KU LEUVEN

An Information Theoretical Approach to EEG Source-Reconstructed Connectivity

Axel Faes

Thesis submitted for the degree of Master of Science in Artificial Intelligence, option Engineering and Computer Science

Thesis supervisor:
Prof. dr. ir. Marc Van Hulle

Assessor:
Mansoureh Fahimi
Prof. dr. Daniele Marinazzo

Mentor:
Mansoureh Fahimi

Academic year 2017 – 2018

Acknowledgements

For me, this thesis was an adventure through different unknown lands. Luckily, I had a guide in the form of my mentor and my supervisor. I was able to become familiar with an extraordinarily interesting field of research and I am excited to be able to continue my research within this field.

First and foremost, I would like to thank my mentor, Mansoureh Fahimi, and my supervisor, prof. Marc Van Hulle. I would also like to thank prof. dr. Daniele Marinazzo for his valuable advice during my research. Finally, an additional thanks goes to the assessors for reading the text.

Axel Faes

Contents

Abstract

Determining how distinct brain regions are connected and communicate with each other will shed light on how behaviour emerges. In EEG studies, interpreting connectivity measures can be problematic, due to the high correlation between signals recorded from the scalp surface, a result of the volume conductance of the scalp and skin. Therefore, meaningful connectivity patterns can be measured only from the from the spatiotemporal distribution of localised cortical sources, generally referred to as source reconstruction. Still, spurious connectivity issues may persist in source reconstructed EEG data, rendering it vital to choose an appropriate measure of connectivity. In this work, an information theoretical approach, which concerns model-free, probability based methods such as mutual information, conditional mutual information and interaction information, is taken. An information theoretical framework for Python is developed in order to operate on source reconstructed activity. This framework is used to perform a connectivity analysis of a high density source reconstructed EEG dataset, which was constructed in an experiment regarding semantic processing of abstract and concrete words.

List of Figures

Chapter 1

Introduction

Within the field of computational neuroscience, extensive research is being done on the connectivity of brain regions. Understanding how distinct brain regions are connected can answer many different questions. It will shed light on how behaviour emerges.

There are multiple ways to study brain connectivity. Brain connectivity can be studied on different levels, such as microscale or macroscale. Another question is what kind of connectivity is being measured. There is a distinction between the structural (or anatomical) connectivity and the functional connectivity. At the macroscale, functional connectivity can be measured with multiple imaging techniques. One popular imaging technique is EEG.

In EEG studies, interpreting connectivity measures can be problematic, due to the high correlation between signals recorded from the scalp surface, a result of the volume conductance of the scalp and skin. Therefore, meaningful connectivity patterns can be measured only from the from the spatiotemporal distribution of localised cortical sources, generally referred to as source reconstruction. Still, spurious connectivity issues may persist in source reconstructed EEG data, rendering it vital to choose an appropriate measure of connectivity.

This thesis takes an information theoretical approach, which concerns model-free, probability based methods such as Conditional Mutual Information and Directed Information. These methods are applied on real data, provided by the lab of computational neuroscience. This dataset contains source-reconstructed EEG data, recorded in an experiment revolving around semantic processing.

The experiment focused on two word categories: abstract words and concrete words. The data contains several brain regions of interest. The semantic processing of abstract and concrete words is slightly different. Some of the semantic processing is done in separate brain regions. There are regions on the pathway of word-processing that are involved in the processing of both concrete and abstract words. In this thesis, they will be referred to as the common regions.

1.1 Research Questions

From the motivation, there are two big aspects that this thesis deals with. There is information theory and source-reconstructed EEG Data. The novelty of this thesis lies in the usage of these information theoretical algorithms for source-reconstructed EEG data.

- How can information theoretical measures be applied to source-reconstructed EEG Data?
- Which observations can be made by applying these information theoretical measures on a high density EEG dataset provided by the lab of computational neuroscience?

1.2 Approach

The goal of this thesis, as well as it's main contribution, is to apply information theoretical measures on a high density EEG dataset provided by the lab of computational neuroscience. There are two steps to accomplish this. First, an implementation has to be made that can perform an information theoretical analysis of source-reconstruted EEG Data. Secondly, the information theoretical measures are applied to a dataset. Important in this step is to decide how to analyse the data.

1. Study of the relevant state-of-the-art literature and theoretical background. This includes source reconstruction, brain connectivity and information theory.
2. Development of framework for information theoretical analysis of source-reconstruted EEG Data.
3. Application of information theoretical measures to high density EEG dataset provided by the lab of computational neuroscience.

1.3 Results

The main novelty lies in the application of information theoretical algorithms for source-reconstructed EEG data. Behind this novelty, there are several different deliverables that made it possible to reach the goal of this thesis.

The implementation of the information theoretical algorithms is a first result of this thesis. The implementation is the backbone behind the thesis. The implementation consists out of three parts. There is the data conversion, which converts the source-reconstructed EEG data into multiple format. Secondly, methods for calculating the correct amount of bins to discretize the continious data have been implemented.

Finally, the actual information theoretical equations have been implemented. The implementation is developed to be of general use, meaning that it isn't too difficult to convert the implementation into an open source connectivity package.

A second result of this thesis is the actual analysis of the high density EEG dataset provided by the lab of computational neuroscience. The main conclusion made in the analysis of the EEG dataset, is that, within the common region, the actual semantic processing is quite different between the semantic processing of abstract words and the semantic processing of concrete words.

1.4 Structure of the Thesis

Chapter 2 discusses source reconstruction. It starts by explaining what source reconstruction entails and why it is important. Afterwards, an explanation is given about the reasons that spurious connectivity issues may still persist in source reconstructed EEG data.

Chapter 3 gives a background of brain connectivity and volume conduction. this chapter hightlights the different kinds of connectivity and briefly discusses them.

Chapter 4 provides a background about information theory. This chapter starts by explaining the concept of information entropy. Using this concept, the different aspects from information theory are discussed, including joint entropy, mutual information, multivariate methods and directed information. Finally, this chapter details methods on how to use information theory on continuous data.

Chapter 5 details the experiment that generated the data which this thesis analyses. This chapter starts by detailing the reasons behind the experiment and then details the experiment itself.

Chapter 6 will go in depth into the actual analysis that has been done on the data. The binning and the different analysis experiments are described.

Chapter 7 goes in depth into the source code that has been developed for the analysis. There are several important aspects to be discussed about the source code, such as the programming language and the implementation of the equations.

Chapter 8 details several interesting directions we can take this research. Considering the conclusion of this thesis, that information theoretical algorithms work well on source-reconstructed EEG data, many new paths open for further investigation. Finally, chapter 9 concludes the thesis.

Chapter 2

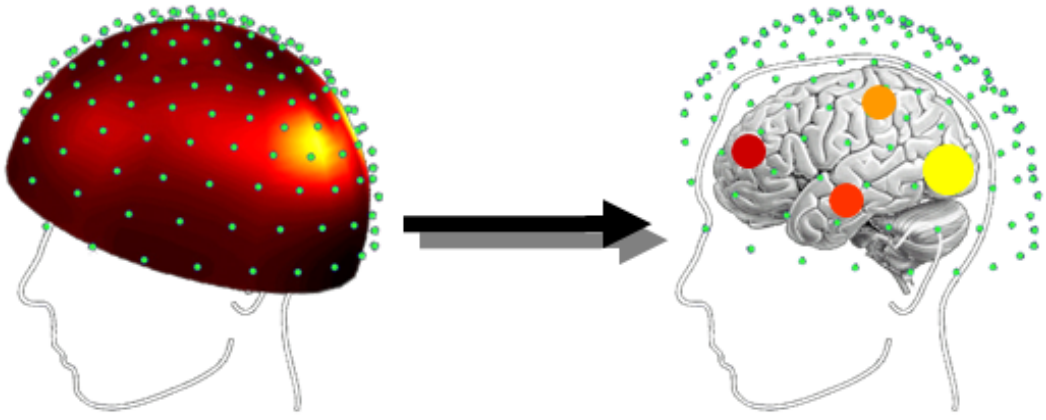
Source Reconstruction

Source reconstruction refers to the localisation of the electrical activity of the brain. In the case of this thesis, we are referring to EEG source localization, which is also called the reverse problem. Electrical activity within the brain is measured on the scalp using EEG electrodes. This is the forward problem. The reverse problem maps the measurements from the EEG electrodes back into the brain. [31]

2.1 Reverse Problem

Figure 2.1 visualizes the reverse problem. This is an ill-posed problem. EEG electrodes measure the scalp, in effect measuring a 2D grid. However, the brain is a 3D object. This means that information is inherently lost when EEG is used. One of the consequences is there are many different solutions to the reverse problem.

Figure 2.1: EEG Source Localisation [30].



Equation 2.1 represents source localisation. X represents the scalp recorded EEG activity. S represents the electrical sources within the brain, a current density vector. L represents

the head volume conductor model. The reverse problem is about finding S, this is represented in equation 2.2.

$$X = LS + n \tag{2.1}$$

$$O(S) = min||X - LS||^2 \tag{2.2}$$

2.2 Head Volume Conductor Model

Several different head volume conductor models can be used. The two most popular are the simple head models and the realistic head models. Simple head models model the brain as a single sphere with a couple layers. This model also assumes a uniform medium within the brain. Using a simple head model is fast and simple. However, it is not accurate.

The realistic head model is an accurate model, but it is computationally much more expensive. Different techniques are used to construct a head model such as finite element or finite boundary techniques.

2.3 Dipoles

The electrical activity from neural populations can be represented as a dipole. The inverse problem is essentially about determining the location, orientation, amplitude and number of these dipoles

There are practical limits to the number of dipoles that can be used. One of the most important limits is the spatial resolution of EEG. EEG can only measure electrical activity with a certain accuracy. This is due to the fact that the electrical activity is measured on the scalp, and each electrode records a linear combination of the underlying neural sources. This happens due to the low conductivity of the layers that are separating the cortex from the surface of the scalp (such as the skull, skin, and cerebrospinal fluid). The phenomena of how different mixtures of sources reach the scalp, how they attenuate and spread through the different layers of the head is referred to as volume conduction. We will elaborate on this further in the thesis.

2.4 Algorithms

There are different kind of algorithms used to solve the inverse problem. There is dipole fitting, which involves only a small number of dipoles. The locations, orientations, and magnitudes of these dipoles needs to be calculated. For some experiments, this method can be good enough, since a single dipole can account for 80% of all electrical activity [8].

Dipole fitting can be performed by a number of different software programs or matlab toolboxes, including BESA, eeglab and fieldtrip [19, 10, 27].

Then there are nonadaptive distributed-source imaging methods and adaptive distributed-source imaging methods. The idea of distributed-source imaging is that thousands of dipoles are placed within the brain on fixed locations with fixed orientations. This leaves only the magnitude to be computed. This is done by assigning electrode weights to each dipole.

Nonadaptive methods compute these electrode weights based on the electrode locations. This means that the weights are fixed over time and frequency. sLORETA is a commonly used nonadaptive inverse-source imaging technique [28]. The provided data used in this thesis has also been source-reconstructed using sLORETA. Other nonadaptive methods include LORETA and minimum-norm estimator [29, 17].

Nonadaptive methods are relatively fast to compute, are applicable to single time points and their result looks like FMRI activation maps. However, there are also some disadvantages. One issue is the number of comparisons that need to be computed in statistical analyses. In the case of 10,000 dipoles over time and frequency with two experiment conditions, there over more than 100 million possible statistical comparisons that need to be computed. [8]

Adaptive distributed-source imaging have a different way of computing the electrode weights. The recorded data is also used to compute the weights. The weights are not fixed over time and frequency. The accuracy of adaptive methods is often quite high, but they are much more complicated with many more parameters to be set. Beamforming is a class of algorithms that is most commonly used adaptive distributed-source imaging method [16, 35].

