxplot and a *t*-test.

The *t*-test confirmed that both of the ICA-based networks could separate the patients and controls groups with a  $p = 0.001$ , whereas the other two traditional methods were not able

to separate the groups  $p = 0.295$  for the traditional method and  $p = 0.14$  for the modified traditional method. The boxplots provided as insets in Figure 15 show that patients have a greater area as compared to controls; however, the boxplots in Figure 16 did not display such a distinction. The differences observed using the ICA-based methods might help explain the effects of pediatric epilepsy on the language networks; therefore, for the rest of the dissertation the focus is set on analyzing the data-driven ICA-based networks and the traditional methods' results are provided as reference for comparative purposes.

### **4.2 Global Network Features**

In order to study the brain language network general dynamics and topology, the global network features were assessed, and the average clustering coefficients were calculated for all networks. The extent of activation network showed a significant clustering difference between the patient and control groups  $p = 0.001$ ; whereas, the intensity of activation network did not yield a significant clustering difference  $p = 0.728$ .

Furthermore, the characteristic path lengths were calculated for both networks. Neither of the networks showed a significant characteristic path length difference between the patients and controls,  $p = 0.962$  for the extent of activation network and  $p = 0.066$  for the intensity of activation network, respectively. Although the two networks did not show any significant characteristic path length difference, they had different path lengths with an average of  $1.35\pm0.03$  for the extent of activation network, and an average of  $1.46\pm0.04$ for the intensity of activation network. Figure 18 summarizes these findings using box plots.



Figure 18: Global network features box plots. (a) Average clustering coefficient of the extent of activation network, (b) Average clustering coefficient of the intensity of activation network, (c) Characteristic path length of the extent of activation network, (d) Characteristic Path Length of the intensity of activation network.

Similarly, the traditional and modified traditional methods did not show any significant average clustering coefficient difference between the patients and controls with  $p = 0.038$ for the traditional method and  $p = 0.310$  for the modified traditional method. Moreover, there was no significant characteristic path length difference between the subject groups with  $p = 0.262$  for the traditional method and  $p = 0.044$  for the modified traditional. Figure 19 illustrates these observations in the form of box plots.

The functional network topology of each subject was investigated by comparing it to 100 Erdős Rényi random networks as explained earlier in section 2.5.3. The average



Figure 19: Global network features box plots. (a) Average clustering coefficient of the traditional network, (b) Average clustering coefficient of the modified traditional network, (c) Characteristic path length of the traditional network, (d) Characteristic Path Length of the modified traditional network.

 $\gamma = C_{net}/C_{random}$  for all subjects was 1.17±0.06 for the extent of activation networks,  $1.27\pm0.01$  for the intensity of activation networks,  $1.18\pm0.07$  for the traditional networks, and 1.13 $\pm$ 0.06 for the modified traditional networks. Similarly, the average  $\lambda$  =  $L_{net}/L_{random}$  of all subjects was 1 $\pm 0.002$  for the extent of activation networks, 1 $\pm 0.01$ for the intensity of activation networks,  $1\pm 0.003$  for the traditional networks, and  $1\pm 0.006$ for the modified traditional networks. Subsequently, each network method thresholded with the optimal wiring cost resulted in small-world indices  $\sigma = \gamma/\lambda$  greater than 1 with  $p = 1.0E-13$  and averaged  $1.16\pm0.06$ ,  $1.28\pm0.08$ ,  $1.181\pm0.07$ , and  $1.131\pm0.06$  for the

extent of activation, intensity of activation, traditional, and modified traditional networks, respectively.

Results have also shown that each of the ICA-based network construction methods generated small-world networks. To distinguish the topology differences between the subject groups, the small-world network parameters were compared between the groups using the Mann-Whitney U nonparametric test. For the extent of activation network the clustering parameter  $\gamma$  was higher for controls as compared to patients with  $p = 0.008$ ; whereas, the path length parameter  $\lambda$  was not different between the subject groups with  $p$  $= 0.739$ . Consequently, the small-world index  $\sigma$  was higher for controls compared to patients  $p = 0.012$ . Similarly, there was no difference between the groups for any of the small-world parameters in the intensity of activation network  $p = 0.524$  for  $\gamma$ ,  $p = 0.832$ for  $\lambda$ , and  $p = 0.524$  for  $\sigma$ . Recall that  $\sigma = \gamma/\lambda$ .

## **4.3 Local Network Features**

After investigating the global network features, the focus was shifted to the individual node features and the effect of epilepsy on such features. Three local network features were used: degree centrality, betweenness centrality, and eigenvector centrality all of which were explained earlier in sections 2.5.4, 2.5.5, and 2.5.6, respectively. Each feature was calculated for every node, and then the subjects were separated into patients and controls groups. A *t*-test was used to highlight nodes that are different between the groups. To control for the multiple testing error rate (90 *t* -tests), the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) was used with three different false discovery rates (FDR) for each network, which are 10%, 25%, and 50%.

Figure 20 together with Table 4 provide the results for the extent of activation network. Figure 21 together with Table 5 provide the results for the intensity of activation network. Similarly, Figure 22 and Table 6 provide the results for the traditional network. Note that Figures 19 through 21 were generated using BrainNet Viewer (Xia et al., 2013). It should be noted that no significant node difference was found at 10% FDR for either ICA-based networks or at any FDR rate for the modified traditional network; therefore, the corresponding figures and tables do not display such results. Figure 20 shows no overlapping significant nodes across the centrality measures within the same FDR rate. On the other hand, Figure 21 and Figure 22 both show several shared significant nodes across the centrality measures at the 50% FDR rate for the intensity of activation networks and at all FDR rates for the traditional networks.

