Predictions of a Model of Language Comprehension Compared to Brain Data

Peter Lindes (plindes@umich.edu)
University of Michigan, Computer Science and Engineering
2260 Hayward Street, Ann Arbor, MI 48109 USA

Keywords: Language comprehension; construction grammar; sentence processing; brain measurements; EEG.

A Model of Human Language Comprehension

Humans understand natural language rapidly in real time. Empirical literature supports the idea that human language comprehension involves *immediate interpretation*. Tanenhaus et al. (1995) show that humans focus their gaze on a particular object in the scene immediately upon hearing a description of that object. Bergen (2012) reviews studies that demonstrate that perceptual and motor areas of the brain are activated dynamically during sentence comprehension. These and other studies show that as soon as a word or phrase that refers to an object or event is processed, its meaning is immediately interpreted and grounded to the situational or dialog context.

Language processing is constrained by the capacity of working memory. Christiansen and Chater (2016) argue that a partial comprehension must be quickly incorporated into larger structures, or it will be lost due to working memory limits. They propose *chunk-and-pass processing*, where the analysis of a sentence is constructed in units they call *chunks*. Whenever possible, chunks already built are composed into larger ones, so that only a few chunks at a time need to be separately maintained in working memory.

Lindes and Laird (2016) have developed a computational theory of language comprehension with immediate interpretation using a chunk-and-pass-like approach. The theory has been implemented in a system called Lucia (Lindes, Mininger, Kirk, & Laird, 2017) that models form-meaning mapping using Embodied Construction Grammar (ECG; B. K. Bergen & Chang, 2013) and is built in the Soar cognitive architecture (Laird, 2012).

An analysis of this model shows that its processing depends on four basic principles:

1. A sentence is comprehended one form-meaning unit, called a *construction*, at a time.
2. Comprehension proceeds as a succession of building these units, or a series of *construction cycles*.
3. Each construction cycle is made up of three phases: *selection*, *integration*, and *grounding*.
4. Each of these phases accesses different types of memory.

In what follows we analyze implications and predictions of this model and compare them to EEG data studies.

Model Predictions

Figure 1 shows a spatiotemporal map of the processing of a simple sentence. Cognitive cycles are grouped into construction cycles, and then word cycles.

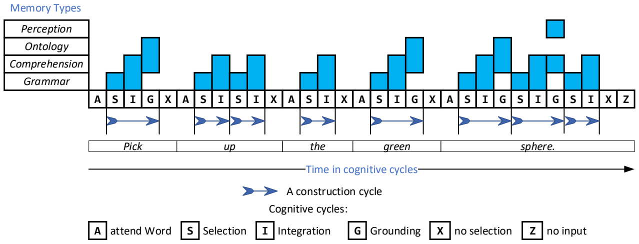

Figure 1: Time course of memory use

In each construction cycle, one construction from the available inventory in the grammar memory is selected (S) and integrated (I) into the comprehension state in working memory. Optionally, it is grounded (G) to the agent's knowledge. A, X, and Z operators perform overhead functions, such as attending to the next word.

The model uses four types of memory. Long-term memories store linguistic knowledge (Grammar) and the agent's long-term knowledge of the environment and its perception and action capabilities (Ontology). Short-term memories store the state of the comprehension process (Comprehension) and the agent's current perception, dialog, and situational states (Perception).

This model implies a time sequence in which different memories are accessed at different times, as Figure 1 shows. We suggest that this spatiotemporal pattern of memory accesses may approximate a similar pattern of activation in the brain. In accordance with standard modeling, the cognitive cycles have a 50ms time course in humans, modulated by long-term memory access. We expect that comparing these predictions to brain data will help understand both the brain and the model better.

Comparison to Brain Data

Figure 2 shows examples of the kind of data reported in the large literature on measurements of the brain during language comprehension (Left: Schwartz & Mitchell, 2019; Right: Hale, Dyer, Kuncoro, & Brennan, 2018). The images show several kinds of Event Related Potentials (ERPs) averaged over many words as they are distributed in time and space. We will compare our model to these and other related data.

Our model predicts that more time is required to process content words than function words due to grounding and the frequent need for multiple constructions. Brennan and Hale (2019) compare several simpler models to EEG data and show processing differences between these two types of words, and that less frequent content words have a stronger N400 response. Further analysis of the details is needed.

