Diathesis Alternations and Selectional Restrictions in Sentence Processing: A fMRI Study

Keywords = Brain and language; syntax; semantics; argument structure; language processing; English

Introduction: Sentence processing is more than decoding linear strings of words. There are hierarchical structures and relationships which affect language comprehension and verbal argument structure is one such example. As Bresnan (1995) states, "argument structure is an interface between the semantics and syntax of predicators (which we may take to be verbs in the general sense). Its function is to link lexical semantics to syntactic structures". Diathesis alternations and selectional restrictions can be used to represent the syntax-semantics interface of argument structure. Diathesis alternations occur when verbs can occur with different subcategorization frames. Selectional restrictions are the semantic restrictions that a word imposes on the environment in which it occurs.

In this neuroimaging study, the goal is to investigate how diathesis alternations and selectional restrictions on a verb and its object plays a role in sentence processing using computational metrics and fMRI data. Specifically, we want to investigate if these two different components of argument structure have differing neural bases during language comprehension.

Methods: Participants (n=51, 32 female) listened to *The Little Prince*'s audiobook for 1h38min. Participants' comprehension was confirmed through multiple-choice questions (90% accuracy, SD = 3.7%). Functional scans were acquired using multi-echo planar imaging sequence (ME-EPI) (TR=2000ms; TE's=12.8, 27.5, 43ms; FA=77 degrees; FOV=240.0mm X 240.0mm; 2X image acceleration; 33 axial slices, voxel-size 3.75 x 3.75 x 3.8mm). Preprocessing was done with AFNI16 and ME-ICAv3.2.

2948 verbs in total were identified using the NLTK toolkit and Stanford POS tagger. Excluding modals, auxiliaries, and gerunds, there are 1970 verbs attested in the story (401 unique). Diathesis alternations for a given verb was calculated from PropBank (Kingsbury, 2002), which consists of all the sentences from the Penn Treebank annotated with subcategorization frames with higher scores indicating more subcategorization frames. Selectional restriction was calculated according to Resnik (1996) by estimating verb-direct object pairs from the Gigaword (Ferraro et al., 2014) & WaCkypedia (Baroni et al., 2009) corpora and then calculating the number of different WordNet semantic classes a given verb's direct objects falls into. Higher selectional restriction scores indicate the verb is more particular about the kinds of arguments it takes as its direct object. While both of these metrics operationalize different aspects of verbal argument structure within a sentence, they also formalize a degree of constraint in terms of sentence processing. The PropBank scores reflect the degree of constraint in terms of syntactic structure and subcategorization frames, whereas the Resnik's scores reflect the degree of constraint in terms of selectional restrictions and the semantic categories. Both of these gradient measures are thus taken as indices of degrees of constraint and correlated with brain activity.

We regressed the predictors described above against fMRI timecourses recorded during passive story-listening in a whole-brain analysis. Along with these two regressors, we entered four regressors of non-interest into the GLM analysis (SPM12): time stamp of each word offset, the log-frequency of each word in movie subtitles (Brysbaert & New, 2009), the pitch (f0) and intensity (rms) of the narrator's voice. The whole-brain main effects were FWE-corrected (T-score>5.3).

Results: The largest clusters for the PropBank scores (formalizing diathesis alternations) were observed in the right Supramarginal Gyrus and Middle Frontal Gyrus and bilateral Precuneus (Fig. 1A). The largest clusters for the Resnik's scores (formalizing selectional restrictions) were observed in the right Superior Temporal Gyrus, Inferior Frontal Gyrus, and Supplementary Motor Area (Fig. 1B).