Table 4: The significant nodes of the extent of activation network at different FDRs. Red indicates patients have a greater value, blue indicates controls have a greater value (not applicable in this case).

	<b>Degree</b> <b>Centrality</b>	<b>Betweenness Centrality</b>	<b>Eigenvector Centrality</b>
<b>25% FDR</b>			Right middle frontal gyrus, orbital part
			Right angular gyrus
<b>50% FDR</b>		Right fusiform gyrus	
			Right middle frontal gyrus, orbital part
			Right angular gyrus



Figure 20: Extent of activation local network features at different False Discovery Rates (FDR). Red nodes indicate patients have a greater value, blue nodes indicate controls have a greater value, and gray nodes indicate no significant difference between the patients and controls groups. The 10% FDR showed no significant nodes.



Figure 21: Intensity of activation local network features at different False Discovery Rates (FDR). Red nodes indicate patients have a greater value, blue nodes indicate controls have a greater value, and gray nodes indicate no significant difference between the patients and controls groups. The 10% FDR showed no significant nodes.



Figure 22: Traditional local network features at different False Discovery Rates (FDR). Red nodes indicate patients have a greater value, blue nodes indicate controls have a greater value, and gray nodes indicate no significant difference between the patients and controls groups.

At the 50% FDR rate for the intensity of activation network, the right inferior f gyrus (opercular and the triangularis parts) were significant in terms of both degre eigenvector centrality measures. Similarly, the right cuneus and the right precuneus both significant in terms of degree and eigenvector centrality measures. Figure 2 shows that the left anterior cingulate gyrus was significant in terms of all three cen measures.

Moreover, Figure 21 shows that except for the left lingual gyrus in the degree cen measures, all the left hemispheric significant nodes had greater values for pa whereas, all right hemispheric significant nodes had greater values for controls. C other hand, Figure 22 shows that the traditional network displays no structure. reorganizational pattern and the hub shifts are sporadic, especially at 50% FDR.

To examine if the local nodal differences or the hemispheric disparity, observed ICA-based networks, can be explained by atypical language laterality, the effect laterality index, as defined in equation (6), on the three centrality measures was test all fourteen ICA-based significant nodes. The results indicate that no effect was fou either of the networks at 10% FDR using the Benjamini-Hochberg method (Benj and Hochberg, 1995). In other word, atypical language laterality does not seem to caused local hub shifts. It should be noted however, that only six patients had language laterality as indicated in Table 3, whereas the rest of the patients and all co had typical left language laterality. Additionally, the epileptogenic focus wa hemispheric for 24 patients, right hemispheric for 4 patients, and bilateral for 1 patie Table 5: The significant nodes of the intensity of activation network at different FDRs. Red indicates patients have a greater value, blue indicates that controls have a greater value.



Table 6: The significant nodes of the traditional network at different FDRs. Red indicates patients have a greater value, blue indicates that controls have a greater value.










Figure 23: Clustering results for (a) intensity of activation network and (b) extent of activation network. The first three lines of each subfigure illustrate the results of the selforganizing maps for the: degree centrality, betweenness centrality, and eigenvector centrality feature vectors in order. The last line shows the final and second level k-means clustering results. The subjects on the left side are patients and the subjects on the right side are controls. A red square represents a subject clustered as a patient; whereas, a blue square represents a subject clustered as a control.

### **4.4 Clustering Results**

Several studies showed that epilepsy affects brain language networks (Campo et al., 2013; Vlooswijk et al., 2010; 2011; J. Wang et al., 2013a; You et al., 2011; 2013). In this dissertation three graph measures were used as feature vectors in a dual-level clustering scheme. Figure 23 part (a) and part (b) illustrate the results of the unsupervised clustering for the intensity of activation network and extent of activation network, respectively. These results show that the intensity of activation network produced better clustering outcomes compared to those produced from the extent of activation network. Table 7 enumerates the confusion matrices of both networks and confirms this observation. The table shows that the intensity of activation network clustering has 70% precision, 72.41% sensitivity, 70% specificity, and 71.19% accuracy. On the other hand, the extent of activation network clustering has 51.52% precision, 58.62% sensitivity, 46.67% specificity, and 52.54% accuracy.



Figure 24: The first two principal components of the different centrality measures. (a) The degree centrality of the extent of activation network. (b) The degree centrality of the intensity of activation network. (c) The betweenness centrality of the extent of activation network. (d) The betweenness centrality of the intensity of activation network. (e) The eigenvector centrality of the extent of activation network. (f) The eigenvector centrality of the intensity of activation network.

<b>Intensity of Activation</b>				<b>Extent of Activation</b>				
	<b>Clustered as</b>					<b>Clustered as</b>		
ctual		<b>Patients</b>	<b>Controls</b>	ಡ		<b>Patients</b>	<b>Controls</b>	
	<b>Patients</b>	21	8		믕	<b>Patients</b>	17	12
	<b>Controls</b>	Q	21		e.	<b>Controls</b>	16	14
	<b>Precision</b>	70.00%				<b>Precision</b>	51.52%	
	<b>Sensitivity</b>	72.41%				<b>Sensitivity</b>	58.62%	
	<b>Specificity</b>	70.00%				<b>Specificity</b>	46.67%	
	<b>Accuracy</b>	71.19%				<b>Accuracy</b>	52.54%	

Table 7: Clustering confusion matrix for the intensity of activation and extent of activation networks.

