Map-like Representations of Lexical Features during Movie Watching

Siyang Li (siyali@zhejianglab.com)
Zhejiang Lab, Kechuang Road
Hangzhou, Zhejiang 311100 China

Huan Zhang (zhanghuanzh@zhejianglab.com)
Zhejiang Lab, Kechuang Road
Hangzhou, Zhejiang 311100 China

***Xia Liang (xia.liang@hit.edu.cn)**
Laboratory for Space Environment and Physical Sciences,
Harbin Institute of Technology
Harbin, Heilongjiang 150001, China

***Yu Zhang (yuzhang@zhejianglab.com)**
Zhejiang Lab, Kechuang Road
Hangzhou, Zhejiang 311100 China

***Tianzi Jiang (jiangtz@nlpr.ia.ac.cn)**
Institute of Automation, Chinese Academy of Sciences
Beijing 100190 China.

Abstract:

Growing focus has converged on how the brain responses to naturalistic stimuli, yet the underlying neural mechanisms remain poorly understood. Cognitive map theory proposes that human memory encodes external information internally by establishing relationships between experience fragments. In this study, we investigated brain's cognitive map-like representations during movie watching and their alignment with the lexical characteristics of movie dialogue. Using functional magnetic resonance imaging (fMRI), we uncovered two brain states with dynamic alternations between internal (*State 1*) and external (*State 2*) focuses during movie watching. Specifically, *State 1* was characterized by more globally organized internal representations and exhibited continuous learning of lexical features, mirroring the modular structures of lexical network. In contrast, *State 2* was primarily engaged in the sensory processing of external stimuli. Using successor representation (SR) technique, we recognized typical place fields in brain state spaces, and demonstrated a significant association between the predictive skewness of the SR field and individual vocabulary comprehension capacity. Our findings offer novel insights into the neural mechanisms underlying the processing of naturalistic stimuli in the human brain, and support to the involvement of cognitive map theory in lexical feature encoding.

Keywords: naturalistic stimuli; cognitive map; fMRI; brain state; lexical feature; predictive representation

Introduction

Understanding the neural substrates of human cognition is a central goal of neuroscience research. With the aid of modern imaging techniques, such as functional magnetic resonance imaging (fMRI), neuroscientists have shown growing interests on how the brain processes naturalistic stimuli such as movie watching (Kringelbach et al., 2023). Research has demonstrated that the brain segments continuous streams of stimuli into discrete events and constructs narrative graphs linking these segments, which consequently influence memory performance and behavior (Lee & Chen, 2022). Yet, the precise characterization and neural representations of such naturalistic information remain unclear.

According to the theory that brain encodes and organizes experiences into relational maps (Behrens et al., 2018; Tolman, 1948), our study aims to investigate cognitive map-like representations in the human brain during movie watching. Firstly, we identified two brain states with distinct spatiotemporal representations during movie watching. Secondly, we demonstrated map-like representations of *State 1* aligned with the lexical features of movie dialogue and significantly predicted vocabulary comprehension capacities.

Materials and Methods

We adopted 7T movie-watching fMRI data consisting of 184 subjects (age: 22-35; 112 females) acquired from the Human Connectome Project (HCP) dataset (Elam et al., 2021). The minimal preprocessed fMRI data were downloaded and used in the following analysis.

Construction of brain-state spaces on cognitive-map networks

We first identified different brain states that showed distinct spatiotemporal patterns during movie watching (Figure 1). A similar approach has been proposed for resting-state fMRI in our recent paper (Li et al., 2024). Initially, for each brain state, we calculated the similarity matrix (W) of hippocampal activity between identified occurrences (a series of short sequences). Next, we mapped brain-state occurrences into a two-dimensional (2D) space by using t-SNE, and clustered the occurrences using k-means. After that, we calculated transitions between occurrence clusters and explored representation organization of each brain state.

