Linking LFG to tiered models of processing

Stephen Jones, University of Oxford

This paper extends previously-presented work on cognitive modelling of processing based on LFG, which developed a model representation analogous to LFG f-structure and showed that GF information could function as memory retrieval cues. It introduces a proposal to include processing tiers in the model analogous to s-structure and the discourse context, describes challenges that an incremental approach presents for mapping theory, and proposes possible solutions.

Language processing is often assumed to require repeated applications of cue-based memory retrievals, with the retrieved memory begin combined with encoded language input to create mental representations. These retrievals are assumed to be not only from longer-term memory (e.g. the semantic associations or combinatorial requirements/constraints of a particular word) but also from working memory, integrating new content into an emerging representation. Under this model, one explanation of increased processing time is the burden of choosing between multiple candidates activated by the retrieval cues at a given stage of a parse.

Many phenomena have been identified (Lewis and Phillips, 2015) where grammar appears to influence processing speed, including identifying the coreferents of "filler-gap" long distance dependencies, constraints on anaphora reference, garden path effects, and illusory comparatives such as *More people have been to Moscow than I have.* Computational models of parsing these effects (e.g. Lewis and Vasishth, 2005; Hale, 2014; Stewart et al., 2014; VanWagenen et al., 2014; Engelmann, 2016) have generally aimed to build constituency-based syntactic structure, either as a tree representation, or using a Construction Grammar approach.

LFG accounts of the relevant phenomena usually use levels of representation other than c-structure to account for binding constraints, long-distance dependencies, scope ambiguity etc. Even where c-structure constraints have been included in an account, e.g. Bresnan (1995) on weak crossover, other levels of representation are included and phenomena such as garden-pathing may refer to lexical or discourse content to account for why one phrase structure is preferred over another on-line.

Christiansen and Chater (2016) and Kuperberg and Jaeger (2016) have proposed informal models of processing under which representations of language are encoded at increasing tiers of abstraction: the product of one processing tier providing input to the next. In these models, processing speed may vary not only with memory retrieval, but also if anomalies or multiple possibilities arise at the interface between tiers. Kuperberg and Jaeger go further and propose a predictive model, where probabilistic inferences from a higher tier influence later processing outcomes at a lower tier through pre-activation of likely candidates for retrieval in memory.

Christiansen and Chater's model refers to a form of Construction Grammar in building the syntactic element of representation. However LFG offers a potential alternative in which there is a unified formalism to represent multiple tiers and account for constraints that apply between them.

The model

The model is implemented in ACT-R 7 (Anderson et al., 2001). It follows Asudeh (2012) in assuming that a full LFG analysis can be developed after each word is processed. It follows Findlay (2016) in mapping grammatical functions directly to argument positions in s-structure.

Examples (1) and (2) show the incremental development of s-structure alongside f-structure for two simple example sentences from Findlay. An immediate challenge is the assignment of the arguments. in (1a) and (2a): the ARG positions present in s-structure and the order of their assignment is not clear until information on the REL and any template such as @PASSIVE has been obtained from the verb. An initial assignment of the first argument to ARG1 is unproblematic for verbs in active voice, or where there is no fronting of non-subject constituents. However, in other cases, additional processing will be required to reassign the argument to ARG2 if necessary.

(1) Kim devoured the cake.

Word	f-structure	s-structure
(a) Kim	$\begin{bmatrix} \text{PRED} & \text{`\ldots'} \\ \text{UDF} \mid \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`Kim'} \end{bmatrix} \end{bmatrix}$	$\begin{bmatrix} \text{REL} & \ldots \\ \text{ARG}_? & k{:}\begin{bmatrix} \; \end{bmatrix} \end{bmatrix}$
(b) devoured	$\begin{bmatrix} \text{PRED} & \text{`devour'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`Kim'} \end{bmatrix} \end{bmatrix}$	$\begin{bmatrix} \text{REL} & \text{devour} \\ \text{EV} & ev{:}\begin{bmatrix} \; \end{bmatrix} \\ \text{ARG}_1 & k{:}\begin{bmatrix} \; \end{bmatrix} \\ \text{ARG}_2 & \end{bmatrix}$
(c) the cake	$\begin{bmatrix} \text{PRED} & \text{`devour'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`Kim'} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`cake'} \\ \text{DEF} & + \end{bmatrix} \end{bmatrix}$	$\begin{bmatrix} \text{REL} & \text{devour} \\ \text{EV} & ev{:}\begin{bmatrix} \; \end{bmatrix} \\ \text{ARG}_1 & k{:}\begin{bmatrix} \; \end{bmatrix} \\ \text{ARG}_2 & c{:}\begin{bmatrix} \; \end{bmatrix} \end{bmatrix}$

(2) The cake was devoured by Kim.