The first two principal components of each centrality measure were computed and plotted to justify the use of the dual-level clustering system. Figure 24 shows that the first two principal components capture between 16% and 18% of the total variance in the data; yet, none of the centrality measures in either of the ICA-based networks show separable clusters for the patients and controls groups. In other words, the subject groups cannot be easily separated by a linear clustering algorithm or even a nonlinear one. Nevertheless, the dual-level clustering scheme, which was used instead, was able to achieve decent precision values of 70% for the intensity of activation network and 51.52% for the extent of activation network.

### **CHAPTER 5**

#### Discussion

In this dissertation two novel data-driven network construction methods were introduced, and were compared to a traditional method and a modified version of it. These proposed methods were used to study the effects of epilepsy on the whole brain language network during task. The results showed that epilepsy is associated with network changes where the patients showed less efficient networks as compared to controls. These network changes did not manifest on the global level but rather on the topological and local nodal levels. Topologically, all networks were shown to have small-world network architecture with subtle differences between the subject groups. On the local level, the intensity of activation based networks showed ipsilateral nodal centrality reorganization. An unsupervised clustering system was able to objectively separate the patients and controls validating the correlation of epilepsy with the nodal hub reorganization.

The thresholding scheme used in this dissertation showed a difference between the patient and control groups in the ICA-based networks but not in the traditional and modified traditional networks. In the ICA-based networks patients had greater area under the curve for both construction methods. The greater area under the curve for the patients implies that they can achieve the language task with less dense graphs. In other words, the patients reach a connected graph with smaller densities than controls. Hence, the whole brain language network is less efficient in patients compared to controls. Controls tend to compartmentalize the brain into several smaller connected components as opposed to employing the whole brain to achieve a single language task. This decrease in efficiency agrees with other studies that demonstrated effects of epilepsy on memory networks (Campo et al., 2013; Vlooswijk et al., 2011; Voets et al., 2009), on resting state networks (Bettus et al., 2010; 2009; Doucet et al., 2012; Liao et al., 2010a; Mankinen et al., 2012; Morgan et al., 2010; Waites et al., 2006; Z. Zhang et al., 2010), and on language networks (Karunanayaka et al., 2011; Vlooswijk et al., 2010).

Karunanayaka et al., (2011) used a semantic/tone decision fMRI task (Binder et al., 1997) to study the different semantic networks, detected by ICA, and their correlation with the performed task. They concluded that epilepsy negatively affects the left hemispheric language network in patients with left focused temporal lobe epilepsy; furthermore, they showed that epilepsy also altered other nodes in the network in both left and right temporal lobe epilepsy. Likewise, Vlooswijk et al., 2010 used a covert word-generation and text reading paradigms to study the effect of epilepsy on language networks. They constructed their networks by connecting highly active regions identified by the modeldependent GLM analysis. Similar to this dissertation, they reported a decrease in functional connectivity in the language areas, and general reduction in language performance for patients.

Unlike these aforementioned studies, the findings presented here are based on whole brain functional networks constructed using an auditory decision descriptive task. The ICA-based networks were data-driven and proved to be useful at detecting brain dynamics while preserving its temporal characteristics throughout task performance.

In order to explain the difference that was observed between the patients and controls in the ICA-based networks, global network features were assessed for understanding the network dynamics and defining its topology; additionally, these global features were computed for the traditional and modified traditional methods for comparative purposes. For both ICA-based network methods, there was no difference between the subject groups except the average clustering coefficient of the extent of activation network. Even though the global features of both networks did not explain the patients/controls difference observed before, the two networks showed variances. These variances were especially obvious in the characteristic path length as the two networks demonstrated different path lengths. Similarly, the traditional and modified traditional methods did not show any significant difference on the global level. Consequently, the ICA-based networks express different information because they were constructed on different bases, namely the extent of activation and intensity of activation.

Moreover, for all network construction methods, the patient and control networks showed a small-world network topology. Nonetheless, the extent of activation network showed a lower small-world clustering parameter compared to controls indicating the tendency of the patients' network towards a more random network. On the other hand, the intensity of activation network did not show any difference between the subject groups. The smallworld network topology has higher information transfer efficiency and better synchronization when compared to random networks; hence, the extent of activation network showed that patients have reduced information transfer efficiency matching the inefficiency that was observed through thresholding. These findings match other studies that find typical and atypical brain networks have a small-world architecture and that neurological disorders introduce changes to the network while maintaining the overall small-world network architecture (Bernhardt et al., 2011; Hosseini and Kesler, 2013; Z. Zhang et al., 2011).

Regardless of the graph global features and topology, it is unrealistic to assume that neurons are optimally and intricately connected knowing the whole brain network topology and structure. It is fair, however, to assume that neurons act on a local information optimization to achieve a certain task as explained in (Fornito et al., 2013). Therefore, three centrality measures were studied to identify the local differences in language networks between children with epilepsy and normal controls.

A node centrality emphasizes its importance as a hub in the brain information highway. The different centrality measures used in this dissertation convey certain common, yet non-redundant information (Zuo et al., 2012). In this dissertation, the extent of activation networks did not reveal shared significant nodes between the centrality measures even at the 50% false discovery rate. On the other hand, the intensity of activation networks highlighted several shared significant nodes. Similarly, the traditional networks showed several shared nodes between the centrality measures across different FDR rates as opposed to the modified traditional method that showed no significant local difference even at high FDR rates.