2.5 Practical Limits

Within a simulated environment, high spatial localization accuracy can be obtained. However, in practive, there are many issues. There are always uncertainties regarding electrode positions, brain anatomy, head movement and scalp conductivity. This means that spatial accuracy is often a few centimeters in size. The smallest voxels which are created by source reconstruction are typically 5-10 mm^3 in size. [15]

Without knowledge of the inverse problem, it might seem that source reconstruction is not a big deal. It might even seem to only bring advantages, since you can work within the actual cortical areas. However, source reconstruction comes with its fair share of problems and inaccuracies.

Within the context of this thesis, knowledge of the inverse problem is essential. While the data was already source-reconstructed, for the reasons above, it is important to know the inverse problem.

Chapter 3

Brain Connectivity

The human brain is anatomically a conglomeration of different brain regions. Functionally, the human brain is a giant network of specialised units connected by dynamically configurable communication pathways. [12]

There are several reasons why scientists would want to measure the way the brain is connected. Understanding brain connectivity aids our knowledge of the working of the brain. It allows us to better understand sensorimotor and cognitive tasks that are performed. This can lead to improved diagnosing of various diseases such as aphasia. [20]

A connectivity analysis refers to any analysis where multiple signals are utilised at the same time. The signal can represents several different sources. The signals could be the recordings from different electrodes or, in the case of this thesis, the source-reconstructed EEG data.

3.1 Connectivity Measures

Different connectivity measures can highlight different aspects about the connectivity. This means that there is not one superior connectivity method. The different connectivity measures are: [13, 8]

- Phase-Based Connectivity
- Power-Based Connectivity
- Cross-Frequency Coupling
- Graph Theory
- Granger Causality
- Information Theory

3.1.1 Phase-based Connectivity

Phase-based connectivity analyses utilise the phase differences between different signals. This is a very popular connectivity measure and this is partly because it has a strong neurophysiological interpretation. The timing of neural populations, as measured through phase, become synchronized. This method is computationally fast, does not make many assumptions on the data and is insensitive to time lag. However, it relies on precise temporal relationships.

3.1.2 Power-based Connectivity

Power-Based Connectivity analyses utlise time-frequency power correlations. These correlations are computed accross time or over trials. Power-Based Connectivity analyses are fairly resistant against temporal noise.

3.1.3 Cross-Frequency Coupling

Cross-Frequency Coupling is a statistical relationship between signals in different frequency bands. When utilised on electrodes, it can infer local connectivity when measured at a single electrode. With multiple electrodes, longer-range connectivity can be inferred. One advantage is that findings can be linked across species. Cross-frequency coupling is also very strong at identifying task-related high-frequency power. Computationally speaking, there are some disadvantages to cross-frequency coupling. The search space is huge, which makes it computationally very expensive.

3.1.4 Graph Theory

Graph Theory is a mathematical framework for studying graphs. A graph, containing vertices and edges, can be used to model networks. These networks can be models for brain connectivity. When electrode signals are used, each node can represent an electrode and connectivity is represented by edges. Graph-theoretical analyses are often easy to interpret. Graph theory presents a generic framework to analyse networks, making it easy to apply the same analyses to different kinds of data. The main task becomes the construction of a network. Graph theory is relatively new in the context of computational neuroscience. This makes graph theory an interesting research direction, but this is also a disadvantage. Since graph theory within computational neuroscience isn't as well established, it can be difficult to compare findings to different studies.

3.1.5 Granger Causality

Granger Causality is a statistical method that utilises signal variance. It tests whether variance from one signal can predict variance in another signal at a certain point in time. One big advantage to granger causality is that it can find directional connectivity. It can ignore simultaneous connectivity. This makes it less susceptible to volume conduction. Granger causality is a computationally expensive method to perform. [4]

3.1.6 Information Theory

Information theory is the main focus of this thesis. It is a simple, yet flexible and robust method for computing connectivity. One useful aspect from information theory is mutual information. Mutual information computes shared information between two variables. This computation is based on the (joint) distributions of values within variables. Information theory has several advantages. It can detect multiple kinds of relationships between different signals. It can detect both linear and non-linear relationships. Information theory also contains many different constructs which are useful for a connectivity analysis. Mutual information is one of them. Another is multivariate generalisations of mutual information, which can work with more than two signals. Channel-coding theory can be used for signal transmission integrity.

Mutual information is, by far, the most popular method from information theory. One disadvantage is that mutual information does not provide any information on what kind of relationship there is between variables. Information theory is generally defined for discrete distributions, with different techniques to discretize continuous data. However, the method and the chosen parameters for discretizing continuous data have a big effect on the computed results. Information theory is generally also a computationally expensive operation to perform. Finally, there are no clear neurophysiological interpretations for information theoretical constructions. Information theory is discussed much more in-depth in the following chapter, chapter 4.

3.2 Connectivity Interpretation

There are several key aspects to keep in mind when interpreting results from a connectivity analysis. Different connectivity measures may or may not take these aspects into consideration. In order to provide a proper connectivity analysis, these key aspects and the properties of the connectivity measures need to be known.

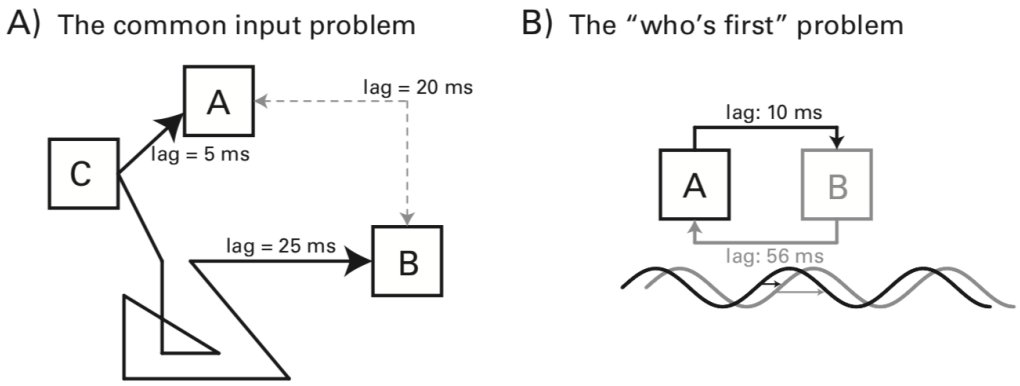
Typically, two different signal are not fully in sync. When an analysis is performed on the signal from multiple electrodes, there may be a phase lag between these electrodes. Most connectivity measures do not take this phase lag into consideration. This does not mean that that these connectivity measures cannot be used, as long as the phase lag is consistent. [20]

Figure 3.1 shows a second key aspect. In this example, region C is connected with A and B. Because C→A is faster than C→B, it may appear that there is a phase lag, even though there is no causal or direct relation between A and B.

3.3 Volume Conduction

The head volume conductor model, discussed in chapter 2, is quite interesting. The brain conducts electrical activity, this is how the electrical activity can be measured on the

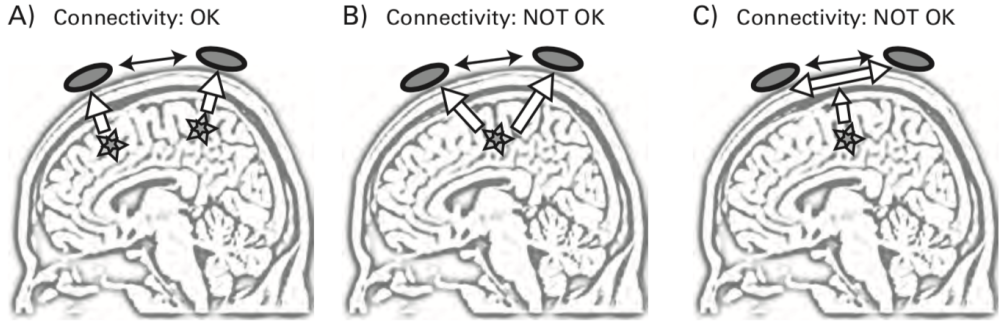
Figure 3.1: Time Lag [8].



scalp. Volume conduction refers to this process of conducting electrical activity through a medium. [5]

Figure 3.2 shows several problems that the reverse problem has to deal with. In situation A, there would be no problem. In this situation, every EEG electrode measures one and only one electrical source within the brain. However, this is not what happens in reality.

Figure 3.2: Volume Conduction [8].



Reality is a combination of situations B and C [5]. Situation B shows that electrical sources in the brain generate large electromagnetic fields which are recorded by more than one EEG electrode. Situation C shows that the scalp also conducts electricity. These two situations have a big effect on connectivity measures.

Chapter 4

Information Theory

Information theory provides us the tools to study the information processing capabilities of different systems. These systems include computers, artificial intelligence and also the brain. The story of Information Theory begins with Shannon who provided the tools to estimate information [32]. The most fundamental aspect of information theory is the concept of the 'bit'.

As a standard, information theory deals with 'bits'. One bit of information represents a choice between two equally probable options. A perfectly balanced coin toss contains one bit of information. It has 50% chance of landing on heads and 50% chance of landing on tails. A bit is a measure of information and a measure of uncertainty. Thus uncertainty and information are tightly intertwined. If we are completely certain about a certain event, there is no information to be gained. [22]

This chapter will cover the most important aspects of information theory. First and foremost, this means discussing entropy. Once we have an understanding of entropy, extensions can be discussed. These include joint entropy and conditional entropy. With these tools, we can go into the area of mutual information, which plays an important role in this thesis. Another important aspect is the generalisation into multivariate systems and dealing with continuous (as opposed to discrete) systems. Finally, some explanation is provided of coding theory and cybernetics.

4.1 Entropy

The next step is entropy. There are different kinds of entropy, such as thermodynamic entropy. Here, we are talking about information entropy. Information entropy is the average uncertainty associated with a random variable. In other words, information entropy is the average rate of information from a random or stochastic variable.

In order to calculate the entropy, we require some discrete random variable X, with $x_1 ... x_n$ the different values from X. $P(x)$ is the probability mass function. The entropy $H(X)$ can be calculated as:

$$H(X) = -\sum_{i=1}^{n} P(x_i) log_2(P(x_i)) \tag{4.1}$$

Interesting to note is that $log_2(P(x_i))$ is the information about value x_i. This formula can be extended into the conditional entropy of two events X and Y. The entropy $H(X|Y)$ is the entropy of random variable X given that the outcome of Y is known.

$$H(X|Y) = \sum_{i,j} P(x_i, y_j) log_2(\frac{P(y_j)}{P(x_i, y_j)}) = -\sum_{i,j} P(x_i, y_j) log_2(\frac{P(x_i, y_j)}{P(y_j)}) \tag{4.2}$$

If random variables X and Y are independent of each other, then $H(X|Y) = H(X)$. If the variables are completely independent, knowing anything about Y, will not change anything we know about X. Similarly, if $H(X|Y) = 0$, then X is completely determined by Y.

The rule of Bayes is also applicable to conditional entropy:

$$H(X|Y) = H(Y|X) + H(X) - H(Y) \tag{4.3}$$

4.2 Joint Entropy

$H(X)$ and $H(X|Y)$ are basic notions of information measures. Figure 4.1 visualises the notion of entropy. This figure also contains two information measures that are not yet described, $I(X, Y)$ and $H(X, Y)$. Respectively, they are the mutual information and the joint entropy.

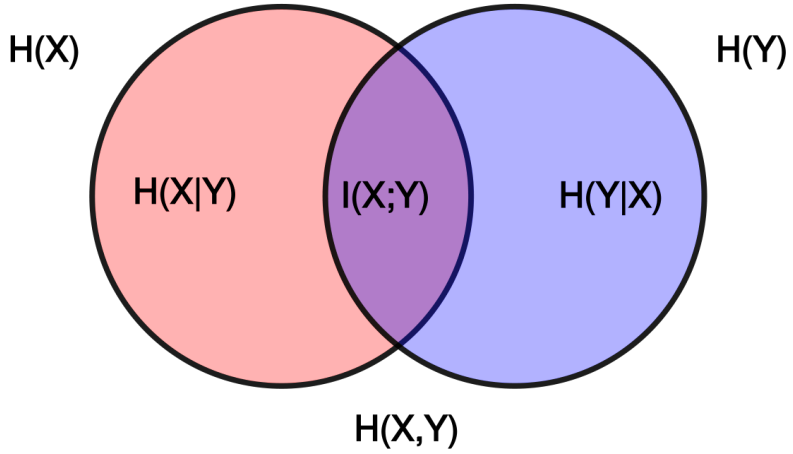
In the figure, $H(X, Y)$ is the complete information content of X and Y. Using the following equation, we can calculate $H(X, Y)$:

$$H(X, Y) = -\sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) log_2(P(x_i, y_j)) \tag{4.4}$$

Using figure 4.1, we can see some interesting relations. $H(X, Y)$ is related to $H(X|Y)$ and $H(X)$:

$$H(X, Y) = H(X|Y) + H(Y) \tag{4.5}$$

Figure 4.1: Venn diagram for information measures.



