Word Length Processing via Region-to-Region Connectivity

Mariya Toneva Machine Learning Department, Neural Computation Carnegie Mellon University

Data Analysis Project DAP committee: Tom Mitchell, Robert Kass, Avniel Ghuman

May 7, 2018

Abstract

A previous magnetoencephalography (MEG) study found that many features of stimuli can be decoded around the same time but at different places in the brain, posing the question of how information processing is coordinated among brain regions. We aim to directly relate the processing of information content in one region to the connectivity of that region. We examine whether, during presentation of a word stimulus, the length of the word relates to how strongly the region that best encodes word length left lateraloccipital cortex (LOC-lh) - connects to other regions, at different times.

Using representational similarity analysis, we find that changes in strengths of the LOC-lh connectivity network across stimuli correlate significantly with changes in word length. This significant correlation occurs at the very end of the decodability period of word length in LOC-lh between 200 - 220ms post stimulus onset. We further investigate whether any individual regions contribute significantly to this correlation. We find that there are only 3 regions which have connectivity strengths with LOC-lh that rank correlate significantly with word length during this time period. All of these regions are located in the left temporal lobe. Two of them – the left temporal pole (also known as anterior temporal lobe) and the left superiortemporal gyrus – have previously been implicated in language processing.

The high temporal resolution of MEG combined with the set of techniques we present can enable the tracking of information flow during language processing.

1 Introduction

A previous MEG study found that many features of stimuli can be decoded around the same time but at different places in the brain [Sudre et al., 2012], posing the question of how information processing is coordinated between brain regions.

Previous approaches to this question are of two types. The first uses a classifier or regression to uncover the relative timings of feature decodability in different brain regions. While addressing when and where information is processed, this approach does not specify how information is coordinated. The second type estimates when the connectivity between regions changes. While this approach assumes that information is coordinated by communication, it



Figure 1: Recreated from Sudre et al. [2012]. Experimental paradigm. While in the MEG, subjects are presented with a question, followed by 60 concrete nouns along with their line drawings in random order. A stimulus is presented until the subject presses a button in response to the question presented in the beginning of the block. There is 1s of fixation between nouns and questions.

does not directly relate to the information content. We aim to more directly relate processing of information content to connectivity.

We examine whether, during presentation of a word stimulus, the length of the word relates to how strongly the region that best encodes word length - left lateraloccipital cortex (LOC-lh) - connects to other regions, at different times.

We focus on word length because it is decodable earlier than other features at around 100ms post stimulus onset [Sudre et al., 2012]. This limits the contribution of other features to the processing activity near 100ms. We further investigate the connectivity of the region that best encodes word length - LOC-lh [Sudre et al., 2012].

2 Data

We analyze MEG data from an experiment in which 10 subjects were presented 60 concrete nouns along with their line drawings. All 60 stimuli were presented in randomized order with 20 repetitions (see Figure 1).

The data were first preprocessed using the Signal Space Separation method (SSS) [Taulu and Simola, 2006] in order to isolate the signal components that originate inside of the sensor



Figure 2: Recreated from Desikan et al. [2006]. Segmentation into 67 anatomical regions according to the Freesurfer 'aparc' notation.

array. This method was followed by its temporal extension (tSSS) to align the measurements of the head position before each block to a common space. The MEG signal was then filtered using a low-pass filter at 150Hz and notch filters at 60Hz and 120Hz to remove contributions from electrical line noise and other very high frequency noise. Next, the Signal Space Projection method was applied to remove eye blinks, residual movement, and other artifacts.

The MEG data for each stimulus were subsequently source localized using the average of the 20 MEG trials. The 3D model needed for the localization was constructed from the structural MRIs of each subject using Freesurfer. Using each subject's anatomical data, we further segmented the 67 regions of interest (ROIs) according to the Freesurfer 'aparc' annotation (see Figure 2). All ROIs were used in the analyses in this work. Finally, the Minimum Norm Estimates (MNE) method [Hamalainen and Ilmoniemi, 1994] was used to estimate the source brain activity from the recorded MEG data. The test sources were evenly distributed (5mm between neighboring sources, loose orientation constraint of 0.2). All analyses in this work were conducted using the magnitude of the estimated source vectors. Each source dipole was standardized to have a mean of 0 and standard deviation of 1 in the time period -250 to 1750ms relative to stimulus onset (at 0ms).