In the intensity of activation network the shared significant nodes showed an interesting phenomenon where the left hemispheric nodes had greater centrality values for the patients while the right hemispheric nodes displayed greater centrality values for the controls. This observation might be attributed to the effect of pediatric epilepsy on the reorganization of the language network (You et al., 2013; 2011). On the other hand, the significant nodes exhibited by the traditional method seemed to be randomly organized with no regulated pattern. It was also shown that language laterality was not correlated with this nodal reorganization in the intensity of activation network; however, the small number of subjects with atypical right language laterality might have biased the observation.

To verify that the local centrality differences were associated with focal epilepsy, the three centrality measures were used as input vectors to an unsupervised clustering system. The data-driven clustering system was not able to correctly group the subjects into patients and controls for the extent of activation network as demonstrated in Figure 23. These results matched the lack of significant hub difference. On the other hand, the intensity of activation network yielded clusters with fair precision, which is close to other studies that reported precision values around 75% in: epilepsy (J. Zhang et al., 2012), attention-deficit hyperactivity disorder (Colby, 2012; Dai et al., 2012), autism (Anderson et al., 2011), Alzheimer's disease (Chen et al., 2011; J. Wang et al., 2013b), schizophrenia (Bassett et al., 2012), and unipolar or severe depression (Lord et al., 2012; Zeng et al., 2012).

It was shown that neither of the centrality measures could achieve acceptable clustering results alone; in addition, the first principal components of these centrality measures did not show any separable or grouped subject clusters. Therefore, the dual-level clustering system that was based on the SOM neural networks and the k-means algorithm was necessary to achieve the fair precision values noted above.

Consequently, as in previous studies (You et al., 2013; 2011), it can be concurred that pediatric epilepsy can cause, or is associated with, local hub shifts in the language network, similar to the shifts shown in the intensity of activation network.

### **CHAPTER 6**

### Conclusion and Future Work

In retrospect to the research aims of this dissertation, new functional brain network construction methods were introduced. The networks were then analyzed using graph theory to investigate the posed research questions. Pediatric epilepsy was shown to be associated with language network inefficiency. The source of this inefficiency was not a change on the global level of the network but rather on the local level. The local level changes were shown to be as ipsilateral hub reorganization. Furthermore, pediatric epilepsy did not affect the language networks topology and it was shown to have smallworld network characteristics similar to those of normal controls.

The data-driven network construction methods were novel in that they assumed a brain network with independent compartments communicating with each other in a network mode. The communication patterns, or edges, were based on different aspects associated with achieving a certain task. The identification of the communication patterns was performed without any a priori assumptions or models. Furthermore, the methods allowed network construction on the individual level. In other words, each individual subject had individual networks in accordance with the methods used. This individual granularity is clinically important because it can help in early detection of the disease or even in studying the effects of certain medications on the overall brain function for individual subjects.

Pediatric epilepsy is a progressive network disease. The main findings of the dissertation concurred that the disease induces local network shifts while preserving the global and topological features. This global feature preservation masks the local shifts and makes it harder to understand the causes of the disease; hence, graph theoretical analysis was proven to be a valuable framework for studying such latent changes.

Future investigations on this research endeavor should not be limited to language networks. For example, the methodology used to construct the networks can be utilized with memory networks or even resting state networks, which are perceived by the scientific community as major research initiatives. In addition, the procedure can be used to explore the effects of other neurological syndromes on different functional networks. For instance, the functional networks can be used to study the effects of Alzheimer's disease on the different memory networks, or the effects of autism on certain cognitive networks and functions.

The analyses performed in this dissertation can be also extended to study pediatric epilepsy as an attack on the functional network. For example, if the major network hubs in controls subjects can be removed, then the overall network performance can be compared to that of patients to determine any similarities or differences. This approach might help in identifying epileptogenic zones in focal epilepsy. Such network attacks can be also performed on structural networks to study the functional/structural network coupling. For instance, identifying key structural hubs then removing these hubs from the

functional networks can help in predicting the functional/structural network linkage in direct relation to the neurological disorder.

### **References**

- Anderson, J.S., Nielsen, J.A., Froehlich, A.L., Dubray, M.B., Druzgal, T.J., Cariello, A.N., Cooperrider, J.R., Zielinski, B.A., Ravichandran, C., Fletcher, P.T., Alexander, A.L., Bigler, E.D., Lange, N., Lainhart, J.E., 2011. Functional connectivity magnetic resonance imaging classification of autism. Brain 134, 3742–3754.
- Bassett, D.S., Nelson, B.G., Mueller, B.A., Camchong, J., Lim, K.O., 2012. Altered resting state complexity in schizophrenia. NeuroImage 59, 2196–2207.
- Beckmann, C., Smith, S., 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. Medical Imaging, IEEE Transactions on 23, 137–152.
- Beckmann, C.F., 2012. Modelling with independent components. NeuroImage 1–11.
- Bell AJ, Sejnowski TJ. An information-maximization approach to blind separation and blind deconvolution. Neural computation (1995) vol. 7 (6) pp. 1129-1159
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 289–300.
- Berl, M.M., Mayo, J., Parks, E.N., Rosenberger, L.R., Vanmeter, J., Ratner, N.B., Vaidya, C.J., Gaillard, W.D., 2012. Regional differences in the developmental trajectory of lateralization of the language network. Hum Brain Mapp.
- Bernhardt, B.C., Chen, Z., He, Y., Evans, A.C., Bernasconi, N., 2011. Graph-Theoretical Analysis Reveals Disrupted Small-World Organization of Cortical Thickness Correlation Networks in Temporal Lobe Epilepsy. Cerebral Cortex.
- Bettus, G., Bartolomei, F., Confort-Gouny, S., Guedj, E., Chauvel, P., Cozzone, P.J., Ranjeva, J.-P., Guye, M., 2010. Role of resting state functional connectivity MRI in presurgical investigation of mesial temporal lobe epilepsy. Journal of Neurology, Neurosurgery & Psychiatry 81, 1147–1154.
- Bettus, G., Guedj, E., Joyeux, F., Confort-Gouny, S., Soulier, E., Laguitton, V., Cozzone, P.J., Chauvel, P., Ranjeva, J.-P., Bartolomei, F., Guye, M., 2009. Decreased basal fMRI functional connectivity in epileptogenic networks and contralateral compensatory mechanisms. Hum Brain Mapp 30, 1580–1591.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging.