This can be validated:

$$
\begin{aligned}
H(X,Y) &= -\sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j)log_2(P(x_i,y_j)) \\
&= -\sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j)log_2(P(y_i)P(x_i|y_j)) \\
&= -\sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j)log_2(P(y_i)) - \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j)log_2(P(x_i|y_j)) \\
&= -\sum_{j=1}^{m} P(y_j)log_2(P(y_i)) - \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j)log_2(P(x_i|y_j)) \\
&= H(Y) + H(X|Y)
\end{aligned}
$$

This relation is useful for the actual implementation of information theoretical algorithms. An important note about entropy, conditional entropy and joint entropy is that they are non-negative. It would not make sense for a random variable to have a negative information content. Joint entropy is also always greater or equal to individual entropy and joint entropy is smaller or equal to the sum of individual entropies.

$$H(X) \geq 0 \tag{4.6}$$

$$H(X|Y) \geq 0 \tag{4.7}$$

$$H(X,Y) \geq 0 \tag{4.8}$$

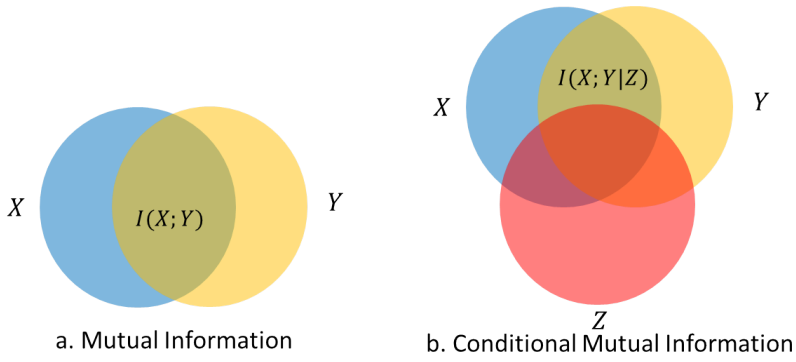
$$H(X,Y) \geq H(X) \tag{4.9}$$

$$H(X,Y) \leq H(X) + H(Y) \tag{4.10}$$

4.3 Mutual Information

Mutual information is the amount of information that is common between two variables. Figure 4.1 shows the mutual information as the intersection between $H(X)$ and $H(Y)$. This can also be seen in figure 4.2. The figure also shows how mutual information can be computed. The individual entropies are summed and the joint entropy is subtracted. Another equal interpretation of mutual information is that mutual information measures the reduction in uncertainty or information when we observe one of the variables.

Figure 4.2: Venn diagram for mutual information.



$$I(X,Y) = H(X) + H(Y) - H(X,Y) \tag{4.11}$$

$$I(X,Y) = H(X) - H(X|Y) \tag{4.12}$$

Mutual information can also be calculated using the probability mass functions directly. In this case we get:

$$I(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j) log_2(\frac{P(x_i,y_j)}{P(x_i)P(y_j)}) \tag{4.13}$$

While mutual information does show whether there is a relationship or correlation between two variables, it does not give information about the "shape" of the relationship. Making the analogy with figure 4.1, mutual information shows how much overlap there is, but it does not explain anything else. Mutual information makes no assumptions of what the distribution of the variables X and Y is like.

Mutual information, like entropy, is non-negative. Additionally, just like there is conditional entropy, there is also conditional mutual information. Figure 4.2 shows the conditional mutual information. It refers to the entropy in common between X and Y, with the exclusion of Z. Conditional mutual information indicates the mutual information given another variable is given:

$$I(X,Y|Z) = \sum_{k=1}^{o}\sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j,z_k)log_2(\frac{P(x_i,y_jz_k)P(z_k)}{P(x_i,z_k)P(y_jz_k)}) \tag{4.14}$$

4.4 Multi-Variate Information Theory

In the previous sections, information theory has been observed through a bivariate lens. These methods can be generalised to multivariate variations. When comparing different brain regions, we do not want to restrict ourselves to a bivariate case.

Joint entropy can be easily extended to a multivariate case. In the bivariate case, the joint probability mass function was used and a summation over both random variables was done. The multivariate case simply generalises this equation:

$$H(X_1,...,X_n) = -\sum_{x_1}...\sum_{x_n} P(x_1,...,x_n)log_2(P(x_1,...,x_n)) \tag{4.15}$$

Equation 4.5 can also be extended into a multivariate case. In this case, we get:

$$H(X_1,...,X_n) = -\sum_{k=1}^{n} H(X_k|X_{k-1},...,X_1) \tag{4.16}$$

The multivariate case of conditional entropy becomes:

$$H(Y|X_1,...,X_n) = H(Y,X_1,...,X_n) - H(X_1,...,X_n) \tag{4.17}$$

We can also make a formula to calculate the entropy of multiple variables conditioned on a single variable:

$$H(X_1,...,X_n|Y) = H(Y,X_1,...,X_n) - H(Y) \tag{4.18}$$

Having extended the different forms of entropy into a multivariate case, we can make a multivariate method for mutual information. The multivariate mutual information becomes a recursive function:

$$I(X_1, ..., X_{n+1}) = I(X_1, ..., X_n) - I(X_1, ..., X_n|X_{n+1}) \tag{4.19}$$

The multivariate mutual information can also be decomposed into a som of entropies, which makes it easier to calculate:

$$I(X_1, ..., X_n) = \sum_{T \subseteq \{1,...,n\}} (-1)^{|T|} H(T) \tag{4.20}$$

$$I(X_1, ..., X_n|Y) = \sum_{T \subseteq \{1,...,n\}} (-1)^{|T|} H(T|Y) \tag{4.21}$$

With these formulas, the discussed entropies are formulated in a multivariate way. This specific formulation of multivariate mutual information is also called interaction information in the literature.

Another interesting note is that most equations can be reduced into a sum of (joint) entropies. Equation 4.21 is formulated as a sum of conditional entropies. But with equation 4.18, we can reduce equation 4.21 into a sum of entropies without conditional variables.

Being able to reduce most equations into a sum of entropies becomes a useful ability during the implementation of these equations.

4.5 Directed Information

Regular mutual information shows the amount of entropy that is shared between two or more random variables. The mutual information can capture both linear and non-linear relationships. However, mutual information cannot show the information flow between different variables.

In order to show the information flow, an extension to mutual information has to be made. This extension is called directed information. Directed information uses processes rather than variables. A process X^n is a sequence, a vector of data $X_1, ..., X_n$. A time series, by its nature, can be seen as a process. [2]

$$I(X^n \to Y^n) = \sum_{i=1}^{n} I(X^i, Y_i|Y^{i-1}) \tag{4.22}$$

An additional note is that directed information is computationally expensive to compute. Multivariate mutual information has to be calculated n times. This makes computation of long processes slow to compute.

4.6 Continuous Data

Up until this point, the discussed equations assumed that discrete data was available. However, within the context of neuroscience and the brain, data is almost always continuous in nature. In the case of EEG research, the data comprises of time series of electrical activity.

In the case of this thesis, the data is a time series of source-reconstructed EEG data. In order to compute entropy with continuous data, the data needs to be discretized [21]. There are multiple ways to discretize data:

- Histogram analysis
- Correlation analysis
- Clustering analysis
- Decision tree analysis
- Probability modelling
- Binning

These ways utilise different tools, such as histograms, correlation, clustering, decision trees, probability distributions, in order to discretize the data.

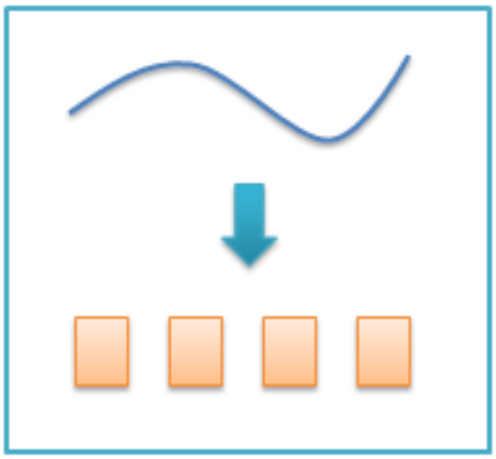
The default choice is to utilise binning. Binning is a relatively simple process that is fast to compute. This is what makes binning attractive. Utilisation of other ways to discretize data is left for future work.

4.6.1 Binning

Binning is the process of putting the continuous data into a predetermined number of bins. In order to avoid adding any kind of noise or bias to the data, every bin should have an equal width. In order to map the continuous data onto the bins, every bin needs to represent a certain interval. The interval of the first bin starts at the minimum value of the continues data. The interval of the last bin ends at the maximum value of the continuous data. The width of the bin is decided by the number of bins. [21]

The big question is how many bins should be used to discretize the data. If too little bins are used, information from the continuous data is lost. On the other hand, with a

Figure 4.3: Binning representation.



very large amount of bins, each bins will only contain one value. In this case, generality is lost.

The number of bins used has a large impact on the entropy, so making the correct choice is important. There are multiple ways to make the correct choice. The number of bins can be estimated empirically by plotting the entropy with varying bin sizes. Using this method, the issues that come with a too small or too large bin size can be vizualised.

There is another, more statistically sound, method for determining the appropriate amount of bins. There is the Freedman-Diaconis rule. This rule states that the optimal amount of bins is related to the interquartile range Q_x, the amount of data point n and the minimum $min(x)$ and maximum $max(x)$ elements of the data:

$$nbins = \frac{max(x) - min(x)}{2Q_x n^{-1/3}} \tag{4.23}$$

4.7 Summary

This chapter provides the necessary mathematical background in order to accomplish the information theoretical analysis of the EEG source-reconstructed data. As explained in section 4.4, most equations can be reduced into a sum of (joint) entropies, which becomes important for the implementation.

To summarize, the two most important aspects from this chapter are:

1. Multivariate mutual information
2. Dealing with continuous data

In order to accomplish the analysis from this thesis, the continuous data has to be discretized. More specifically, this thesis uses the binning method in order to discretize the data. The analysis itself focusses on multivariate mutual information. Chapter 6 will go in depth into the actual analysis that has been done on the data. The binning and the different analysis experiments are described. Chapter 7 goes in depth into the source code that has been developed for the analysis. There are several important aspects to be discussed about the source code.

Chapter 5

EEG Experiment

Applying information theoretical algorithms on a high density EEG dataset is the main focus of this thesis. The EEG dataset is provided by the lab of computational neuroscience of KU Leuven. In order to perform an analysis of a dataset, it is important to understand what kind of data we are dealing with.

This chapter discusses the experiment used to generate the dataset. The experiment and the reasons behind the experiment are introduced. Afterward the actual experiment is discussed. This starts with the an explanation of the EEG experiment itself and the paradigm. Afterwards the source localization is discussed. Finally, the final structure of the data is brievely discussed.

5.1 Introduction

The experiment resolves around the human brain's representation of semantic categories. By using high density EEG recordings, the semantic processing of words within the human brain was captured. Within the literature, there are several different theories and concepts that explain the representation of semantic categories within the human brain. The human brain is a complicated organ with many different cortical areas. Each theory focusses on different cortical areas and explain how they are involved with the representation of semantic categories. These theories have been further developed with the aid of different neuroimaging experiments.

