A usage-based constructionist approach to grammar: Semantic network analyses of words, constructions, and alternations

Alvin Cheng-Hsien Chen[a] and Hung-Kuan Su[b]
[a,b]National Taiwan Normal University
[a]alvinchen@ntnu.edu.tw [b]jacksugood@gmail.com

Keywords: network analysis, usage-based, construction grammar, contingency learning

Grammatical competence features a development of native intuition in "predicting the linguistic constructions that are most likely to be relevant in the ongoing discourse context" (Gries & Ellis, 2015, p. 236). Choosing a proper construction from a set of notionally equivalent constructions is essential to one's grammatical competence. This study provides a case study on the functional variations of three near-synonymous space particle constructions (SPC) encoding CONTAINMENT in Chinese: [*zai* NP *li/nei/zhong*] and presents semantic network analyses on the inter-relationships between these partially schematic constructions (i.e., SPC) and the words occupying the NP position of the SPC (i.e., their co-occurring landmark, LM). We first extracted all relevant SPC tokens from Academia Sinica Balanced Corpus of Mandarin Chinese. We determined two types of relationships: (1) LM-SPC, and (2) LM-LM. The former specifies the faithfulness of the contingency between a word serving as the LM of the SPC and the SPC; the latter specifies the semantic similarity between words occupying the LM position of the SPC. The LM-SPC association was identified using the "distinctiveness" of the distinctive collexeme analyses (Gilquin, 2006; Gries & Stefanowitsch, 2004); the LM-LM association was determined based on the cosine similarity of the effective statistical learning of the word-embedding modeling (Pennington, Socher, & Manning, 2014). We utilized these metrics to build semantic networks in which the nodes represent symbolic units (e.g., LM words and SPC constructions) and the edges strong LM-SPC and LM-LM associations. We ask three questions: (1) For each LM, how prototypical is it of the meaning of the SPC? (2) For each SPC, how semantically cohesive are its LM exemplars? (3) What are the functional differences of these three SPCs? Network algorithms of betweenness centrality and detection of communities were used to find the patterns underlying the semantic networks. Our results suggest that LI is a more unmarked SPC in encoding CONTAINMENT, co-occurring with more heterogenous LMs. NEI shows a strong preference for LMs denoting temporal concepts. This metaphorical use often implies a preplanned objective in the proposition, with the LM as an intended deadline. Finally, ZHONG shows a strong connection to LMs denoting high-dynamicity events. This extended use often comes with a marked aspectual reading of the LM. Our network analyses bring to the foreground the importance of repeated language experiences in the shaping and entrenchment of linguistic knowledge. Mental grammar represents rich implicit knowledge of the contingency learning of these distributions in language use (Ellis & Ogden, 2017).

References

Ellis, N. C., & Ogden, D. C. (2017). Thinking about multiword constructions: Usage‐based approaches to acquisition and processing. Topics in Cognitive Science, 9, 604-620. doi:10.1111/tops.12256

Gilquin, G. (2006). The verb slot in causative constructions: Finding the best fit. Constructions, S1, 1-46.

Gries, S. T., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective on 'alternations'. International Journal of Corpus Linguistics, 9, 97-129. doi:10.1075/ijcl.9.1.06gri

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global vectors for word representation. Proceedings of Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), 1532-1543.