The Journal of neuroscience : the official journal of the Society for Neuroscience, 17(1), 353–362.

- Bollabas B. Random Graphs (Academic, London, 1985).
- Bonacich, P., 1987. Power and centrality: A family of measures. American journal of sociology 1170–1182.
- Bonilha, L., Nesland, T., Martz, G.U., Joseph, J.E., Spampinato, M.V., Edwards, J.C., Tabesh, A., 2012. Medial temporal lobe epilepsy is associated with neuronal fibre loss and paradoxical increase in structural connectivity of limbic structures. Journal of Neurology, Neurosurgery & Psychiatry 83, 903–909.
- Brandes, U., 2001. A faster algorithm for betweenness centrality\*. The Journal of Mathematical Sociology 25, 163–177.
- Brodmann K. Vergleichende Lokalisationslehre der Grosshirnrinde. Leipzig : Johann Ambrosius Bart, (1909)
- Calhoun, V., & Adali, T. (2012). Multi-subject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery. IEEE REVIEWS IN BIOMEDICAL ENGINEERING.
- Campo, P., Garrido, M.I., Moran, R.J., García-Morales, I., Poch, C., Toledano, R., Gil-Nagel, A., Dolan, R.J., Friston, K.J., 2013. Network reconfiguration and working memory impairment in mesial temporal lobe epilepsy. NeuroImage 72C, 48–54.
- Cardoso JF. Infomax and maximum likelihood for blind source separation. Signal Processing Letters, IEEE (1997) vol. 4 (4) pp. 112-114
- Chartrand, G., 1985. Introductory graph theory.
- Chen, G., Ward, B., Xie, C., Li, W., Wu, Z., Jones, J., Franczak, M., Antuono, P., Li, S.- J., 2011. Classification of Alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional MR imaging. Radiology 259, 213–221.
- Colby, J.B., 2012. Insights into multimodal imaging classification of ADHD 1–18.
- Collins, D.L., Zijdenbos, A.P., Kollokian, V., Sled, J.G., Kabani, N.J., Holmes, C.J., Evans, A.C., 1998. Design and construction of a realistic digital brain phantom. IEEE Transactions on Medical Imaging, vol.17, no.3, pp.463,468. Doi: 10.1109/42.712135

Comon, P., 1994. Independent component analysis, a new concept? Signal processing.