One of the most prominant theories of semantic word processing is the grounded cognition model. The model is also called the embodied cognition model. This model explains that semantic knowledge is kept within high-level perception and motor representation systems in the brain.

This means that a word is comprehended based on modality-specific neural systems. In other words, elements such as visual and auditory features are used to define words. As an example, animate objects cluster in the more lateral aspect of the fusiform gyrus, whereas activations associated with inanimate or man-made objects cluster in the more

medial aspect of the fusiform gyrus [25]. There are other studies that show that damage to the brain's modal system creates category-specific deficits [1, 7].

These deficits can be explained by using the grounded model, natural objects such as animals, vegetation, etc, are distinguished primarily from their visual semantic properties. On the other hand, man-made objects such as tools and vehicles are distinguished primarily by their function. This is also shown by fMRI and PET imaging studies [11].

The grounded cognition model can be criticized from several aspects. Stimuli that are conceptual in nature are much more difficult to control, as the associated brain activity is almost completely subject-dependent [23]. Most importantly, the grounded cognition model only explains features that are related to a physical object. More abstract concepts, such as concepts related to emotions, are not explained by the model.

Looking at the difference in the processing of abstract and concrete words, there are some theories. The main theories are the dual coding theorem and the context availability theorem [24, 36]. The dual coding theorem states that there are two separate systems within the brain, a nonverbal "imagery" system and a verbal "linguistic" system. The nonverbal system implement is very alike to the grounded cognition model. The verbal system is involved with the abstract nature of language.

The context availability theorem is quite different from the dual coding theorem. Rather than having two distinct systems, the context availability theorem states that the processing of concrete and abstract words never happen in isolation and that context is important. Context determines how words are processed. Concrete words are contextually related to their physical referents. Abstract words are more variable and are contextually related to previous experience.

Neither theory has been proven or disproven by scientific literature. Most research into these theories have been done using functional neuroimaging techniques such as fMRI and PET, which is limited in terms of temporal resolution [3]. EEG, by virtue of its excellent temporal resolution, has become popular to probe the brain's detailed processing of objects and words. Several studies using EEG/ERP recordings have successfully distinguished different word categories. This has been done on different stages of semantic processing [18]. One disadvantage to using EEG is the low spatial resolution, which makes it fall short in detecting cortical network activiation differences.

In the experiment described in this chapter, high density EEG recordings were used to capture the fast dynamics involved in semantic processing. The neural activity has been localized on the cortex with an accuracy in the range of millimetres and milliseconds. This provides an interesting and unique opportunity to analyse the activity during the semantic processing of abstract and concrete words.

5.2 Materials and methods

The experiment was done in a sound-attenuated, darkened room with a constant temperature of 20 degrees, sitting in front of an LCD screen at a distance of about 70cm.

EEG data is recorded using 128 active Ag/AgCl electrodes (SynampsRT, Compumedics, France), according to the international 10-20 system. Two of these electrodes serve as ground (AFz) and reference (FCz). The EEG signal is recorded at a 2 KHz sampling rate and downsampled to 500 Hz. All electrodes are mounted in an electrode cap that is placed on the subject's head (Easycap, Germany).

5.3 Experimental Paradigm

Choosing the correct paradigm is important. The paradigm has to make sure that the subjects are involved in lexical access and semantic processing. Therefore, a categorization paradigm was selected. This paradigm consists out of 600 Dutch words, all of which are nouns, taken from the database of concreteness ratings for 30,000 Dutch words [6].

These words are split up into two groups, abstract and concrete words. Abstracts words have a concreteness rating of maximum 2.5, and concrete words have a concreteness rating of minimum 3.5. The concreteness rating of abstract and concrete words are statistically different, as tested by a t-test.

Both groups are controlled for word length and word frequency, having no significant difference. Word length and word frequency are computed from the Dutch CLEARPOND software. The words are also pseudo randomly organized in order to make sure that no two consecutive words have a high forward association, and forward associations is controlled for all consecutive words. Forward associations was taken from the Dutch free association network created by De Deyne et al [9].

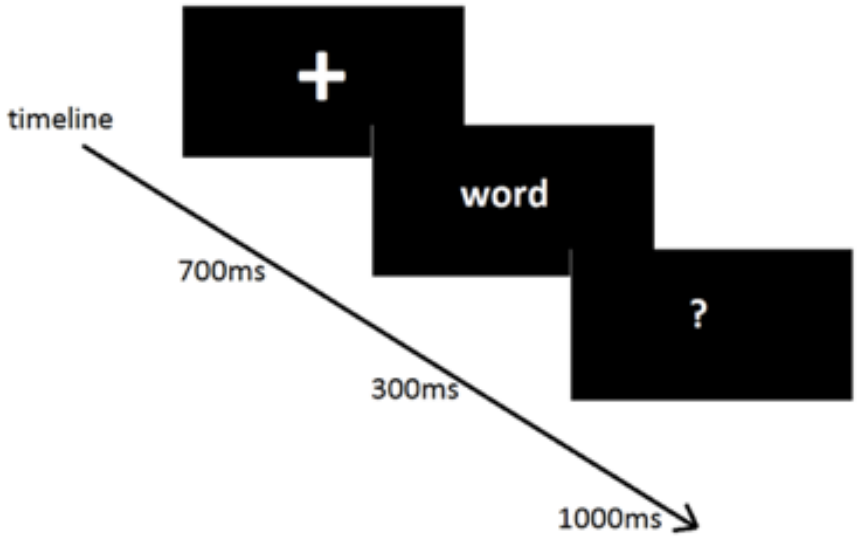
Figure 5.1 shows the general flow for each trial. In each trial, a single word in white text is presented on a fully black screen. The word is shown for 300ms. This is followed by a question mark that is shown for a duration of 1 second. Each trials starts by showing a fixation cross to cue subjects to focus on the middle part of the screen.

Subjects are asked to press the mouse button as soon as they the question mark and the word they are shown is a colour. The colour category is used as a non-target, or filler, and are not included in the results. After the button press, the subjects are given visual feedback on the button press. They are shown "kleur!" (colour) if they pressed correctly, and "fout!" (wrong) otherwise.

5.4 Source localization

In order to source reconstruct the EEG data, the Brainstorm toolbox is used [33]. The Brainstorm toolbox is freely available under the GNU general public license. The default

Figure 5.1: General Flow for each Trial.



anatomy is based on the ICBM-152 template.

OpenMEEG BEM model is used for the forward model [14]. In this model, the cortex is divided into 15,000 dipoles. By merging the matrices computed from the baseline of all selected trials, the noise covariance and data covariance matrices are obtained.

Section 2.4 discussed different algorithm that can be used for the inverse modelling method. In this case, sLORETA is used [28]. sLORETA yields zero localization error. Source orientation is constrained to be orthogonal to the cortical surface. The signal-to-noise ratio (SNR) is kept at the default suggested value, which is 3. Sulci are not taken during the analysis. Brainstorm's documentation states that accurate source localization in these regions is implausible.

Using different source localization algorithms, the correctness of the procedure is verified [26]. The results are analysed by taking the average over all trials regardless of semantic features, using the four methods available in the brainstorm toolbox: wMNE, dSPM, sLORETA, and unconstrained sLORETA.

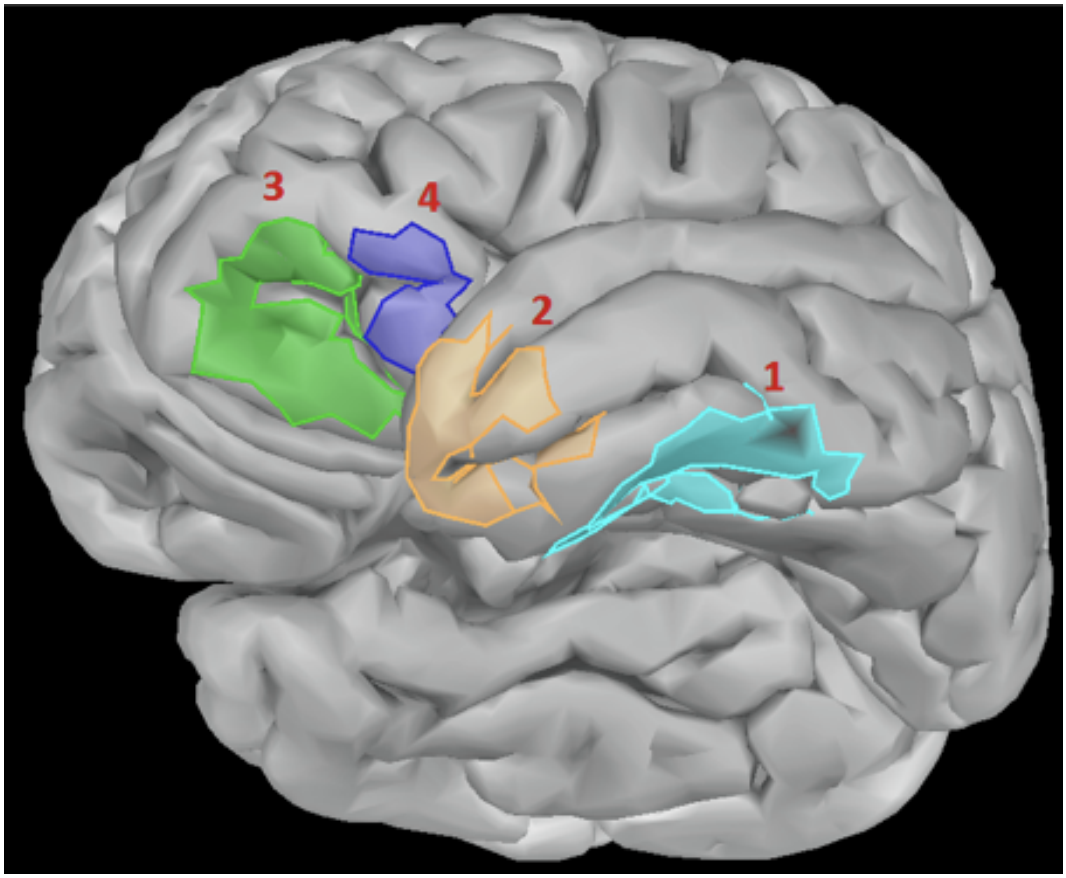
5.5 Region of Interest selection

In order to be able to focus on a very specific set of data, the delivered data only contained several regions of interest. The procedure for selecting the regions of interest starts from the source reconstructed data, obtained by using sLORETA. The current density maps were normalized with respect to a reference level in order to provide a

statistical map. This statistical map is essentially the signal to noise ratio of the current estimate as a function of location. The normalized source maps are used to evaluate the significance of the data.

The sLORETA statistical maps are averaged over all trials and activity below 75% of the maximum activity is eliminated. This is done for both paradigms (abstract and concrete). The two obtained maps are overlapped and from this the most active regions with an estimated square size of larger than $3cm^2$ are selected. This resulted in four regions on the inferior temporal gyrus, temporal pole, inferior frontal, and anterior orbital gyrus. Figure 5.2 visualizes the selected regions.

Figure 5.2: Visualization of the Selected Brain Regions.



Chapter 6

Analysis and Discussion

This chapter discusses the information theoretical analysis of the dataset. The source code developed for the analysis has been implemented in the Python programming language. The actual implementation will be discussed in-depth in chapter 7.

The dataset contains source-reconstructed data. The trials within the dataset encompass 13 subjects and contains 3404 trials. Each trial consists of 340 data points, or samples, because each trial is 1.7 seconds long with a 200Hz sampling rate. The dataset represents 4 cortical regions. For these 4 regions, there are three kinds of time series, data recorded from these regions during abstract word representation, concrete word representation, and rest. These time series are used for the analysis.

The analysis described in this chapter is performed on all 4 regions. The results are consistent across the 4 regions. Therefore, the figures in this chapter focus on a single region in order to remove redundancy.

First, this chapter will discuss the procedure for finding the amount of bins that will be used throughout the analysis. In order to be able to utilise the binning method for entropy estimation, the correct number of bins needs to be computed.

