Is argument structure processing syntactic or semantic in nature? Evidence from computational modeling on naturalistic fMRI

Tianze Xu[1, 2], Jixing Li[3], Xiaoming Jiang[2]
[1] University of Zurich, [2] Shanghai International Studies University, [3] City University of Hong Kong
tianze.xu@uzh.ch

Background: A long-standing theoretical debate exists in linguistics concerning argument structure processing, with separationism focusing on syntactic structure and projectionism on semantic properties[1]. To investigate whether argument structure processing is primarily influenced by syntactic structure or semantic properties, this study employed integrative neurocomputational modeling[2, 3] to link brain functions with explicitly defined computational models.

Method: We analyzed naturalistic functional magnetic resonance imaging (fMRI) data from participants listening to a story[4], with a focus on *subject noun phrase + verb* chunks. The methodological framework integrated a general linear model (GLM) analysis of the fMRI data with computational modeling using natural language processing algorithms. These components were integrated using representational similarity analysis (RSA), allowing us to assess the relatedness of two symbolic computational models—one relying on syntactic information[5] from parse trees and the other based on semantic selectional preference information of verbs[6]—to brain activities.

Results: The GLM analysis identified significant neural correlates of argument structure processing largely consistent with previous findings, including the precuneus, the right superior temporal gyrus, and the right middle temporal gyrus. Some deviations from previous studies likely reflect the naturalistic nature of the stimuli and our contrast design. The RSA results favored the model utilizing semantic information—a finding further supported by effects observed in brain regions associated with argument structure processing in the literature and by an additional RSA comparing constructions with varying levels of transitivity.

Discussion: These findings suggest that during naturalistic story listening, humans rely heavily on semantic information to interpret argument structure. This study demonstrates an alternative method to engage with the debate on argument structure, highlighting a collaborative effort between theoretical neuroscientific, and computational linguistics.

Keywords: Argument structure processing; Computational modeling; Naturalistic fMRI; Syntax; Semantics

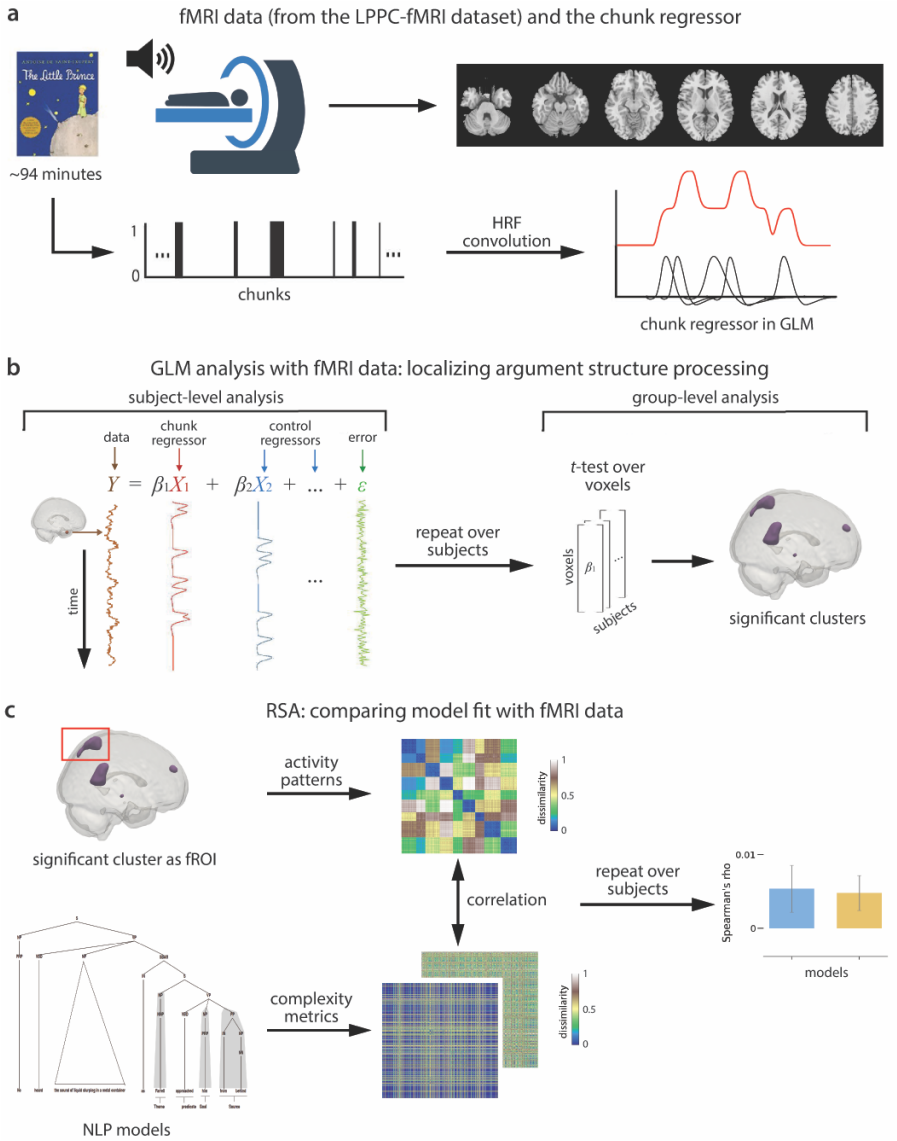

Figure 1. Overview of the analysis pipeline. (**a**) Acquisition of the fMRI data (from the LPPC-fMRI dataset[4]) and creation of the chunk regressor for GLM. (**b**) Localization of brain activity for argument structure assignment using GLM to identify fROIs, following established protocols[3]. (**c**) Comparison of NLP model fit with fMRI activity patterns within each fROI using RSA. Abbreviations: LPPC-fMRI, *Le Petit Prince* fMRI Corpus; GLM, general linear model; RSA, representational similarity analysis; fROI, functional regions of interest; NLP, natural language processing.

References

[1] Williams, A. (2015). *Arguments in syntax and semantics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511610479

[2] Hale, J. T., Campanelli, L., Li, J., Bhattasali, S., Pallier, C., & Brennan, J. R. (2022). Neurocomputational models of language processing. *Annual Review of Linguistics*, *8*, 427–446. https://doi.org/10.1146/annurev-linguistics-051421-020803

[3] Li, J., Wang, S., Luh, W.-M., Pylkkänen, L., Yang, Y., & Hale, J. (2021). Cortical processing of reference in language revealed by computational models. bioRxiv. https://doi.org/10.1101/2020.11.24.396598

[4] Li, J., Hale, J., & Pallier, C. (2024). *Le Petit Prince: A multilingual fMRI corpus using ecological stimuli* (Version 2.0.5) [Data set]. OpenNeuro. https://doi.org/10.18112/openneuro.ds003643.v2.0.0

[5] Gildea, D., & Jurafsky, D. (2002). Automatic labeling of semantic roles. *Computational Linguistics*, *28*(3), 245–288. https://doi.org/10.1162/089120102760275983

[6] Zapirain, B., Agirre, E., Màrquez, L., & Surdeanu, M. (2013). Selectional preferences for semantic role classification. *Computational Linguistics*, *39*(3), 631–663. https://doi.org/10.1162/COLI_a_00145