- Dai, D., Wang, J., Hua, J., He, H., 2012. Classification of ADHD children through multimodal magnetic resonance imaging. Front. Syst. Neurosci. 6, 1–8.
- Doucet, G., Osipowicz, K., Sharan, A., Sperling, M.R., Tracy, J.I., 2012. Extratemporal functional connectivity impairments at rest are related to memory performance in mesial temporal epilepsy. Hum Brain Mapp 34, 2202–2216.
- Erdős, P., Rényi, A., 1959. On random graphs. Publicationes Mathematicae Debrecen, Vol. 6 (1959), pp. 290-297 6, 290–297.
- Essen, D.C.V., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., Consortium, F.T.W.-M.H., 2013. The WU-Minn Human Connectome Project: An overview. NeuroImage 1–18.
- Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL and the Brain Development Cooperative Group. Unbiased average age-appropriate atlases for pediatric studies. Neuroimage (2011) vol. 54 (1) pp. 313-27
- Fornito, A., Zalesky, A., Breakspear, M., 2013. Graph analysis of the human connectome: Promise, progress, and pitfalls. NeuroImage 1–19.
- Friston KJ, Li B, Daunizeau J, Stephan KE. Network discovery with DCM. Neuroimage (2011) vol. 56 (3) pp. 1202-1221
- Gaillard, W.D., 2004. Functional MR imaging of language, memory, and sensorimotor cortex. Neuroimaging Clin N Am 14, 471–485.
- Gaillard, W.D., Berl, M.M., Moore, E.N., Ritzl, E.K., Rosenberger, L.R., Weinstein, S.L., Conry, J.A., Pearl, P.L., Ritter, F.F., Sato, S., Vezina, L.G., Vaidya, C.J., Wiggs, E., Fratalli, C., Risse, G., Ratner, N.B., Gioia, G., Theodore, W.H., 2007. Atypical language in lesional and nonlesional complex partial epilepsy. Neurology 69, 1761– 1771.
- Gentzsch, W., 2001. Sun Grid Engine: towards creating a compute power grid. First IEEE/ACM International Symposium on Cluster Computing and the Grid. Doi: 10.1109/CCGRID.2001.923173
- Greicius, M.D., Srivastava, G., Reiss, A.L., Menon, V., 2004. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc Natl Acad Sci USA 101, 4637–4642.
- Hagberg, A., Swart, P., S Chult, D., 2008. Exploring network structure, dynamics, and function using NetworkX.
- Hosseini, S.M.H., Kesler, S.R., 2013. Comparing connectivity pattern and small-world organization between structural correlation and resting-state networks in healthy adults. NeuroImage 78C, 402–414.
- Hyvärinen A. Fast and robust fixed-point algorithms for independent component analysis. Neural Networks (1999)
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage 17, 825–841.
- Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. Fsl. NeuroImage 62, 782–790.
- Jezzard P, Mathews PM, Smith, SM. Functional MRI: An Introduction to Methods (2002)
- Jutten C and Herault J, "Blind separation of sources—Part I: An adaptive algorithm based on neuromimetic architecture," Signal Processing, vol. 24, pp. 1–10, 1991.
- Karunanayaka, P., Kim, K.K., Holland, S.K., Szaflarski, J.P., 2011. The effects of left or right hemispheric epilepsy on language networks investigated with semantic decision fMRI task and independent component analysis. Epilepsy & Behavior 20, 623–632.
- Kim KK, Karunanayaka P, Privitera MD, Holland SK, Szaflarski JP. Semantic association investigated with functional MRI and independent component analysis. Epilepsy & Behavior (2011) vol. 20 (4) pp. 613-622
- Lahlou, M., Guillen, M.R., Adjouadi, M., Gaillard, W., 2006. Mediar: An online webbased repository site of fmri medical images and clinical data for childhood epilepsy 1–8.
- Liao, W., Zhang, Z., Pan, Z., Mantini, D., Ding, J., Duan, X., Luo, C., Lu, G., Chen, H., 2010a. Altered Functional Connectivity and Small-World in Mesial Temporal Lobe Epilepsy. PLoS ONE 5, e8525.
- Liao, W., Zhang, Z., Pan, Z., Mantini, D., Ding, J., Duan, X., Luo, C., Wang, Z., Tan, Q., Lu, G., Chen, H., 2010b. Default mode network abnormalities in mesial temporal lobe epilepsy: A study combining fMRI and DTI. Hum Brain Mapp 32, 883–895.
- Lord, A., Horn, D., Breakspear, M., Walter, M., 2012. Changes in Community Structure of Resting State Functional Connectivity in Unipolar Depression. PLoS ONE 7, e41282.
- Mankinen, K., Jalovaara, P., Paakki, J.-J., Harila, M., Rytky, S., Tervonen, O., Nikkinen, J., Starck, T., Remes, J., Rantala, H., Kiviniemi, V., 2012. Connectivity disruptions in resting-state functional brain networks in children with temporal lobe epilepsy. Epilepsy Research 100, 168–178.
- McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ. Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Mapp (1998) vol. 6 (3) pp. 160-88
- Milham, D.A.F.A.M.P., 2013. Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data 1–31.
- Morgan, V., Gore, J., Abou-Khalil, B., 2010. Functional epileptic network in left mesial temporal lobe epilepsy detected using resting fMRI. Epilepsy Research 88, 168–178.
- Penny WD, Friston KJ, Ashburner JT, Kiebel SJ, Nichols TE. Statistical Parametric Mapping: The Analysis of Functional Brain Images (2006)
- Pham DT, Garrat P, Jutten C. "Separation of a mixture of independent sources through a maximum likelihood approach," in Proc. EUSIPCO, (1992), pp. 771–774.
- Reus, M.A. de, van den Heuvel, M.P., 2013. Estimating false positives and negatives in brain networks. NeuroImage 70, 402–409.
- Roberts S and Everson R. (2001). Independent component analysis: Principles and practice. Cambridge: Cambridge University Press.
- Rubinov M and Sporns O. Complex network measures of brain connectivity: Uses and interpretations. Neuroimage (2010) vol. 52 (3) pp. 1059-1069
- Smith, S., Jenkinson, M., Woolrich, M., Beckmann, C., Behrens, T., Johansen-Berg, H., Bannister, P., De Luca, M., Drobnjak, I., Flitney, D., Niazy, R., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J., Matthews, P., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23, S208–S219.
- Smith, S.M., 2002. Fast robust automated brain extraction. Hum Brain Mapp 17, 143– 155.
- Sonty, S.P., Mesulam, M.M., Weintraub, S., Johnson, N.A., Parrish, T.B., Gitelman, D.R., 2007. Altered Effective Connectivity within the Language Network in Primary Progressive Aphasia. J Neurosci 27, 1334–1345.

Sporns, O., (2011a). Networks of the brain. Cambridge, Mass: MIT Press.