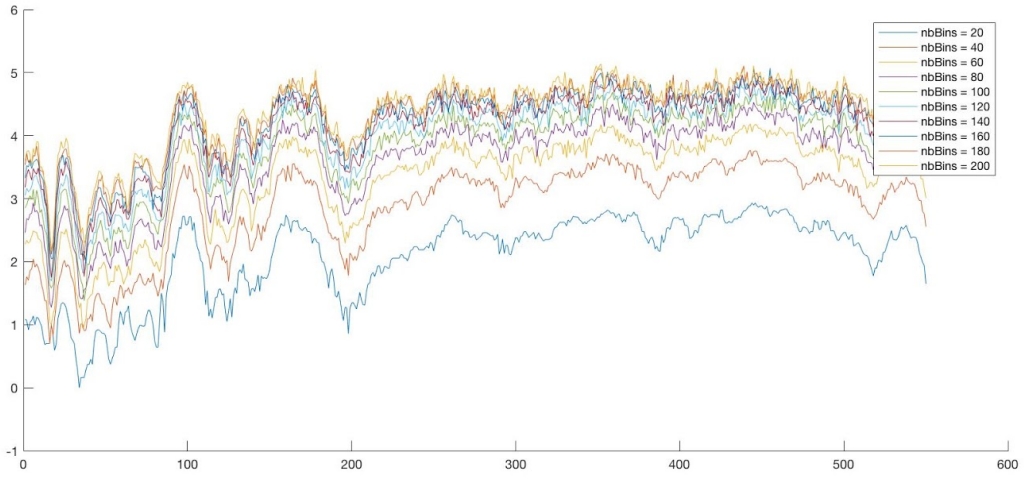
Afterwards, the main analysis is discussed. The cortical regions which were active during both the abstractness and concreteness experiments are the main focus. These are the 4 regions described above. Secondly, the analysis is adapted in order to analyse the different subjects independently. Finally, an analysis was made for a comparison between the active regions and the resting state counter parts.

6.1 Calculating Bin Sizes

Section 4.6 discussed the discretization of continuous data. There are multiple methods to calculate the optimal number of bins. The number of bins can be calculated empirically. This can be seen in figure 6.1. The entropy is calculated for a number of different bin sizes. Using this method, the entropy seems to converge for the bin sizes. The convergence

starts from a bin size of about 100. This gives an indication for the optimal bin size.

Figure 6.1: Entropy comparison between bin sizes



There is also equation 4.23. Using this equation, the optimal bin size is calculated to be 90. Both the equation and empirical method find approximately the same result. For the analysis, a bin size of 90 has been chosen.

6.2 Comparing the Cortical Region

As explained before, the time series recorded from the different regions during abstract word representation, concrete word representation, and rest are the main focus. The main strength of an information theoretical approach is the measurement of mutual information. Mutual information can measure both linear and non-linear relationships.

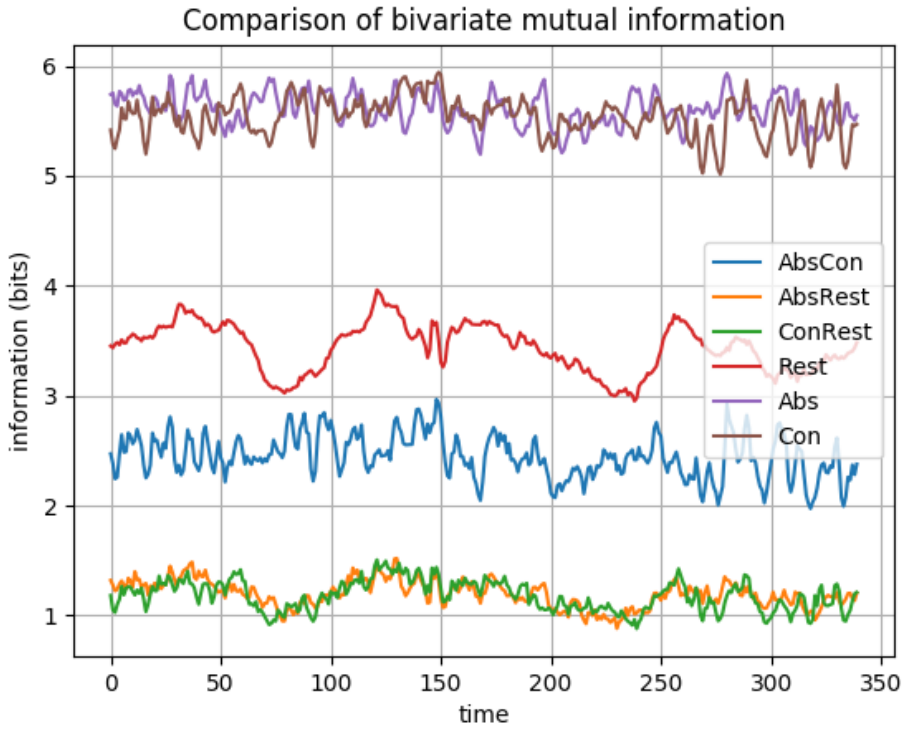
From the activity, we could see that some cortical areas are active during both the abstractness and concreteness experiments. Mutual information can tell us what kind of relationship there is between the abstractness and concreteness. There are several different possibilities. One theory is that the cortical region represents the semantic processing that is common between the abstractness and concreteness experiments.

Figure 6.2 shows the results of the analysis. Abs, Con and $Rest$ represents the information entropy through time. Abs is the information entropy from the abstractness experiment. Con is the information from the concreteness experiment.

Abs and Con have about the same information entropy throughout time. This corresponds to the fact that this region is active and doing semantic processing. $Rest$ is the information entropy during resting state.

$AbsRest$ and $ConRest$ represent the mutual information between, respectively, abstract-

Figure 6.2: Experiment



ness and rest, and, concreteness and rest. The mutual information is much lower than the entropy from the resting state. This indicates that whatever is happening during resting state is very different from the activity during semantic processing. This is a small, but very important observation.

$AbsCon$ represents the mutual information between the abstractness experiment and the concreteness experiment. One observation is that $AbsCon$ is on an equal footing with the information entropy during the resting state. This can, wrongfully, lead to the conclusion that the only equal information between abstractness and concreteness is the activity that occurs during the resting state.

However, considering that $AbsRest$ and $ConRest$ are much lower, this conclusion cannot be made. This analysis finds that there is some mutual information between abstractness, concreteness and resting state. There is also some mutual information between the abstractness and concreteness, which is not completely caused by the resting state activity.

Figure 6.3 roughly visualizes these relationships information. There is some common information (rounded to 1 bits) between all 3 datasets. There is also some information that is unique to each of the datasets. Abstractness and concreteness share some

information, which is unrelated to the resting state. This venn diagram helps visualize some results, but is very deceptive for visualizing multivariate mutual information.

There are some conclusions to be made from this. First of all, some of the semantic processing of abstractness and concreteness in the same cortical area is shared. However, it seems that the same cortical area does not compute exactly the same features during semantic processing of abstract and concrete words.

Figure 6.3: Venn Diagram for Information in Cortical Region [34].

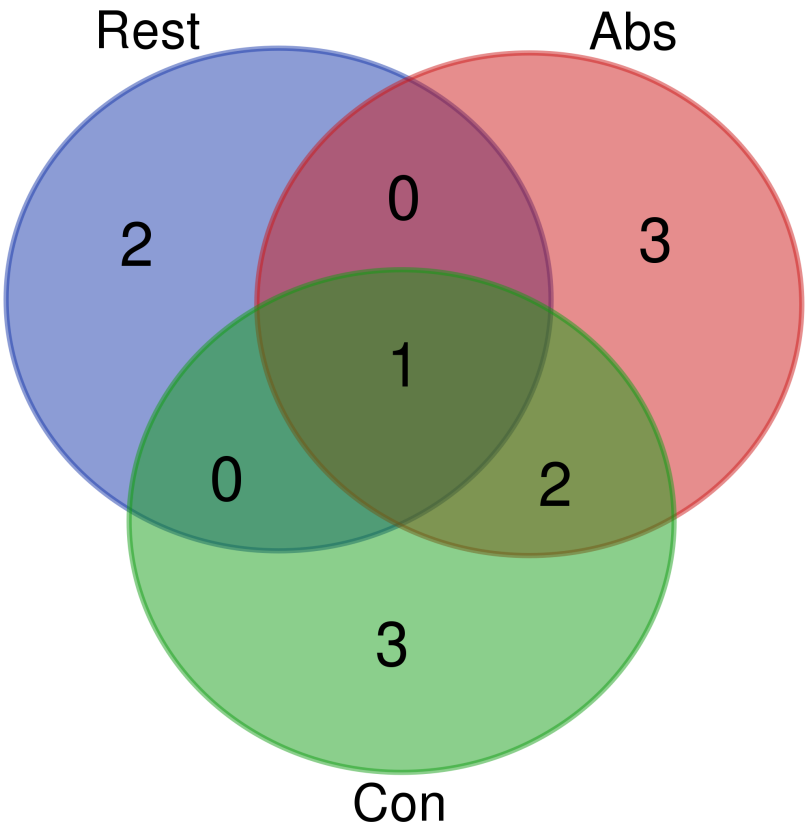
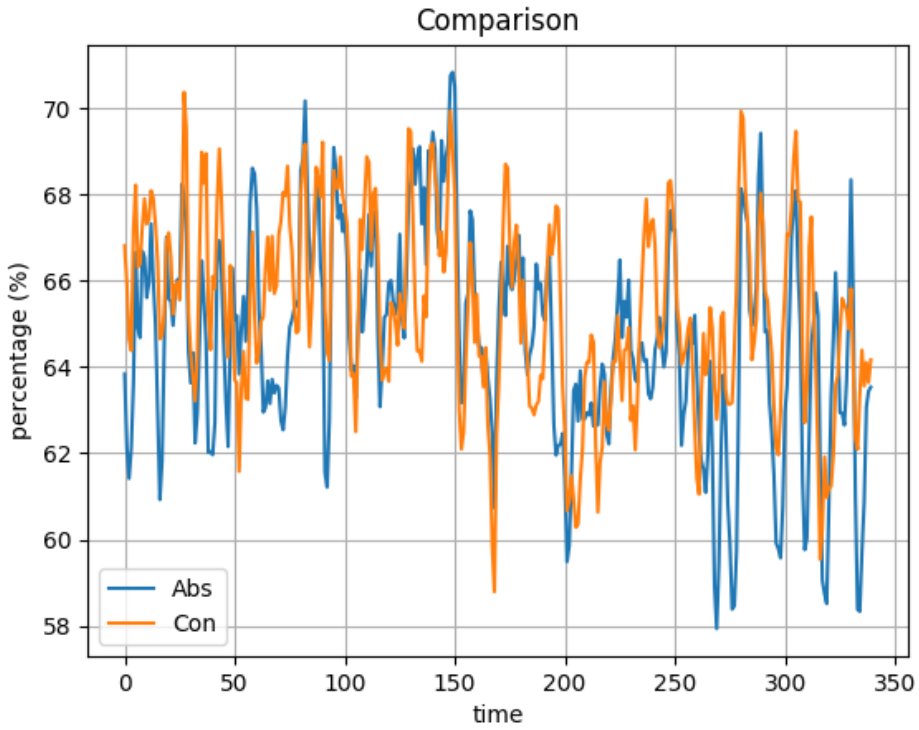


Figure 6.4 shows this relation expressed as percentages. About 60% to 70% of the activity is shared between the semantic processing of abstract and concrete words. This leaves 40% to 30% of the activity that is unique for the semantic processing of abstract words and for the semantic processing of concrete words.

6.3 Comparing the Cortical Region: Multivariate

So far, bivariate mutual information has been utilised. However, there is also the multivariate mutual information. This can be used to calculate the mutual information between the abstractness, concreteness and resting state.

