

Modeling Parsing Constraints in High-Dimensional Semantic Space: On the Use of Proper Names

Curt Burgess, Kay Livesay and Kevin Lund

Department of Psychology
University of California, Riverside
Riverside, CA 92521
curt@cassandra.ucr.edu

The Hyperspace Analogue to Language (HAL) semantic model (Lund & Burgess, in press) demonstrates the impoverished semantics of proper names and the richer semantics of famous names. In addition, proper names used as agentive subjects in a parsing experiment fail to show, in our simulation, an effect of semantic information on parsing due to their impoverished semantics.

The role of semantic context in processing syntactic ambiguity raises important issues concerning the function of the language processor. Modular parsers (Frazier, 1978) are insensitive to contextual information. Conversely, constraint-based approaches (MacDonald, Pearlmutter, & Seidenberg, 1994) propose that language subsystems can more freely interact allowing for semantic influence on initial parsing decisions. Using the HAL model (Burgess & Lund, 1996; Lund & Burgess, in press), we present a semantics of proper names followed by a simulation experiment that suggests that proper names do not contain sufficient semantic information to influence syntactic processing.

Experiment 1: Semantics of Names

Proper names carry little semantic information accounting for the difficulty in their retrieval (Cohen and Burke, 1993). However, famous names, such as Reagan, have a richer semantics. Name semantics can be represented using the HAL model as shown by a nearest neighbor analysis. Near neighbors of *Cathy*, are other common proper names. However, famous names, like *Reagan*, has other politicians and political terms. Likewise, *book* has as its neighbors other related objects. In addition, HAL's high-dimensional semantic vectors can distinguish proper names from nouns and verbs.

Experiment 2: Context Effects and Semantic Distances

HAL's high-dimensional semantics can capture the constraints that predict whether semantic context will facilitate syntactic processing (Burgess and Lund, 1996). Shorter semantic distances were found between noun-verb pairs of stimuli that did not result in a gardenpath (1b, *ransom-paid*) than for sentences that did (1a, *man-paid*).

Verbs like that in sentence (2a) rarely take a direct object; while in sentence (2b), the verb can take either a direct object or a sentence complement. Verbs that do not take a direct object could provide considerable guidance to the parser and avoid syntactic misassignment. However, Ferreira and Henderson (1990, FH) found that reading times for sentences like (2a and 2b) were longer than for unambiguous controls with complementizers. FH conclude that biasing verb information did not effect initial parsing decisions.

- (1a) The man paid by the parents was unreasonable.
- (1b) The ransom paid by the parents was unreasonable.
- (2a) Sue hoped Iowa elected better people.

- (2b) Sue wrote Iowa elected better people.
- (3) The student hoped ...

Another potentially rich source of information in this experiment could be the nature of the noun preceding the verb. The mental model that develops with *Sue hoped* can encompass many possibilities and may not be as constraining as the mental model from (3), which offers the verb crucial additional information during comprehension.

We test the stimuli used in two parsing experiments - one that showed an effect of the NP context (Burgess & Tanenhaus, BT) and one, using proper names, that did not (Ferreira & Henderson, FH). Semantic distances were greater for the stimuli (noun-verbs pairs) that did not show the effect of context (FH) than for the stimuli that did (BT).

Discussion

Models of proper-name retrieval emphasize that name retrieval difficulty is due to the lack of meaningful semantic information in a name, and this can be seen with a name's semantic neighbors. Longer semantic distances with the proper name - verb pairs of FH suggests that proper names do not carry compelling semantic information. That HAL's semantic representations can be used to model the outcome of these two parsing experiments extends the generality of this semantic model to parsing phenomena beyond that of just animacy and general plausibility effects.

Acknowledgments

This research was supported by NSF PFF award SBR-9453406 to Curt Burgess.

References

- Burgess, C., & Lund, K. (1996). *Parsing constraints and high-dimensional semantic space*. Manuscript submitted for publication.
- Burgess, C., & Tanenhaus, M. K. (1992). *Semantic, syntactic and visual factors in syntactic ambiguity resolution*.
- Cohen, G., & Burke, D. M. (1993). Memory for proper names: A review. *Memory*, 1, 249-263.
- Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye-movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 555-568.
- Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. Unpublished doctoral dissertation, University of Connecticut.
- Lund, K., & Burgess, C. (in press). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instrumentation, and Computers*.
- MacDonald, M. C., Pearlmutter, N., & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychology Review*, 101, 676-703.