- Sporns, O., (2011b). The human connectome: a complex network. Ann N Y Acad Sci 1224, 109–125.
- Sporns, O., 2013. The human connectome: Origins and challenges. NeuroImage, Vol 80, pp 53-61.
- Sun, B., Berl, M. M., Burns, T. G., Gaillard, W. D., Hayes, L., Adjouadi, M., Jones, R., A., "Age association of language task induced deactivation induced in a pediatric population", NeuroImage, Vol. 65, pp. 23-33.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15, 273–289.
- Varoquaux, G., Craddock, R.C., 2013. Learning and comparing functional connectomes across subjects. NeuroImage 1–11.
- Vlooswijk, M.C.G., Jansen, J.F.A., Majoie, H.J.M., Hofman, P.A.M., de Krom, M.C.T.F.M., Aldenkamp, A.P., Backes, W.H., 2010. Functional connectivity and language impairment in cryptogenic localization-related epilepsy. Neurology 75, 395–402.
- Vlooswijk, M.C.G., Vaessen, M.J., Jansen, J.F.A., de Krom, M.C.F.T.M., Majoie, H.J.M., Hofman, P.A.M., Aldenkamp, A.P., Backes, W.H., 2011. Loss of network efficiency associated with cognitive decline in chronic epilepsy. Neurology 77, 938– 944.
- Voets, N.L., Adcock, J.E., Stacey, R., Hart, Y., Carpenter, K., Matthews, P.M., Beckmann, C.F., 2009. Functional and structural changes in the memory network associated with left temporal lobe epilepsy. Hum Brain Mapp 30, 4070–4081.
- Waites, A.B., Briellmann, R.S., Saling, M.M., Abbott, D.F., Jackson, G.D., 2006. Functional connectivity networks are disrupted in left temporal lobe epilepsy. Ann Neurol 59, 335–343.
- Wang, J., You, X., Wu, W., Guillen, M.R., Cabrerizo, M., Sullivan, J., Donner, E., Bjornson, B., Gaillard, W.D., Adjouadi, M., 2013a. Classification of fMRI patterns-A study of the language network segregation in pediatric localization related epilepsy. Hum Brain Mapp.
- Wang, J., Zuo, X., Dai, Z., Xia, M., Zhao, Z., Zhao, X., Jia, J., Han, Y., He, Y., 2013b. Disrupted Functional Brain Connectome in Individuals at Risk for Alzheimer's Disease. Biological Psychiatry 73, 472–481.
- Wang, Z., Zhang, Z., Jiao, Q., Liao, W., Chen, G., Sun, K., Shen, L., Wang, M., Li, K., Liu, Y., Lu, G., 2012. Impairments of thalamic nuclei in idiopathic generalized epilepsy revealed by a study combining morphological and functional connectivity MRI. PLoS ONE 7(7): e39701. doi:10.1371/journal.pone.0039701
- Watts, D., Strogatz, S., 1998. Collective dynamics of "small-world" networks. Nature 393, 440–442.
- Wilke, M., Lidzba, K., 2007. LI-tool: A new toolbox to assess lateralization in functional MR-data. J Neurosci Methods 163, 128–136.
- Wilke, M., Schmithorst, V.J., 2006. A combined bootstrap/histogram analysis approach for computing a lateralization index from neuroimaging data. NeuroImage 33, 522– 530.
- Worsley K, Friston K. Analysis of fMRI time-series revisited--again. Neuroimage (1995) vol. 2 (3) pp. 173-81
- Wu, K., Taki, Y., Sato, K., Hashizume, H., Sassa, Y., Takeuchi, H., Thyreau, B., He, Y., Evans, A.C., Li, X., Kawashima, R., Fukuda, H., 2013. Topological Organization of Functional Brain Networks in Healthy Children: Differences in Relation to Age, Sex, and Intelligence. PLoS ONE 8, e55347.
- Wu, T., Wang, L., Chen, Y., Zhao, C., Li, K., Chan, P., 2009. Changes of functional connectivity of the motor network in the resting state in Parkinson's disease. Neuroscience Letters 460, 6–10.
- Xia, M., Wang, J., He, Y., 2013. BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics. PLoS ONE 8, e68910.
- You, X., Adjouadi, M., Guillen, M.R., Ayala, M., Barreto, A., Rishe, N., Sullivan, J., Dlugos, D., Vanmeter, J., Morris, D., Donner, E., Bjornson, B., Smith, M.L., Bernal, B., Berl, M., Gaillard, W.D., 2011. Sub-patterns of language network reorganization in pediatric localization related epilepsy: a multisite study. Hum Brain Mapp 32, 784– 799.
- You, X., Adjouadi, M., Wang, J., Guillen, M.R., Bernal, B., Sullivan, J., Donner, E., Bjornson, B., Berl, M., Gaillard, W.D., 2013. A decisional space for fMRI pattern separation using the principal component analysis-a comparative study of language networks in pediatric epilepsy. Hum Brain Mapp 34, 2330–2342.
- Yu-Feng, Z., Yong, H., Chao-Zhe, Z., Qing-Jiu, C., Man-Qiu, S., Meng, L., Li-Xia, T., Tian-Zi, J., Yu-Feng, W., 2007. Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. Brain and Development 29, 83–91.
- Zeng, L.-L., Shen, H., Liu, L., Hu, D., 2013. Unsupervised classification of major depression using functional connectivity MRI. Hum Brain Mapp n/a–n/a.
- Zeng, L.-L., Shen, H., Liu, L., Wang, L., Li, B., Fang, P., Zhou, Z., Li, Y., Hu, D., 2012. Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis. Brain 135, 1498–1507.
- Zhang, J., Cheng, W., Wang, Z., Zhang, Z., Lu, W., Lu, G., Feng, J., 2012. Pattern Classification of Large-Scale Functional Brain Networks: Identification of Informative Neuroimaging Markers for Epilepsy. PLoS ONE 7, e36733.
- Zhang, Z., Liao, W., Chen, H., Mantini, D., Ding, J.-R., Xu, Q., Wang, Z., Yuan, C., Chen, G., Jiao, Q., Lu, G., 2011. Altered functional-structural coupling of large-scale brain networks in idiopathic generalized epilepsy. Brain 134, 2912–2928.
- Zhang, Z., Lu, G., Zhong, Y., Tan, Q., Liao, W., Wang, Z., Wang, Z., Li, K., Chen, H., Liu, Y., 2010. Altered spontaneous neuronal activity of the default-mode network in mesial temporal lobe epilepsy. Brain Research 1323, 152–160.
- Zhang, Z., Lu, G., Zhong, Y., Tan, Q., Yang, Z., Liao, W., Chen, Z., Shi, J., Liu, Y., 2009. Impaired attention network in temporal lobe epilepsy: A resting FMRI study. Neuroscience Letters 458, 97–101.
- Zuo, X.-N., Ehmke, R., Mennes, M., Imperati, D., Castellanos, F.X., Sporns, O., Milham, M.P., 2012. Network centrality in the human functional connectome. Cereb Cortex 22, 1862–1875.