Figure 6.4: Comparison



Looking at figure 6.3, you would expect the multivariate mutual information to be 1 bit. However, the actual result is -1 bit, as seen in figure 6.5. At first, this seems very strange. Having a negative amount of information is very counter-intuitive.

Looking at the equation that is used for the multivariate mutual information, we get:

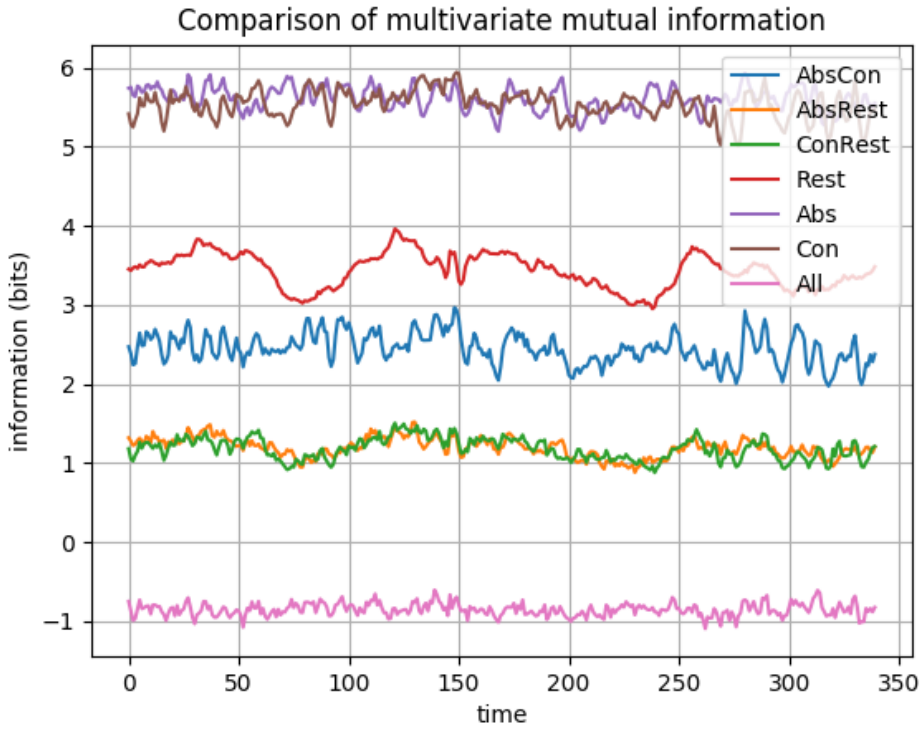
$$I(Abs, Con, Rest) = I(Abs, Con) - I(Abs, Con|Rest) \tag{6.1}$$

Looking at figure 6.5, we can see that $I(Abs, Con)$ has a value between 2 and 3 bits of information. However, $I(Abs, Con|Rest)$ is a bit more difficult. This can also be written as a sum of entropies:

$$I(Abs, Con|Rest) = H(Abs, Rest) + H(Con, Rest) - H(Abs, Con, Rest) - H(Rest) \tag{6.2}$$

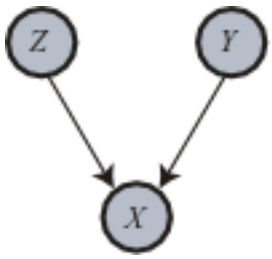
The deceptive point is that the joint entropy $H(Abs, Con, Rest)$ isn't as large as the venn diagram makes it look. The result is that $I(Abs, Con|Rest) > I(Abs, Con)$, which causes a negative multivariate mutual information.

Figure 6.5: Multivariate Experiment



Negative multivariate mutual information indicates a synergy. Figure 6.6 shows a graphical model detailing this. Z and Y represent Abs and Con, while X represents $Rest$. $Rest$ causes a stronger dependency between Abs and Con, than without $Rest$. This shows that multivariate mutual information is quite difficult to interpret.

Figure 6.6: Multivariate Mutual Information Synergy

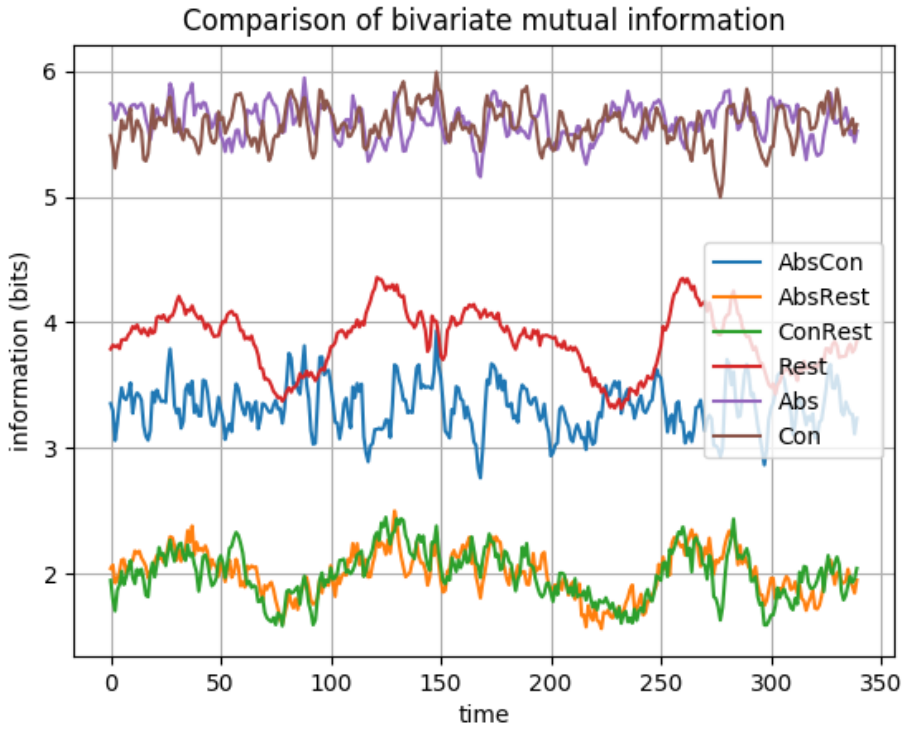


6.4 Per Subject Comparison

A per-subject analysis has also been performed. With a per-subject analysis, we can see whether results would vary between different subjects and whether there are different conclusions to be made.

Figure 6.7 shows the analysis from section 6.2 executed on a single subject. The figure looks very similar to figure all-channel-1. This reaffirms the conclusion on a per-subject basis.

Figure 6.7: Comparison for Subject 1



6.4.1 Effect of Trials Used

Another interesting avenue to pursue was analysing the effect of the amount of trials used within information theoretical equations. In other words, how much data do we need in order to get decent results. We expected that information theory would still deliver good results, even with relatively little data.

Several additional analyses were performed. Each analysis used a different amount of trials. The complete dataset contains 3404 trials, with an average of 262 trials per subject (per timepoint). With binning being done with 90 bins, this means that there are about 3 trials in each bin.

In order to see the effect of the amount of trials that are used, the information theoretical analysis was performed with 10, 40, 80 and 100 trials.

Figure 6.8 shows the result from only using 10 trials. With 90 bins, most bins are completely empty. Due to this, the computed information entropy is very inaccurate.

Figure 6.8: Comparison for Subject 1 - 10 trials

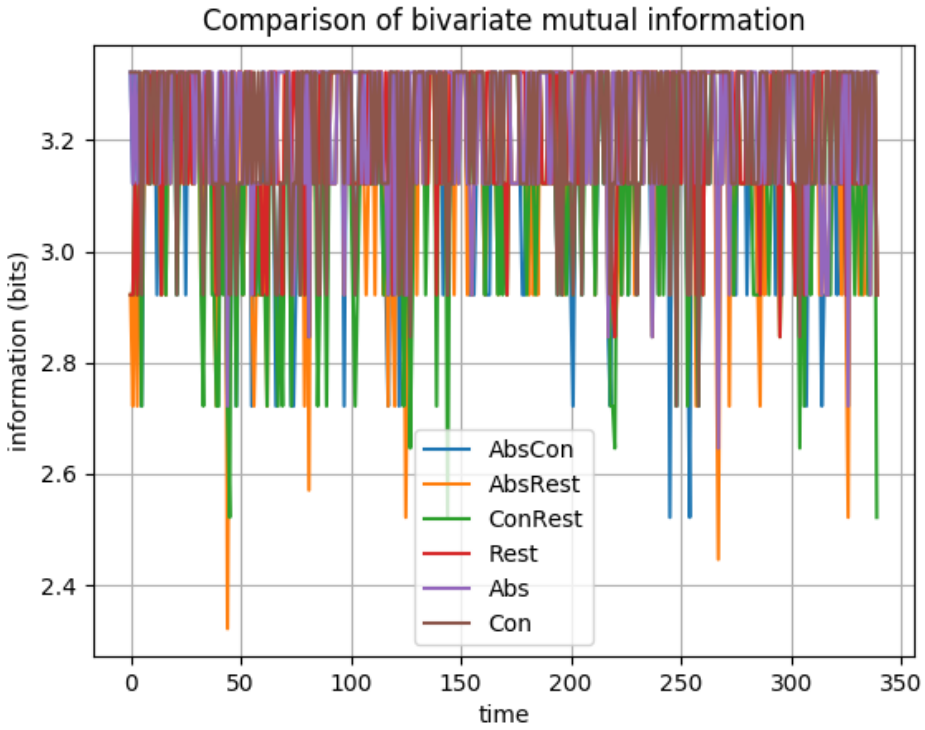


Figure 6.9 shows the result from only using 40 trials. In this case, the result starts becoming more accurate.

Figure 6.10 shows the result from using 80 trials. With 80 trials, nearly every bin should be filled. From this point, the resolution starts becoming much more accurate.

Figure 6.11 shows the result from using 100 trials. The resolution increases yet again. However, the new details that are visible in the graph do not cause more conclusions or results to be drawn.

This analysis shows that, if not enough data is available, information theoretical equations cannot reliably be used. For the 10 and 40 datapoint analyses, this is clearly shown. However, the 80 and 100 datapoint analyses show that even with relatively few trials, relative to the amount of bins used, the resolution is already quite clear.