Word	f-structure	s-structure
(a) The cake	$\begin{bmatrix} \text{PRED} & \text{`\ldots'} \\ \text{UDF} \mid \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`cake'} \\ \text{DEF} & + \end{bmatrix} \end{bmatrix}$	$\begin{bmatrix} \text{REL} & \ldots \\ \text{ARG}_? & c{:}\begin{bmatrix} \; \end{bmatrix} \end{bmatrix}$
(b) was devoured	$\begin{bmatrix} \text{PRED} & \text{`devour'} \\ \text{VOICE} & \text{PASSIVE} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`cake'} \\ \text{DEF} & + \end{bmatrix} \end{bmatrix}$	$\begin{bmatrix} \text{REL} & \text{devour} \\ & \text{@Passive} \\ \text{EV} & ev{:}\begin{bmatrix} \; \end{bmatrix} \\ \text{ARG}_1 & \\ \text{ARG}_2 & c{:}\begin{bmatrix} \; \end{bmatrix} \end{bmatrix}$
(c) by Kim	$\begin{bmatrix} \text{PRED} & \text{`devour'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`cake'} \\ \text{DEF} & + \end{bmatrix} \\ \text{OBL}_{BY} & \begin{bmatrix} \text{PRED} & \text{`Kim'} \end{bmatrix} \end{bmatrix}$	$\begin{bmatrix} \text{REL} & \text{devour} \\ \text{EV} & ev{:}\begin{bmatrix} \; \end{bmatrix} \\ \text{ARG}_1 & k{:}\begin{bmatrix} \; \end{bmatrix} \\ \text{ARG}_2 & c{:}\begin{bmatrix} \; \end{bmatrix} \end{bmatrix}$

Alternatively, if the argument position remains underspecified, all cases will require additional processing to fully specify once the verb and auxiliaries have been processed. The implications of both alternatives are compared.

References

Anderson, J. R., Budiu, R., and Reder, L. M. (2001). A theory of sentence memory as part of a general theory of memory. *Journal of Memory and Language*, 45(3):337–367.

Asudeh, A. (2012). *The logic of pronominal resumption*. Number 35 in Oxford Studies in Theoretical Linguistics. Oxford University Press, Oxford.

Bresnan, J. (1995). Linear order, syntactic rank, and empty categories: On weak crossover. In Dalrymple, M., Kaplan, R. M., Maxwell, J. T., and Zaenen, A. E., editors, *Formal Issues in Lexical-Functional Grammar*, pages 241–274. CSLI Publications.

Christiansen, M. H. and Chater, N. (2016). The now-or-never bottleneck: a fundamental constraint on language. *Behavioral and Brain Sciences*, 39:1–19.

Engelmann, F. (2016). An integrated model of eye movement behaviour in sentence comprehension. `https://github.com/felixengelmann/act-r-sentence-parser-em`.

Findlay, J. Y. (2016). Mapping theory without argument structure. *Journal of Language Modelling*, 4(2):293–338.

Hale, J. T. (2014). *Automaton theories of human sentence comprehension*. CSLI Publications, Stanford, CA.

Kuperberg, G. R. and Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1):32–59.

Lewis, R. L. and Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29:375–419.

Lewis, S. and Phillips, C. (2015). Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research*, 44:27–46.

Stewart, T., Choo, F.-X., and Eliasmith, C. (2014). Sentence processing in spiking neurons: A biologically plausible left-corner parser. In *Proceedings of the Cognitive Science Society*, volume 36.

VanWagenen, S., Brennan, J., and Stabler, E. P. (2014). Quantifying parsing complexity as a function of grammar. In Schütze, C. T. and Stockall, L., editors, *Connectedness: Papers by and for Sarah VanWagenen*, volume 18 of *UCLA Working Papers in Linguistics*. Los Angeles: UCLA.