## **Appendices**

## Appendix I



Table 8: Detailed demographics information for patients and controls.



BCCH: British Columbia's Children's Hospital

CNMC: Children's National Medical Center

CHOA: Children's Healthcare of Atlanta

# Appendix II

Table 9: Detailed seizure information









# Appendix III





## **VITA**

## ANAS SALAH EDDIN



## HONORS AND AWARDS

- 1. Best Paper Finalist (2013) The  $6<sup>th</sup>$  International IEEE EMBS Neural Engineering Conference.
- 2.  $1<sup>st</sup>$  place poster winner (2011) Yearly Scholarly Forum, Biomedical Studies Category, Florida International University.
- 3. Outstanding Graduate Award (2009) Florida International University, College of Engineering and Computing.
- 4. The Fulbright Scholarship (2007-2009) –United States Department of State, Bureau of Educational and Cultural Affairs.

## PUBLICATIONS

- 1. Salah Eddin, A., Wang, J., Wensong, W., Sargolzaei, S., Bjornson, B., Jones, R.A., Gaillard, W.D., Adjouadi, M., 2013. The Effects of Pediatric Epilepsy on a Language Connectome. Submitted, November 2013.
- 2. Sargolzaei, S., Cabrerizo, M., Goryawala, M., Salah Eddin, A., Adjouadi, M., 2013. Scalp EEG based Epileptic Classification using Graph Theory Applied to Functional Connectivity Network. Submitted, November 2013.
- 3. Wang, J., Salah Eddin, A., Wensong, W., Gaillard, W.D., Adjouadi, M., 2013. The effects of seizure focus on language network reorganization in pediatric temporal lobe epilepsy. Submitted, October 2013.
- 4. Salah Eddin, A., Wang, J., Sargolzaei, S., Gaillard, W.D., Adjouadi, M., 2013. ICAbased connectivity on Brain Networks using fMRI. In Proceedings of the  $6<sup>th</sup>$

International IEEE EMBS Conference on Neural Engineering, pp. 391-394, November 6-8, 2013, San Diego, CA.

- 5. Sargolzaei, S., Salah Eddin, A., Cabrerizo, M., Adjouadi, M., 2013. Resting State Functional Connectivity Based on Principal Component Transformation of Cortical fMRI Measurements. In Proceedings of the  $6<sup>th</sup>$  International IEEE EMBS Conference on Neural Engineering, pp. 1501-1504, November 6-8, 2013, San Diego, CA.
- 6. Wang, J., Salah Eddin, A., Zhou, Q., Gaillard, W.D., Adjouadi, M., 2013. Discriminating regional functional networks in pediatric epilepsy. In Proceedings of the  $6<sup>th</sup>$  International IEEE EMBS Conference on Neural Engineering, pp. 1513-1516, November 6-8, 2013, San Diego, CA.
- 7. Salah Eddin, A., Sargolzaei, S., Wang, J., Gaillard, W.D., Adjouadi, M., 2013. Smallworld connectivity in FMRI based functional language networks using ICA in pediatric epilepsy. Proceedings of the 19<sup>th</sup> Annual Meeting of the Organization for Human Brain Mapping. pp. 76, June 16, 2013. Seattle, WA.
- 8. Delgado, J., Salah Eddin, A., Adjouadi, M., Sadjadi, S.M., 2011. Paravirtualization for Scientific Computing: Performance Analysis and Prediction". In Proceedings of the 2011 IEEE International Conference on High Performance Computing and Communications (HPCC '11), pp. 536-543, Banff, AB, Canada, 2011.

## PRESENTATIONS

- 1. Tarquinio, D., Salah Eddin, A., Diaz, M., Stewart, T. 2012. Retrospective subdural EEG source localization of the ictal onset zone in pediatric epilepsy surgery". Proceedings of the 2012 Annual Meeting of the American Epilepsy Society. pp. 63, December 3, 2012. San Diego, CA.
- 2. Salah Eddin, A., Wang, J., Adjouadi. M., 2012. Rest state fMRI extraction from taskdriven data using independent component analysis. Presented at the 2012 SACNAS National Conference. October 11-14, 2012. Seattle, WA.
- 3. Delgado, J., Salah Eddin, A., Sadjadi, S.M., Adjouadi. M., 2012. Deadline-Driven Medical Job Scheduling on Virtualized Systems. Presented at the 2012 SACNAS National Conference. October 11-14, 2012. Seattle, WA.
- 4. Salah Eddin, A., Adjouadi, M., 2011. Improvement of Automatic Brain Masking by Machine Learning. Presented at the 2011 Latin American Grid Summit. Nov. 4, 2011. Boca Raton, FL.
- 5. Salah Eddin, A., Adjouadi, M., 2010. An Assistive Technology Tool for Text Entry Based on N-gram Statistical Language Modeling. Presented at CAHSI, Fourth Annual Poster Session. April 5, 2010, Redmond, WA.