Figure 6.9: Comparison for Subject 1 - 40 trials

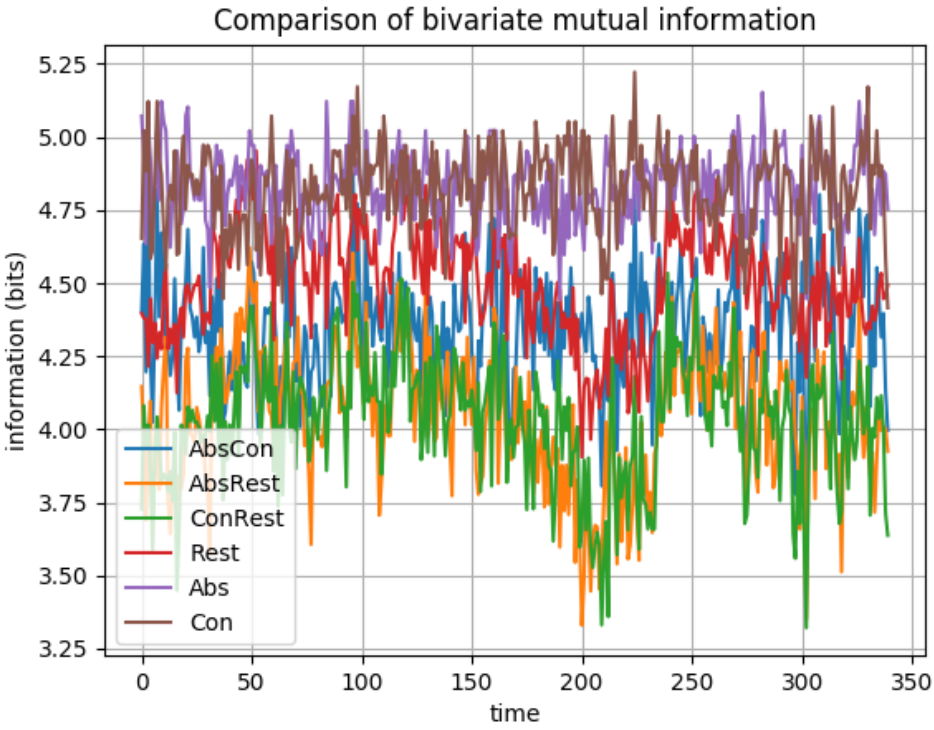


Figure 6.10: Comparison for Subject 1 - 80 trials

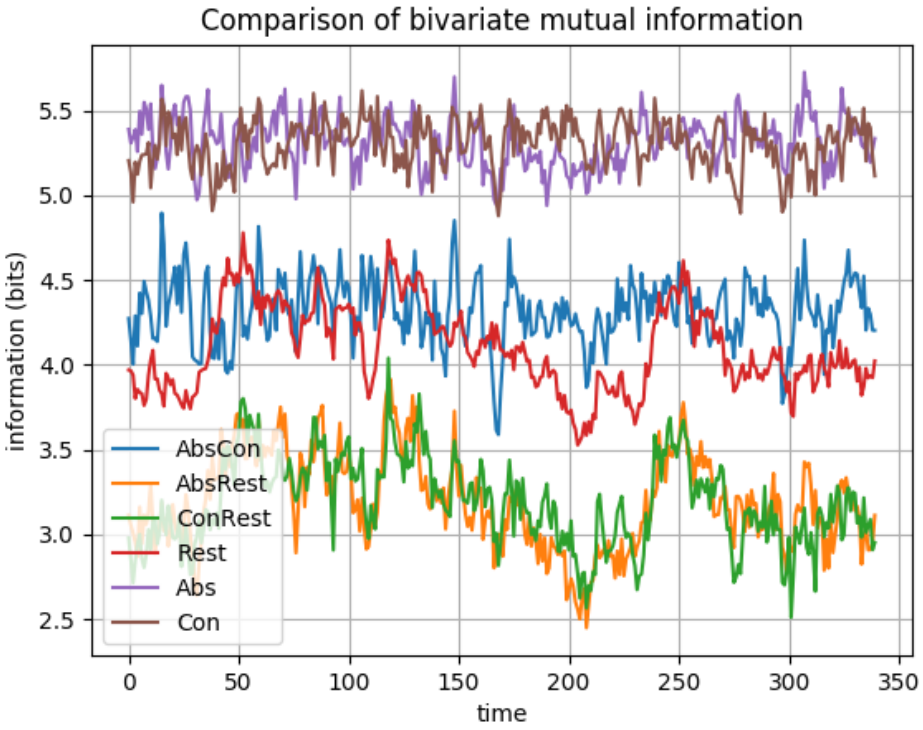
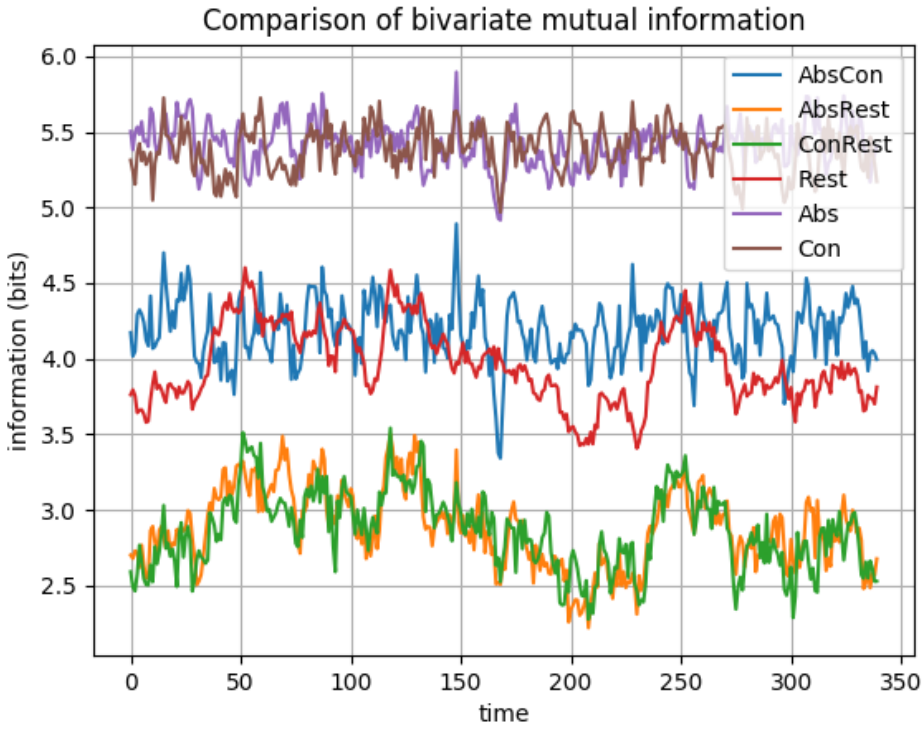


Figure 6.11: Comparison for Subject 1 - 100 trials



Chapter 7

Implementation

This chapter discusses the implementation made for the analysis described in the previous section. The implementation is important for several different reasons. For the sake of reproducability, it is important that it is understood why certain implementation choices were made.

Secondly, we strive for open science. Open science is about making research publicly available for society. A desired outcome is to be able to make the information theoretical analysis framework openly available for others to use. This also means that certain choices need to be made during the implementation.

This chapter starts with a motivation for the choice of the programming language, Python, used for the implementation. The data which needed to be analysed was available in a Matlab format. In order to read the data in Python, the data needed to be converted into a different format. Finally, the implementation of the information theoretical equations are discussed.

7.1 Python Programming Language

The Python programming language is becoming more and more popular in a multitude of different fields. Python looks a lot like typical pseudocode, which makes it highly readable. Python is also modular and has a large standard library. These features are the main cause behind the surge in popularity.

The scientific community is becoming more interested in Python due to its numerical libraries, such as *NumPy*, which are highly optimized. Arguably, these numerical libraries place Python on a equal footing with Matlab, another programming language popular among scientists. Unlike Matlab, Python is completely free and open-source, making it more readily available.

Due to becoming more popular, domain-specific ecosystems of open-source Python software have been developed. The focus on reusable components is an essential aspect

of scientific computing within computational neuroscience.

The different features of Python make it a desirable programming language for information theoretical analysis. The numerical library *NumPy* can easily handle the source-reconstructed EEG data used within this thesis. Secondly, we want to contribute to the open-source community by providing a library for information theoretica analysis of source-reconstructed EEG data.

One immediate disadvantage is that the data is provided as Matlab data. Since Python cannot directly handle Matlab data, a conversion has to be made. Considering the conversion is only a minor inconvenience and happens only once, it was not considered detrimental to the choice of Python.

7.2 Data Conversion

The data is stored in Matlab format which Python cannot work with. In order to be able to utilise Python, the data has to be converted into a format that Python can use. The conversion is done using a Matlab script shown in figure 7.1.

The Matlab script converts a Matlab data structure into JSON. JSON is an universal format. Nearly every programming language has a library available to read a JSON file. The Matlab script has 5 different arguments. loc_mat and $name\_mat$ indicate which Matlab file to load. loc_json and $name\_json$ indicate which JSON file to write to. $name\_data$ indicates which variable from the matlab file has to be exported.

An important note is that JSON is a textual format, not a binary. This causes a small loss of precision. Matlab internally stores real numbers as double-precision floating point numbers. Since JSON is textually, it stores real numbers as a decimal number, as a string. Due to the nature of a floating point representation, converting a float into a string and back into a float is not necessarily lossless. Luckily, the error is negligible.

Figure 7.1: Matlab JSON Encoding

```
function [] = convertToJSON(loc_mat, name_mat, name_data, loc_json,
    ↪ name_json)
    data = load(strcat(loc_mat, name_mat));
    json = jsonencode(data.(name_data));
    fid = fopen(sprintf('%s%s%s', loc_json, name_json, '.json'),'wt');
    fprintf(fid, '%s', json);
    fclose(fid);
end
```

The Matlab script is called from within Python. Matlab provides bindings to Python. Using these bindings, Matlab scripts can be called from within Python. Figure 7.2 shows

the primary Python code used to call the Matlab scripts.

Figure 7.2: Matlab within Python

```
eng = matlab.engine.start_matlab()
for name_mat in mat:
    datasets = mat[name_mat]
    for name_data, name_json in datasets:
        # Check if the json file has been generated already
        if not os.path.isfile(loc_json + name_json + ".json"):
            eng.convertToJSON(loc_mat, name_mat, name_data, loc_json,
                ↪ name_json, nargout=0)
```

So far, we have converted the Matlab data in a JSON format. However, for data analysis, JSON is not a perfect choice. Reading large JSON files is very slow. For this reason, a second conversion is done. The JSON files are converted into pickle files. Pickle is a binary file format that is specifically designed for Python. Pickle files load relatively fast, which makes the experiments faster to compute.

7.3 Information Theoretical Equations

With the data readily available within Python, the only ingredient left before the experiments can begin are the information theoretical equations. As explained in Section 4.4, information theoretical equations can be reduced into a sum of joint entropies. This makes the implementation easier and cleaner.

The main function of importance is the implementation of the entropy equation. Figure 7.3 shows the implementation. The $*X$ notation in Python is used to indicate a variable amount of parameters. This entropy function models the multivariate joint entropy equation (Equation 4.15). The histogram function from numpy is used to accomplish the binning.

Figure 7.3: Multivariate Entropy

```
def entropy(bins, *X):
    # Binning of the data
    data = np.histogramdd(X, bins=bins)
    # Calculate probabilities
    data = data[0].astype(float)/data[0].sum()
    # Compute H(X, Y, ..., Z) = sum(P(x, y, ..., z) * log2(P(x, y, ..., z)))
    return np.sum(-data * np.log2(data+sys.float_info.epsilon))
```

With the entropy function, we can make the other functions. The translation of the

information theoretical equations into Python is very straightforward. The simplicity can be seen in figure 7.4 and figure 7.5. These figures model, respectively, equation 4.5 and equation 4.11.

Figure 7.4: Conditional Entropy

```
def conditionalEntropy(bins, X, Y):
    # Compute H(X|Y) = H(X,Y) - H(Y)
    entro2 = entropy(bins, Y)
    entroJoint = entropy(bins, X, Y)
    return entroJoint - entro2
```

Figure 7.5: Mutual Information

```
def mutualInformation(bins, X, Y):
    # Compute I(X,Y) = H(X) + H(Y) - H(X,Y)
    entro1 = entropy(bins, X)
    entro2 = entropy(bins, Y)
    entroJoint = entropy(bins, X, Y)
    return entro1 + entro2 - entroJoint
```

Utilising the $*X$ notation in Python, all equations can be implemented in their multivariate versions. Figure 7.6 implements the multivariate conditional entropy, seen in equation 4.18 and equation 4.17. Figure 7.7 implements multivariate mutual information, seen in equation 4.20. Finally, figure 7.8 implements multivariate conditional mutual information, seen in equation 4.21. It should be noted that figure 7.4 and figure 7.5 become unnecessary with the multivariate versions available $(mutualInformation(bins, X, Y) == mutualInformationMulti(bins, [X, Y]))$.

Figure 7.6: Multivariate Conditional Entropy

```
def conditionalEntropyMulti(bins, XX, YY):
    # X and Y should be lists of random variables
    # XX = X_1,...,X_n
    # YY = Y_1,...,Y_n
    # Compute H(XX|YY) = H(XX, YY) - H(YY)
    # Compute H(X_1,...,X_n|Y_1,...,Y_n) = H(X_1,...,X_n, Y_1,...,Y_n) - H(
        ↪ Y_1,...,Y_n)
    return entropy(bins, *(XX + YY)) - entropy(bins, *YY)
```

Figure 7.7: Multivariate Mutual Information

```
def mutualInformationMulti(bins, *X):
    subsets = get_subsets(*X)
    entr = 0
    for sub in subsets:
        entr += (-1)**(len(sub)) * entropy(bins, sub)
    return entr
```

Figure 7.8: Multivariate Conditional Mutual Information

```
def mutualInformationConditionalMulti(bins, Y, *X):
    subsets = get_subsets(*X)
    entr = 0
    for sub in subsets:
        entr += (-1)**(len(sub)) * conditionalEntropyMulti(bins, sub, [Y])
    return entr
```

7.4 Summary

This chapter shows the implementation that was developed for the information theoretical analysis. The importance of choosing the correct programming language is highlighted, and the reasons for choosing Python have been discussed.