Estimation of lexical features

We extracted the lexical features of state occurrences by applying wave2vec to the text of movie dialogue, and generated a lexical structure matrix (S) by calculating Pearson correlation between lexical vectors (Figure 1). Next, we evaluated the association of the lexical similarity matrix S and the neural similarity matrix W, and specifically investigated the lexical similarity within and between clusters of state occurrences.

Cognitive map-like representation

We constructed a successor representation (SR) matrix (M) for each brain state, with each column representing the place field (PF) of corresponding state occurrence (Li et al., 2024; Stachenfeld et al., 2017). We then

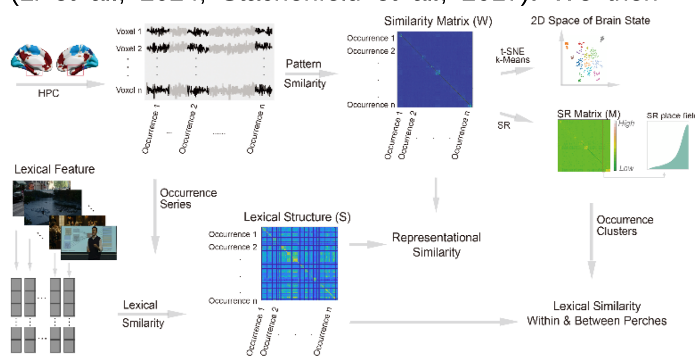

Figure 1: Analytical flowchart of the study.

mapped SR columns of occurrence clusters into the 2D brain-state space and detected typical place fields. Subsequently, we estimated the backward skewness of the SR fields, which has been shown to reveal the predictivity and preference of representation organization (Stachenfeld et al., 2017). Furthermore, we examined the relationship between predictive skewness of SR fields and -participants' language skills.

Results

Spatiotemporal patterns of brain states during movie watching

We identified two brain states using fMRI data during movie watching (Figure 2A). *State 1* predominantly involved default more network (DMN) and hippocampus, while *State 2* exhibited strong activation in sensorimotor regions and hippocampus. By estimating the coherence of hippocampal activity (W) within each state, we uncovered that *State 1* exhibited significantly higher variability or lower similarity in neural representations over the course of movie watching ($t=8.85$, $p<10^{-15}$). By mapping brain-state occurrences into a 2D state space, we found that the organization of occurrence clusters in *State 2* more closely followed the proximity principle, i.e., cluster hubs, as indicated by larger nodal size in Figure 2A, positioned closer together in the brain-state space (a significant interaction between 'distance between clusters' and 'state' as detected in the two-way ANOVA, $F(695,1)=13.6$, $p<10^{-3}$).

Lexical features of brain state representation

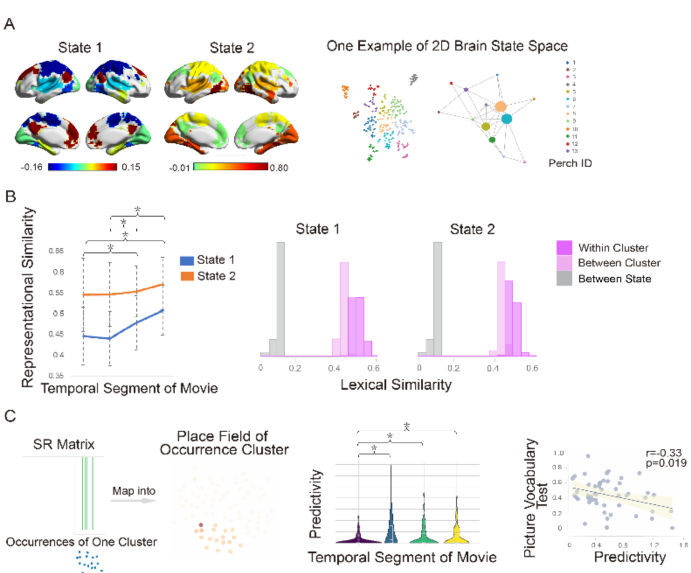

Figure 2: Representational and lexical features of brain state space during movie watching.