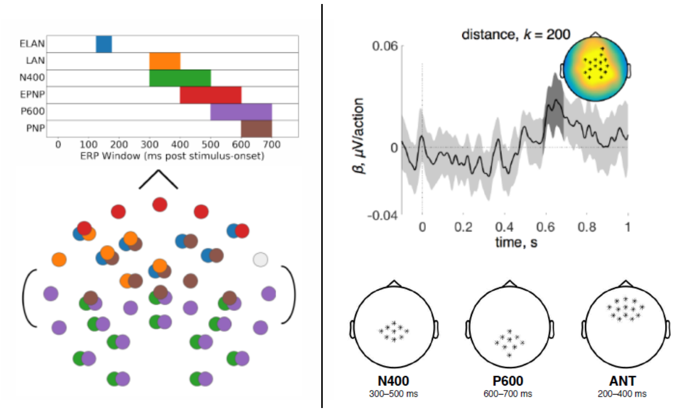

Figure 2: Examples of EEG data

Our model shows that the number of construction cycles per word varies, and the length of each cycle varies depending on whether grounding is needed. The model implies that processing of each word runs to completion before the next word is attended to. The EEG data shows word processing extending to beyond 600ms, even to almost a second (see Fig, 2). Human language input proceeds typically in a range of 150-250 words/minute, for an average time between words of 240-400ms. This implies that word processing often continues in parallel with the processing of subsequent words. This sort of parallelism is lacking in our model. Figure 3 gives a suggestion of what the processing pattern might look like. How to accomplish this within the Soar architecture is an open question.

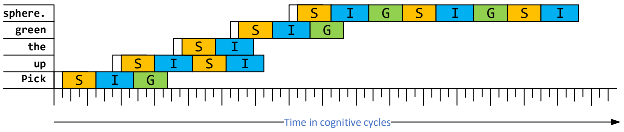

Figure 3: A possible parallel processing pattern

Bornkessel-Schlesewsky and Schlesewsky (2019) present a new, unified way of looking at the N400 response, arguing that it happens when the brain needs to modify its predictive model of the sentence. Our system restructures its sentence model each time it adds a construction, and when it performs a local repair. A general prediction function, which is lacking from our current model, will need to be added for it to better reflect the brain's processing.

Our model suggests there are different memory areas involved in language comprehension, and that there is a repetitive time sequence in their accesses. The data in Figure 2 show temporal patterns in the spatial distribution of brain activation. There is the potential here to improve both our understanding of the brain and our model by further analysis of the relationship between these two things.

This abstract suggests ways in which cognitive models of comprehension and brain measurements can be compared to improve both the models and understanding of the brain. The analysis here is very preliminary and superficial; much work is needed to explore these areas in detail.

Acknowledgments

The work described here was supported by the AFOSR under Grant Number FA9550-18-1-0168. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressly or implied, of the AFOSR or the U.S. Government.

References

Bergen, B. K. (2012). *Louder Than Words: The New Science of How the Mind Makes Meaning*. New York: Basic Books.

Bergen, B. K., & Chang, N. C. (2013). Embodied Construction Grammar. In T. Hoffman & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 168–190). Oxford University Press.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Toward a Neurobiologically Plausible Model of Language-Related, Negative Event-Related Potentials. *Frontiers in Psychology*, *10*(FEB). https://doi.org/10.3389/fpsyg.2019.00298

Brennan, J. R., & Hale, J. T. (2019). Hierarchical structure guides rapid linguistic predictions during naturalistic listening. *PLoS ONE*, *14*(1). https://doi.org/10.1371/journal.pone.0207741

Christiansen, M. H., & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, 1–72. https://doi.org/10.1017/S0140525X1500031X

Hale, J., Dyer, C., Kuncoro, A., & Brennan, J. R. (2018). *Finding Syntax in Human Encephalography with Beam Search*. (2014). Retrieved from http://arxiv.org/abs/1806.04127

Laird, J. E. (2012). *The Soar Cognitive Architecture*. Cambridge, MA: The MIT Press.

Lindes, P., & Laird, J. E. (2016). Toward Integrating Cognitive Linguistics and Cognitive Language Processing. *International Conference on Cognitive Modeling (ICCM)*.

Lindes, P., Mininger, A., Kirk, J. R., & Laird, J. E. (2017). Grounding Language for Interactive Task Learning. *Proceedings of the First Workshop on Language Grounding for Robotics*, 1–9. https://doi.org/10.18653/v1/w17-2801

Schwartz, D., & Mitchell, T. (2019). *Understanding language-elicited EEG data by predicting it from a fine-tuned language model*. Retrieved from http://arxiv.org/abs/1904.01548

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K., & Sedivy, J. C. (1995). Integration of Visual and Linguistic Information in Spoken Language Comprehension. *Science*, *268*(5217), 1632.