3 Methods

Our main goal is to more directly relate the word length information encoded in the left lateraloccipital (LOC-lh) to the connectivity between the LOC-lh and other brain regions. For this purpose, we need a way to estimate the following: 1. the connectivity of LOC-lh with other regions 2. the time course of word length processing in LOC-lh 3. the relationship between the connectivity network strengths during a stimulus presentation and the length of that stimulus.

3.1 Connectivity Network of the Left Lateraloccipital Cortex (LOClh)

We first average the activity in a region of interest (ROI) during a single stimulus presentation across all dipoles within this region. The activity in an ROI for a single stimulus presentation now corresponds to a one-dimensional time series.

Now we estimate the connectivity of LOC-lh to all other brain regions in a specific time window during a particular stimulus presentation. For each of the remaining 66 ROI, we correlate the same 50ms time window of activity between the one-dimensional activity in this ROI and the one-dimensional activity in LOC-lh. For example, if we consider the time window of 50 to 100ms after the onset of stimulus 'house', we now have a 66-dimensional vector, where each element describes the correlation strength between LOC-lh and the corresponding region during 50 to 100ms of presentation of 'house'. We refer to this 66-dimensional vector as the strengths of the connectivity network of LOC-lh during 50 - 100ms of 'house'.

We can also estimate how the connectivity strengths of LOC-lh to all other regions change across time by sliding this 50ms window across the time of stimulus presentation and recomputing the correlations as described above. Note that we do not consider any lags between the time series in these analyses.

3.2 Representational Similarity Analysis (RSA)

To address the last two necessities, we use Representational Similarity Analysis (RSA) [Kriegeskorte et al., 2008]. In general, RSA enables the comparison between models with different representations. The most common use of RSA in neuroscience is to understand what candidate model representation of a stimulus explains the activity in the brain during this stimulus. Comparing the brain representation of a stimulus (i.e. activity corresponding to this stimulus in a particular region of interest) to the candidate model's representation is often complicated by the fact that they have different dimensionality. For example, a word length model represents the stimulus 'house' as a single number '5', whereas the brain activity in LOC-lh during the presentation of 'house' has 100 dimensions. To address this difficulty, instead of considering how the representations of the same stimulus compare across different models, RSA compares how the distances between pairs of stimuli change across models. Computing the pairwise distances for a particular model's representation across all stimuli results in a representational dissimilarity matrix (RDM) for that model. The final step in RSA is to compute a measure of correlation (either Pearson or Spearman) between two model RDMs. The intuition is that if the representations of the same stimulus across dif-



Figure 3: Illustration of representational dissimilarity matrices across three models: brain activity in LOC-lh, word length, and connectivity network strengths of LOC-lh. A complete pairwise distance matrix is called a 'representational dissimilarity matrix' (RDM).

ferent models contain similar information, then the pairwise distances across models should also be similar (i.e. the corresponding RDMs should correlate significantly).

In our analyses, we consider three models of stimulus representation (see Figure 3). The first one represents a stimulus by the brain activity in LOC-lh in a certain time window during the presentation of that stimulus. We refer to the corresponding model RDM as the 'activity RDM'. The second model we consider represents a stimulus by its word length. We refer to the corresponding model RDM as the 'word length RDM'. The last model represents a stimulus by the strengths of the connectivity network of LOC-lh in a certain time window during the presentation of that stimulus. We refer to the corresponding model RDM as the 'connectivity RDM'. Note that there is one activity RDM and one connectivity RDM per time window.

We use both Euclidean and cosine distance as distance metrics to construct the activity and connectivity RDMs, as there is not a clear best one in the literature and we don't have a prior hypothesis as to whether the magnitude of the vectors (Euclidean distance) or the angle between them (cosine distance) is of more importance. Both the activity and connectivity RDMs are constructed on the individual subject level and then averaged over subjects for subsequent analyses.

We conduct the representational similarity analysis by correlating two RDMs using the Mantel test for distance matrices [Mantel, 1967]. The Mantel test computes the Pearson correlation between the original pairs of RDMs and compares this original correlation to that between RDMs with randomly permuted rows/columns. We correct the resulting p-value for multiple comparisons across time and models using the Benjamini-Hochberg-Yekutieli procedure for arbitrarily correlated tests [Benjamini and Yekutieli, 2001].

