ISBFC

Dynamics of Semantic Representations in the Human Brain

Hiromu SAKAI[1,2]

1. Faculty of Science and Engineering, Waseda University, Tokyo, Japan.
2. Waseda Research Institute for Science and Engineering, Waseda University, Tokyo, Japan.

E-mail: *hsakai@waseda.jp*

Abstract: Neural decoding using MEG data reveals dynamics of semantic representations in the human brain. We first examine when perception of stimuli with different types of modalities leads to shared semantic representations of modality-independent concepts. We then reveal how contexts modulate the semantic representations of objects in the high-dimensional feature space.

Keywords: Context, Modality, Picture Naming, Word Read-Aloud, Neural Decoding

Introduction

Research on neural mechanisms of language and concepts has focused mainly on identifying the location of the "language network" in the human brain since its starting points as studies of aphasia in the 19th century. From the beginning of the 21st century, researchers have shifted their attention to how linguistic information is encoded in a broader range of cortical regions. Following such trends, we report recent research in our laboratory on how neural codes for language are revealed by so-called multivariate pattern analysis of neural activities measured by non-invasive brain function measurement techniques. We demonstrate how magnetoencephalography (MEG), a non-invasive technology equipped with special resolution and temporal resolution, can reveal temporal dynamics and contextual variability of semantic representations.

In the first study, we examined when and how modality-independent concepts are represented as multivariate patterns. When we see a cat image or read a written word "cat", our intuition tells us that a mental image of a cat are represented in our brain. Is this intuition supported by neurological evidence, and, if it is supported, what is the time-course of emergence of such "modality-independent" representation in the human brain? A cross-decoding analysis using MEG data was conducted to shed new light on this question[1, 2]. We train a machine learning classifier model using the neural data obtained when participants performed either picture-naming or word read-aloud tasks. After training, the model's performance is tested upon the data of the same task (within modality) or different (across modality). The above-chance level accuracy of cross-modal decoding indicates the existence of modality-independent representations of semantic information within the human brain.

In the second study, we point out that modulation of semantic representations by the presentation context is reflected in the high-dimensional feature space. Even if we saw the same image of a dog, it would evoke significantly different reactions depending on the situation and context. Semantic representations in the human brain should therefore be sensitive to the presentation contexts of objects. In this study, using a blocked-cycle picture naming paradigm[3, 4], we manipulated the presentation context of objects and examined how semantic representations were modulated through decoding analysis of MEG signals.

Methods

Study 1: Eleven participants were recruited. Colored photo images and black-and-white line drawing pictures were prepared for 12 objects selected from 4 categories: animal, plant, food, and tool. An equal number of word stimuli was prepared using two different types of hand-written fonts. Participants were asked either (i) to name the picture orally or (ii) to read aloud the words presented at the center of the monitor screen. MEG signals were recorded when participants were conducting these two tasks using a superconducting self-shield MEG (Sumitomo Heavy Industry Inc.) [5]. Pre-processing and data analysis were performed using the MNE-Python software package. Decoding analysis was conducted using the scikit-learn Python machine learning library. Decoding of categories was conducted using logistic regression models as a classifier and normalized magnetic field amplitude measured by each sensor as features. Accuracies were calculated at every combination of time points during the trial epochs. A 5-fold cross-validation was conducted. Within modality decoding analysis, both training data and test data were selected from the data measured from the same task, whereas across modality decoding, training data and test data were selected from those from different tasks.

Study 2: Seventeen participants were recruited through the participant recruitment system at the Center for Information and Neural Network (Osaka, Japan). Colored photo images and black-and-white line drawing pictures were prepared for 5 objects selected from 5 categories: animal, plant, food, tool, and clothes. Participants were asked to name the picture presented in the monitor screen orally. In homogeneous (HOM) blocks, objects from the same category were presented. In heterogeneous (HET) blocks, objects selected from each of the five categories were presented. Each object was presented 30 times during the blocks. The MEG signals were recorded using MEG (Electa Neuromag360, MEGIN Oy) with 120 magnetometers and 240 gradiometers with a sampling frequency of 1000Hz. Auditory responses were recorded by DR-100MKII PCM recorder (TASCOM). MEG data were corrected for head movements and external interferences by the Signal Space Separation (SSS) method implemented in MaxFilter software. Further pre-processing was conducted using the MNE-Python software package.