Compared to *State 2*, we observed significant lower association between the representational similarity of lexical features and hippocampal activity in *State 1* ($t=-18.93$, $p<10^{-23}$). Notably, the representational similarity of *State 1* gradually enhanced over the course of movie watching (Figure 2B left), exhibiting a significant interaction effect between 'movie segment' and 'brain state' as detected in the two-way ANOVA ($F(1143,3)=6.02$, $p<10^{-3}$). Moreover, we observed higher lexical similarity within occurrence clusters than between clusters for both brain states (*State 1*: $t=39.19$, $p>10^{-30}$; *State 2*: $t=-51.72$, $p<10^{-30}$), with a more pronounced effect in *State 1* (Figure 2B right).

Cognitive map-like representations in brain state space

We identified typical 'place fields' of occurrence clusters in the brain-state space using SR analysis (Figure 2C left), suggesting a cognitive map-like representation in brain state space. We assessed the predictive skewness (i.e., predictivity) of SR fields and observed a significant increase in the predictivity of *State 1* ($F(707, 3)=4.7$, $p=0.003$; post-hoc t test, $ps<0.05$). Besides, the predictivity of *State 1* also demonstrated a significant correlation with behavioral performance on the Picture Vocabulary Test (Figure 2C right).

Conclusion

The present study identified two brain states associated with internal and external focus during movie watching. *State 2* responds to immediate sensory processing of external lexical features, while *State 1* integrates abstract relational representations of movie segments into internal knowledge maps. Specifically, *State 1* was dominated by DMN and hippocampus, and exhibited a longer temporal receptive window compared to *State 2*, may contribute to the formation of long-term memory. Over the course of movie watching, the cognitive map-like representations of *State 1* gradually converged on lexical features, may support the continuous learning of complex structures and knowledge. Using the SR approach, we revealed typical 'place fields' patterns in both brain state spaces, mirroring 'modules' in the lexical structure of movie dialogues. Moreover, the predictivity of SR place fields was significantly associated with participants' vocabulary skills and verbal comprehension. Our results provide supporting evidences for cognitive map theory in the human brain, such that the internal neural representations of lexical features are organized into a cognitive map during movie watching.

Acknowledgments

This work was partially supported by the STI2030-Major Projects 2021ZD0200200, 2022ZD0211500, the National Natural Science Foundation of China (Grant Nos. 62201519, 52307259, 62327805, 82151307, 82072000, 81671769), the Fundamental Research Funds for the Central Universities (HIT.OCEF.2023015), the Natural Science Foundation of Heilongjiang Province, China (YQ2023H016).

References

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. Neuron, 100(2), 490-509.

Elam, J. S., Glasser, M. F., Harms, M. P., Sotiropoulos, S. N., Andersson, J. L. R., Burgess, G. C., Curtiss, S. W., Oostenveld, R., Larson-Prior, L. J., Schoffelen, J. M., Hodge, M. R., Cler, E. A., Marcus, D. M., Barch, D. M., Yacoub, E., Smith, S. M., Ugurbil, K., & Van Essen, D. C. (2021). The Human Connectome Project: A retrospective. Neuroimage, 244, 118543.

Kringelbach, M. L., Perl, Y. S., Tagliazucchi, E., & Deco, G. (2023). Toward naturalistic neuroscience: Mechanisms underlying the flattening of brain hierarchy in movie-watching compared to rest and task. Science Advances, 9(2), eade6049.

Lee, H., & Chen, J. (2022). Predicting memory from the network structure of naturalistic events. Nature Communications, 13(1), 4235.

Li, S., Li, Z., Liu, Q., Ren, P., Sun, L., Cui, Z., & Liang, X. (2024). Predictable navigation through spontaneous brain states with cognitive-map-like representations. Prog Neurobiol, 233, 102570.

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. Nat Neurosci, 20(11), 1643-1653.

Tolman, E. C. (1948). Cognitive maps in rats and men. Psychological Review, 55(4), 189-208.