The data conversion has been discussed in depth for several reasons. It shows that the choice of the programming language for the analysis does not necessarily need to depend on the specific data format. The work required to convert data outweighs the disadvantages of working with a programming language that is not desired.

Finally, the discussion of the implementation of the information theoretical equations shows the elegance of Python. It shows that equations can be fairly easily converted into Python and it shows the usefulness of formulating the equations as a sum of entropies. As a final message, this chapter shows that the implementation itself is just as important as the actual data analysis.

Chapter 8

Future Work

Even though this thesis has come to a close, there are still many opportunities to take and avenues to explore. The results from this thesis shows that there is a lot of promise for an information theoretical approach for EEG source-reconstructed data.

8.1 Directed Information

In section 4.5, directed information was described. This is a very interesting algorithm that can lead to many interesting results. Directed information allows information flow between two processes to be calculated.

This is especially useful for information flow and connectivity. In the case of the source-reconstructed EEG data used in this thesis, directed information can be used to compute the flow of information between different regions in the brain.

This could be useful to further investigate what the activity within the common region (between abstractness and concreteness) means. One of the possible questions could be, does the flow of information start within the common region, or does it start in the seperate regions and then flow to the common region.

$$I(X^n \to Y^n) = \sum_{i=1}^{n} I(X^i, Y_i | Y^{i-1}) \tag{8.1}$$

8.2 Open Source Connectivity Package

The implementation has been described in-depth in chapter 7. This was explained in-depth for multiple purposes. Being able to use the source code to develop an open-source connectivity package is one of the main reasons.

Currently, the implementation is not in a state where it can be released as an open-source toolbox. Other connectivity measures need to be added to the source code and the

current implementation needs to be cleaned-up.

8.3 Comparison with Granger Causality

In this thesis, the information theoretical equations have not been compared to other connectivity measures. Most notably, the granger causality could be compared against.

8.4 Multivariate Mutual Information Alternatives

Multivariate mutual information can be difficult to interpret. This is especially the case when the multivariate mutual information is negative. The main cause is that multivariate mutual information heavily shows synergies and redundancies.

However, the multivariate mutual information described in this thesis, interaction information, is not the only generalisation for mutual information. There is also total correlation and dual total correlation. These are non-negative generalizations of mutual information.

Total correlation measures the divergence of the joint entropy to the independent entropies.

$$C(X^1, ..., X^n) = [\sum_{i=1}^{n} H(X^i)] - H(X^1, ..., X^n) \tag{8.2}$$

Dual total correlation is bounded by the joint entropy.

$$D(X^1, ..., X^n) = H(X^1, ..., X^n) - \sum_{i=1}^{n} H(X^i | X^1, ..., X^{i-1}, X^{i+1}, ..., X^n) \tag{8.3}$$

Both of these methods are rarely used within the context of computational neuroscience. Due to the strength of being a non-negative generalizations of mutual information, the measure is easier to interpet. It would therefore be interesting to see how useful total correlation and dual total correlation can be when dealing with EEG source-reconstructed data.

Chapter 9

Conclusion

The main goal of this thesis is to apply information theoretical measures of brain connectivity to a high density EEG source-reconstructed dataset. Information theoretical algorithms are becoming more and more utilised within the field of computational neuroscience.

Information theoretical algorithms are model-free and probability based. They make no assumptions on what kind of distribution the data has. This makes information theoretical algorithms very versatile and an excellent tool to measure connectivity.

This thesis mainly focussed on 2 different information measurements. The first measurement is bivariate mutual information. The second measurement is a generalisation of mutual information, called multivariate mutual information or interaction information.

The provided EEG dataset was produced by an experiment resolving around semantic processing. More specifically, the experiment compared the brain activity between the semantic processing of concrete words and abstract words. The provided EEG dataset had already been source-reconstructed.

One cortical area of interest was an area that was active during both semantic processing of abstract and concrete words. Analysis showed that the actual activity in this area differed between the semantic processing of abstract and concrete words.

The implementation has been carefully developed such that the source code can be used in open-source projects. In the future, we want to further develop the implementation so it can be released as an open-source connectivity package which can be used by other scientists.

In conclusion, an analysis of a high density EEG dataset has been computed and an implementation has been developed. The implementation provides a platform for further information theoretical analysis and research.

Bibliography

[1] L. W. Barsalou, W. K. Simmons, A. K. Barbey, and C. D. Wilson. Grounding conceptual knowledge in modality-specific systems. *Trends in cognitive sciences*, 7(2):84–91, 2003.

[2] K. J. Blinowska. Review of the methods of determination of directed connectivity from multichannel data. *Medical & biological engineering & computing*, 49(5):521–529, 2011.

[3] S. Bookheimer. Functional mri of language: new approaches to understanding the cortical organization of semantic processing. *Annual review of neuroscience*, 25(1):151–188, 2002.

[4] S. L. Bressler and A. K. Seth. Wiener–granger causality: a well established methodology. *Neuroimage*, 58(2):323–329, 2011.

[5] C. Brunner, M. Billinger, M. Seeber, T. R. Mullen, and S. Makeig. Volume conduction influences scalp-based connectivity estimates. *Frontiers in computational neuroscience*, 10:121, 2016.

[6] M. Brysbaert, A. B. Warriner, and V. Kuperman. Concreteness ratings for 40 thousand generally known english word lemmas. *Behavior research methods*, 46(3):904–911, 2014.

[7] A. Caramazza and B. Z. Mahon. The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends in cognitive sciences*, 7(8):354–361, 2003.

[8] M. X. Cohen. *Analyzing neural time series data: theory and practice*. MIT press, 2014.

[9] S. De Deyne and G. Storms. Word associations: Network and semantic properties. *Behavior Research Methods*, 40(1):213–231, 2008.

[10] A. Delorme and S. Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, 2004.

[11] J. T. Devlin, R. P. Russell, M. H. Davis, C. J. Price, H. E. Moss, M. J. Fadili, and L. K. Tyler. Is there an anatomical basis for category-specificity? semantic memory studies in pet and fmri. *Neuropsychologia*, 40(1):54–75, 2002.

[12] L. Faes, R. G. Andrzejak, M. Ding, and D. Kugiumtzis. Methodological advances in brain connectivity. *Computational and mathematical methods in medicine*, 2012, 2012.

[13] K. J. Friston. Functional and effective connectivity: a review. *Brain connectivity*, 1(1):13–36, 2011.

[14] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, and M. S. Hämäläinen. Mne software for processing meg and eeg data. *Neuroimage*, 86:446–460, 2014.

[15] R. Grech, T. Cassar, J. Muscat, K. P. Camilleri, S. G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, and B. Vanrumste. Review on solving the inverse problem in eeg source analysis. *Journal of neuroengineering and rehabilitation*, 5(1):25, 2008.

[16] J. Groß, J. Kujala, M. Hämäläinen, L. Timmermann, A. Schnitzler, and R. Salmelin. Dynamic imaging of coherent sources: studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences*, 98(2):694–699, 2001.

[17] M. S. Hämäläinen and R. J. Ilmoniemi. Interpreting magnetic fields of the brain: minimum norm estimates. *Medical & biological engineering & computing*, 32(1):35–42, 1994.

[18] O. Hauk, M. H. Davis, M. Ford, F. Pulvermüller, and W. D. Marslen-Wilson. The time course of visual word recognition as revealed by linear regression analysis of erp data. *Neuroimage*, 30(4):1383–1400, 2006.

[19] K. Hoechstetter, H. Bornfleth, D. Weckesser, N. Ille, P. Berg, and M. Scherg. Besa source coherence: a new method to study cortical oscillatory coupling. *Brain topography*, 16(4):233–238, 2004.

[20] B. Horwitz. The elusive concept of brain connectivity. *Neuroimage*, 19(2):466–470, 2003.

[21] R. A. Ince, B. L. Giordano, C. Kayser, G. A. Rousselet, J. Gross, and P. G. Schyns. A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula. *Human brain mapping*, 38(3):1541–1573, 2017.

[22] N. Kannathal, M. L. Choo, U. R. Acharya, and P. Sadasivan. Entropies for detection of epilepsy in eeg. *Computer methods and programs in biomedicine*, 80(3):187–194, 2005.

[23] D. Kemmerer. Visual and motor features of the meanings of action verbs: A cognitive neuroscience perspective. In *Cognitive science perspectives on verb representation and processing*, pages 189–212. Springer, 2015.

[24] J. Kounios and P. J. Holcomb. Concreteness effects in semantic processing: Erp evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4):804, 1994.

[25] M. J. Landrum, J. M. Lee, M. Benson, G. Brown, C. Chao, S. Chitipiralla, B. Gu, J. Hart, D. Hoffman, J. Hoover, et al. Clinvar: public archive of interpretations of clinically relevant variants. *Nucleic acids research*, 44(D1):D862–D868, 2015.

[26] K. Mahjoory, V. V. Nikulin, L. Botrel, K. Linkenkaer-Hansen, M. M. Fato, and S. Haufe. Consistency of eeg source localization and connectivity estimates. *Neuroimage*, 152:590–601, 2017.

[27] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen. Fieldtrip: open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational intelligence and neuroscience*, 2011:1, 2011.

[28] R. D. Pascual-Marqui et al. Standardized low-resolution brain electromagnetic tomography (sloreta): technical details.

[29] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *International Journal of psychophysiology*, 18(1):49–65, 1994.

[30] W. Penny. Meg and eeg analysis, 2018.

[31] J.-M. Schoffelen and J. Gross. Source connectivity analysis with meg and eeg. *Human brain mapping*, 30(6):1857–1865, 2009.

[32] J. W. Shin and S. J. Kim. A mathematical theory of communication. *University of Ulinois Ress*, 1949.

[33] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis, and R. M. Leahy. Brainstorm: a user-friendly application for meg/eeg analysis. *Computational intelligence and neuroscience*, 2011:8, 2011.

[34] B. . E. G. UGent. Calculate and draw custom venn diagrams, 2018.

[35] B. D. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on biomedical engineering*, 44(9):867–880, 1997.

[36] J. Wang, J. A. Conder, D. N. Blitzer, and S. V. Shinkareva. Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human brain mapping*, 31(10):1459–1468, 2010.

Master's thesis filing card

Student: Axel Faes

Title: An Information Theoretical Approach to EEG Source-Reconstructed Connectivity

UDC: 681.3*I20

Abstract:

Determining how distinct brain regions are connected and communicate with each other will shed light on how behaviour emerges. In EEG studies, interpreting connectivity measures can be problematic, due to the high correlation between signals recorded from the scalp surface, a result of the volume conductance of the scalp and skin. Therefore, meaningful connectivity patterns can be measured only from the from the spatiotemporal distribution of localised cortical sources, generally referred to as source reconstruction. Still, spurious connectivity issues may persist in source reconstructed EEG data, rendering it vital to choose an appropriate measure of connectivity. In this work, an information theoretical approach, which concerns model-free, probability based methods such as mutual information, conditional mutual information and interaction information, is taken. An information theoretical framework for Python is developed in order to operate on source reconstructed activity. This framework is used to perform a connectivity analysis of a high density source reconstructed EEG dataset, which was constructed in an experiment regarding semantic processing of abstract and concrete words.

Thesis submitted for the degree of Master of Science in Artificial Intelligence, option Engineering and Computer Science

Thesis supervisor: Prof. dr. ir. Marc Van Hulle

Assessor: Mansoureh Fahimi
Prof. dr. Daniele Marinazzo

Mentor: Mansoureh Fahimi