3.2.1 Time Course of Word Length Processing in LOC-lh

To characterize the time course of word length processing in LOC-lh, we compare the word length RDM with a series of activity RDMs corresponding to sliding windows of 10ms of LOC-lh activity starting from stimulus onset (0ms) and ending at 300ms with no overlap.

3.2.2 Relating Connectivity Network of LOC-lh to Word Length

To relate the strengths of the LOC-lh connectivity network during stimulus presentation to the word length of that stimulus, we compare the word length RDM to a series of connectivity RDMs. Each connectivity RDM corresponds to the 50ms time window where the correlations between the LOC-lh and all other regions were estimated, as described in Section 3.1.

3.3 Alternative Word Length Decoding

Instead of using RSA to characterize the time course of word length processing in LOClh, one can use a method similar to the one described by Sudre et al. [2012]. We train a ridge regression model to predict the word length of a stimulus from sliding windows of 10ms activity in LOC-lh during the stimulus presentation. Then, using leave-one-out cross validation, we can predict the word length for a held out stimulus across folds, and compute the percent of variance explained (POVE) in the model, which is defined as:

$$POVE_w = 1 - \frac{\sum_i (w_i - \hat{w}_i)^2}{\sum_i (w_i - \bar{w})^2}$$
(1)

, where w_i is the true word length of the stimulus held out in fold i, \hat{w}_i is the predicted word length for that stimulus, and \bar{w} is the mean word length across all stimuli. We refer to this POVE as the decodability of word length. We contrast the two methods of characterizing the time course of word length processing in LOC-lh.

4 Results

4.1 Word Length vs Activity in LOC-lh

Using the Mantel test for distance matrices, we correlate the word length RDM with the average subject LOC-lh activity RDM across time. Each LOC-lh activity RDM represents the stimuli using 10ms of concatenated activity across all dipoles within the LOC-lh during the presentation of that stimulus. We investigate both cosine and Euclidean distances as distance metrics in the activity RDM.

Changes in activity in LOC-lh across stimuli significantly correlate with changes in the stimuli word lengths for both distance metrics (see Figure 4). However, there is a difference between the timings and magnitudes of the resulting correlations. Activity RDMs constructed using Euclidean distance seem to correlate more strongly with the word length RDM, and the significant period is earlier than that of RDMs constructed using cosine distance. In addition, it appears that out of the two distance metrics, the Euclidean distance is the one that produces results most similar to those from ridge regression (see Figure 5).



Figure 4: Correlations between word length RDM and LOC-lh activity RDMs computed using both cosine and Euclidean distance metrics. Stars indicate statistically significant correlations at alpha=0.01 when FDR corrected for multiple comparisons over time and model.



Figure 5: Comparison of ridge regression performance to RSA when the activity RDM is computed using Euclidean distance. The percent of variance in word length explained by the ridge regression model tracks the RSA correlations between the word length RDM and the activity RDM constructed using Euclidean distance.



Figure 6: Correlations between word length RDM and connectivity RDMs compared through time against correlations between word length RDM and activity RDM. RDMs computed using two different distance metrics - (a) Euclidean distance and (b) cosine distance. Stars indicate statistically significant correlations at alpha=0.01 when FDR corrected for multiple comparisons over time and model.

4.2 Word Length vs Connectivity Network of LOC-lh

Using the Mantel test for distance matrices, we correlate the word length RDM with the mean subject LOC-lh connectivity RDM across time. Each LOC-lh connectivity RDM is created using correlations of 50ms time windows. We investigate both cosine and Euclidean distances as distance metrics in the connectivity RDMs.

Changes in strengths of LOC-lh connectivity network across stimuli correlate significantly with changes in word length (see Figure 6). This significant correlation occurs at the very end of the decodability period of word length in LOC-lh between 200 - 220ms post stimulus onset.

4.3 Connectivity of Specific Regions with LOC-lh

Finally, we investigate whether any individual regions contribute significantly to the correlation around the 200 - 220ms period of interest. For this analysis, we rank correlate each individual region's connectivity with LOC-lh across stimuli during a particular time window with the word lengths of the stimuli. The intuition behind this experiment is that if there is a region that has stronger (weaker) connections with LOC-lh during longer words, there will be significant positive (negative) rank correlations.

We find that there are only 3 regions whose connectivity strengths to LOC-lh rank correlate significantly with word length (see Figure 7). Furthermore, this significant rank correlation is negative, indicating that these three regions are more strongly connected with LOC-lh during the presentation of shorter words.