Decoding analysis was conducted using the scikit-learn Python machine learning library. Pair-wise decoding of 25 objects was performed using the support vector machine as a classifier over normalized magnetic field amplitude measured by each sensor as features. Latencies of oral responses were detected using a custom MATLAB program after manual discarding of faulty trials.

Results

Study 1: Within-modality decoding shows that significant decoding accuracies are achieved after 300 ms in picture naming and after 500 ms in word read-aloud tasks. In across-modality decoding, accuracies reached significance only when picture data were used for training and word data were used for testing. It is also revealed that significant decoding accuracies were achieved by using the picture data obtained around 700 ms to 800 ms and the word data obtained around 600 ms.

Study 2: Response times were significantly faster in the HET blocks compared to those in the HOM blocks, showing an interference effect reported in the previous literature [3, 4]. Significant correlation between response latencies and decoding accuracies were observed in either of the blocks. That is, higher decoding accuracies were correlated with faster response times. Decoding accuracies within each of the five categories showed a significant contrast between the HOM blocks and the HET blocks, whereas no contrast was observed across the five categories. To put it differently, within-category decoding becomes difficult in the HOM blocks compared to the HET blocks, whereas no such contrast was observed in across-category decoding.

Discussion

Study 1: The result of within-modality decoding suggests that the visual information of objects is represented in earlier timing compared to the linguistic information of words in the human brain. The result of cross-modality decoding shows that modality-independent semantic representation was reached about 100 ms earlier from word stimuli than from picture stimuli.

Study 2: The results indicate that semantic representations within a high-dimensional feature space were modulated by the presentation context. The observation that the decoding accuracies were significantly low in the HOM blocks shows that representations of objects in the HOM blocks must be less separable by the SVM classifiers. This is becasue the features shared among the members of the same categories become more salient in the HOM blocks as a result of repeated presentation of objects in the same categories. This makes semantic representations less separable from each other in the HOM blocks. The nature of interference effects can thus be derived from contextual dynamics of semantic representations.

Conclusions

Putting the results of two studies together, we conclude that MPVA of MEG data provides a promising approach toward the goal of revealing how and when semantic representations emerge dynamically in the human brain. The development of new methodologies will open up a new perspective in cognitive neuroscience of language in the 21st century.

Acknowledgements

This research was conducted as a part of project research at Waseda Institute for and Engineering Research and supported by JSPS Grant-in-Aid for Scientific Research (S) [PI: Hiromu SAKAI; #23H05493], JSPS Grant-in-Aid for Challenging Research (Exploratory) [PI: Hiromu SAKAI; #23K17272], and JSPS Grant-in-Aid for Specially Promoted Research [PI: Reiko Mazuka; #20H05617].

References

[1] E. Leonardelli, E. Fait, and S. L. Fairhall, "Temporal dynamics of access to amodal representations of category-level conceptual information," *Scientific Reports*, 9(1), 239, Jan. 2019, doi: 10.1038/s41598-018-37429-2.

[2] J. Dirani and L. Pylkkänen, "The time course of cross-modal representations of conceptual categories," *NeuroImage*, 277, Aug. 2023, 120254, doi: 10.1016/j.neuroimage.2023.120254.

[3] J. F. Kroll and E. J. Stewart, "Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations," *Memory and Language*, 33(2), pp. 149-174, Apr.1994, doi: 10.1006/jmla.1994.1008.

[4] B. Maess et al., "Semantic category interference in overt picture naming: Sharpening current density localization by PCA," *Journal of Cognitive Neuroscience*, 14(3), pp. 455-462, Apr. 2002, doi: 10.1162/089892902317361967.

[5] K. Tanaka et al., "Superconducting Self-Shielded and Zero-Boil-Off Magnetoencephalogram Systems: A Dry Phantom Evaluation," Sensors, 24(18), 6044, Sep. 2024, doi: 10.3390/s24186044.