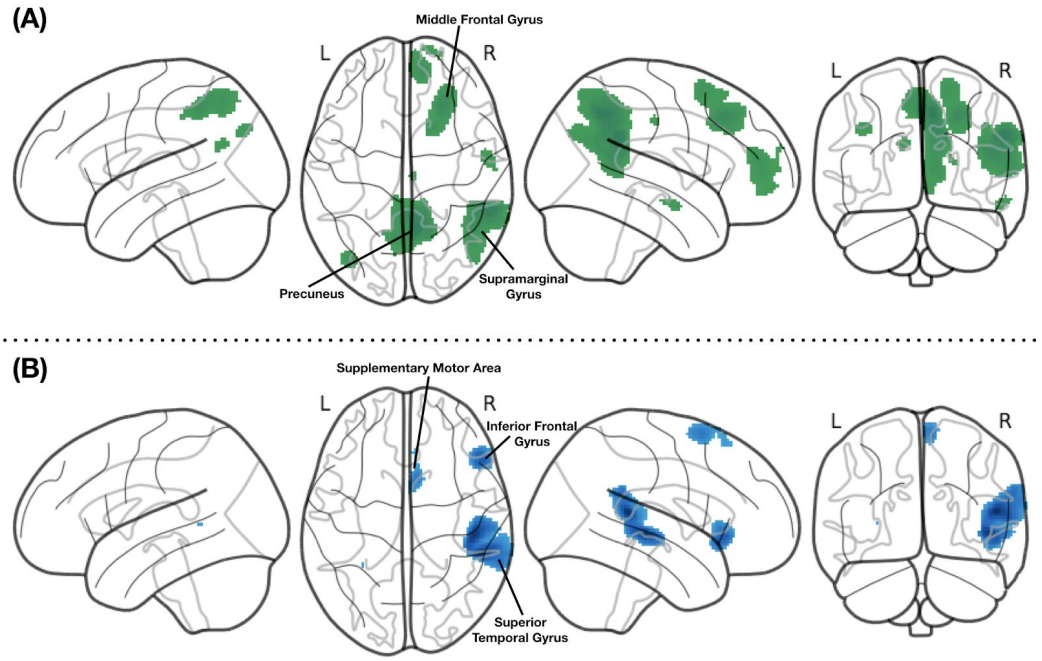

Figure 1 (A): Whole-brain contrasts for PropBank scores in green (B): Whole brain contrasts for Resnik's score in blue

Conclusion: The results for diathesis alternations corroborate previous neuroimaging studies related to subcategorization (Shetreet et al., 2006; Thompson et al., 2010; Thompson et al., 2007; Meltzer-Asscher et al., 2013). However, this study differs in that the neural bases of subcategorization was investigated in an ecologically valid setting within a naturalistic language comprehension study. The results for selectional restrictions are consistent with other neuroimaging studies related to lexical-semantic processing (Kuperberg et al., 2000; Baker et al., 2001; Zempleni et al., 2007). Based on these results, we can see different patterns of activation for syntactic and semantic subprocesses respectively. Overall, this study sheds light on the brain areas involved in argument structure and its syntax-semantics interface during sentence processing.

Selected References

Bresnan, J. (1995). Lexicality and argument structure. In *Paris Syntax and Semantics Conference* (Vol. 12).

Kingsbury, P., & Palmer, M. (2002, May). From TreeBank to PropBank. In *LREC* (pp. 1989-1993).

Kuperberg, G. R., McGuire, P. K., Bullmore, E. T., Brammer, M. J., Rabe-Hesketh, S., Wright, I. C., & David, A. S. (2000). Common and distinct neural substrates for pragmatic, semantic, and syntactic processing of spoken sentences: an fMRI study. *Journal of Cognitive Neuroscience*, *12*(2), 321-341.

Meltzer-Asscher, A., Schuchard, J., den Ouden, D. B., & Thompson, C. K. (2013). The neural substrates of complex argument structure representations: Processing "alternating transitivity" verbs. *Language and cognitive processes*, *28*(8), 1154-1168.

Resnik, P. (1996). Selectional constraints: An information-theoretic model and its computational realization. *Cognition*, *61*, 127-159.

Shetreet, E., Palti, D., Friedmann, N., & Hadar, U. (2006). Cortical representation of verb processing in sentence comprehension: Number of complements, subcategorization, and thematic frames. *Cerebral Cortex*, *17*(8), 1958-1969.

Thompson, C. K., Bonakdarpour, B., Fix, S. C., Blumenfeld, H. K., Parrish, T. B., Gitelman, D. R., & Mesulam, M. M. (2007). Neural correlates of verb argument structure processing. *Journal of Cognitive Neuroscience*, *19*(11), 1753-1767.

Zempleni, M. Z., Renken, R., Hoeks, J. C., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *Neuroimage*, *34*(3), 1270-1279.