5 Discussion

We presented evidence that changes in activity in LOC-lh across stimuli correlate significantly with changes in the stimuli word lengths. This result shows that we can replicate the findings



Figure 7: Time courses of rank correlations between word length and connectivity of 3 regions with LOC-lh. Stars indicate statistically significant correlations at alpha=0.05 when FDR corrected for multiple comparisons over time and model. During the period of interest between 200 - 220ms, these regions are the only ones that have connectivities with LOC-lh that significantly rank correlate with word length.

of Sudre et al. [2012] using the presented RSA method. We also show that the choice of distance metric in computing the activity RDMs makes both a quantitative and qualitative difference. The results using the Euclidean distance metric best relate to the results obtained by ridge regression.

We also show that the changes in strengths of LOC-lh connectivity network across stimuli also correlate significantly with changes in word length. This significant correlation occurs at the very end of the decodability period of word length in LOC-lh between 200 - 220ms post stimulus onset. For this analysis, the choice of metric does not seem to make as much of a difference.

Finally, we investigate whether any individual regions contribute significantly to this correlation around the 200 – 220ms period of interest. We find that there are only 3 regions whose connectivity strengths to LOC-lh rank correlate significantly with word length. All of these regions are located in the left temporal lobe. Two of them – the left temporal pole (also known as anterior temporal lobe) and the left superior temporal gyrus – have previously been implicated in language processing. The temporal pole has been found to be critical for semantic memory [Binney et al., 2010], while the superior temporal gyrus contains Wernicke's area and has been linked to language comprehension [Dronkers et al., 2004].

6 Future work

In the future, we would like to address the current limitation of using stimuli that contain both words and line drawings by conducting similar analyses on a word-only experiment. It would also be valuable to contrast the current results with an analysis of pronounceable nonwords and false fonts in order to understand whether the connectivity with LOC-lh depends at all on semantics.

One very natural extension of the current analyses is to incorporate a more biologicallyinspired measure of connectivity, such as phase locking value. Such an approach would enable us to relate information processing to a more mechanistic understanding of the communication between regions.

Another important direction is to move beyond the LOC-lh as a seed region in our analyses, and apply similar techniques to examine the pairwise connectivities between all brain regions and their relationship to the length of the stimulus and other word features. Such an analysis would get us closer to our goal of mapping the flow of information across the brain during language processing.

References

- G. Sudre, D. Pomerleau, M. Paratucci, L. Wehbe, A. Fyshe, R. Salmelin, and T. Mitchell. Tracking neural coding of perceptual and semantic features of concrete nouns. *NeuroImage*, 62(1):451–463, 2012.
- S. Taulu and J. Simola. Spatiotemporal signal space separation method for rejecting nearby interference in meg measurements. *Physics in Medicine & Biology*, 51(7):1759, 2006.
- R.S. Desikan, F. Ségonne, B. Fischl, B.T. Quinn, B.C. Dickerson, D. Blacker, R.L. Buckner, B.T. Hyman A.M. Dale, R.P. Maguire, and M.S. Albert. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage*, 31(3):968–980, 2006.
- M. Hamalainen and R. Ilmoniemi. Interpreting magnetic fields of the brain: minimum norm estimates. Medical & biological engineering & computing, 32(1):35–42, 1994.

- N. Kriegeskorte, M. Mur, and P. Bandettini. Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 2:4, 2008.
- N. Mantel. The detection of disease clustering and a generalized regression approach. *Cancer research*, 27(2):209–220, 1967.
- Y. Benjamini and D. Yekutieli. The control of the false discovery rate in multiple testing under dependency. Annals of statistics, pages 1165–1188, 2001.
- R. J. Binney, K. V. Embleton, E. Jefferies, G. J. Parker, and M. A. Lambon Ralph. The ventral and inferolateral aspects of the anterior temporal lobe are crucial in semantic memory: evidence from a novel direct comparison of distortion-corrected fmri, rtms, and semantic dementia. *Cerebral Cortex*, 20(11):2728–2738, 2010.
- N. F. Dronkers, D. P. Wilkins, and B. B. Redfern R. D. Van Valin Jr, and J. J. Jaeger. Lesion analysis of the brain areas involved in language comprehension. *Cognition*, 92(1):145–177, 2004.