Abs-231

Matthew Brook O'Donnell, Ute Römer, and Nick C. Ellis (all University of Michigan)

Exploring Zipfian distributions in English verb argument constructions: Corpus and psycholinguistic evidence

Faced with novel utterances such as "the ball MANDOOLS across the ground" and "the teacher SPUGGED the boy the book", how are we able to arrive at similar interpretations? You know that MANDOOL is a verb of motion and have some idea of how MANDOOLING works and similarly that SPUGGING involves some sort of transfer. The larger configuration of words (the construction) has come to carry meaning as a whole (Goldberg, 2003, 2006). MANDOOL inherits its interpretation from the echoes of the verbs that occupy this Verb Argument Construction (VAC) -- words like 'come', 'walk', 'move', ..., 'scud', 'skitter' and 'flit'. Small-scale studies of a handful of constructions in both first- (Goldberg, 2006) and second- (Ellis & Ferreira-Junior, 2009) language acquisition have argued that the Zipfian distributions (Zipf, 1949) of verbs in VACs helps make them learnable. The aim of our research is to explore these factors for VAC acquisition in a larger number of VACs in both a large-scale corpus analysis and through psycholinguistic experiments.

In the corpus analysis, we are building an inventory of specific VACs identified in the COBUILD Grammar Patterns volume (Francis, Hunston, & Manning, 1996) using the BNC. We define searches against part-of-speech categories and dependency relations produced by three different parsers (RASP [Briscoe et al. 2006], Stanford [de Marneffe et al 2006] and C&C [Curran et al. 2007]) and take a consensus of whether a sentence matches the VAC pattern. Then for each VAC, such as V across n, we record the distribution of the verb types, their token frequencies and sentence contexts. We determine the degree to which the distributions are Zipfian (e.g. come 474 ... spread 146 ... throw 17 ... stagger 5). Statistical analyses (MI, Delta-P, Chi-Square) examine the associations between verbs and constructions (e.g. scud, skitter, sprawl, flit have the strongest association with V across n). WordNet and other semantic resources are used to measure the cohesion of the types in each distribution (e.g., semantic fields TRAVEL and MOVE most frequent for V across n). These data allow us to make predictions regarding language users' knowledge of verbs in constructions.

In psycholinguistic experiments we use free association tasks to test these predictions. We have native and non-native speakers of English think of the first word that comes to mind to fill the V slot in a particular VAC frame. The range of the verbs that they generate, and their speed of access, inform us about the representation of these VACs in the human mind. For each VAC, we compare the results from the experiments and the corpus analysis in terms of verb selection preferences. This research demonstrates the productive synergy of corpus linguistic and psycholinguistic methods and findings.

References

Briscoe, E., Carroll, J., & Watson, R. (2006). The Second Release of the RASP System. Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions, Sydney, Australia.

Curran, J.R., Clark, S., & Bos, J. (2007). Linguistically Motivated Large-Scale NLP with C&C and Boxer. Proceedings of the ACL 2007 Demonstrations Session (ACL-07 demo), 33-36.

Ellis, N.C., & Ferreira-Junior, F. (2009). Constructions and their acquisition: Islands and the distinctiveness of their occupancy. Annual Review of Cognitive Linguistics, 111-139.

Francis, G., Hunston, S., & Manning, E. (Eds.). (1996). Grammar Patterns 1: Verbs. The COBUILD Series. London: Harper Collins.

Goldberg, A. E. (2003). Constructions: a new theoretical approach to language. Trends in Cognitive Science, 7, 219-224.

Goldberg, A. E. (2006). Constructions at work: The nature of generalization in language. Oxford: Oxford University Press.

de Marneffe, M., MacCartney, B., & Manning, C.D. (2006). Generating Typed Dependency Parses from Phrase Structure Parses. In LREC 2006.

Zipf, G. K. (1949). Human behaviour and the principle of least effort: An introduction to human ecology. Cambridge, MA: Addison-Wesley.