Linking LFG to tiered models of processing

Stephen Jones, University of Oxford

LFG18 Conference, University of Vienna
18 July 2018

This paper extends previously-presented work (Jones, 2017) 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.

1 Background

Language processing is often assumed to require repeated applications of cue-based memory retrievals, with the retrieved memory being 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 variations in processing time for particular words is the differing 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 "filler" and "gap" in 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 relies on 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.

2 The model

The model assumes the LFG architecture proposed by Mycock and Lowe (2013), associating lexical entries and tiers of representation with elements of an ACT-R module as shown in Figure 1. 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.

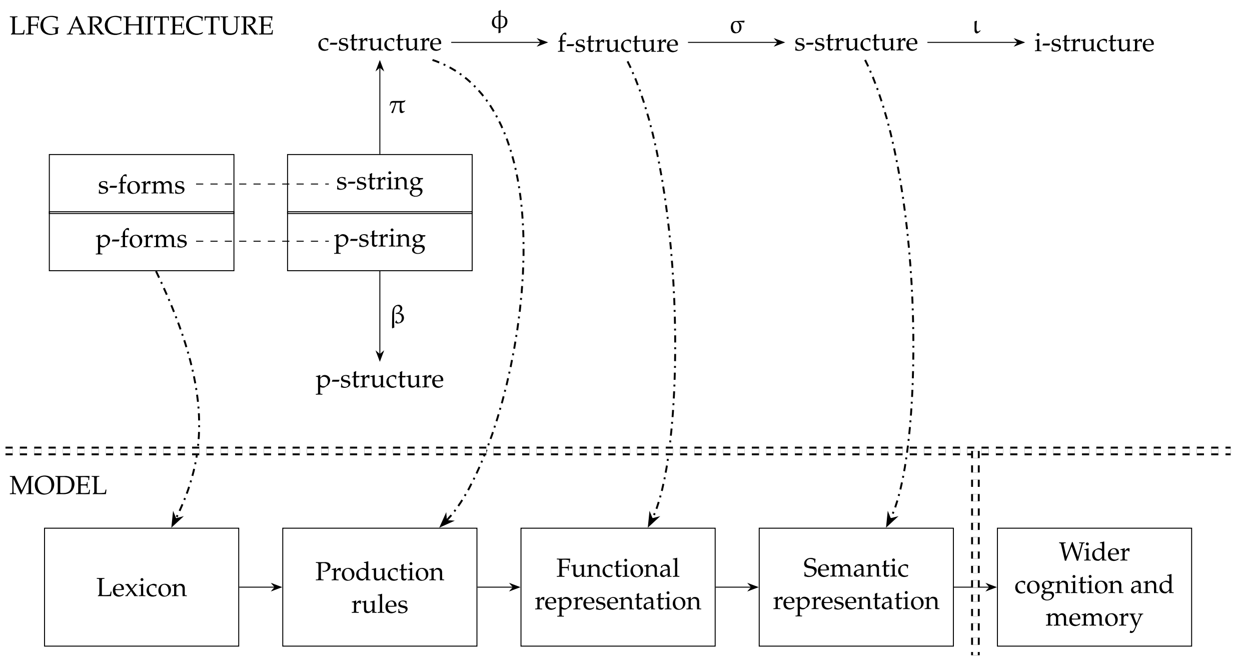

Figure 1: Associating the model with LFG architecture

3 Incorporating c-structure without building a representation

Building and retrieving a mental representation of c-structure is costly, but models of Korean and English show that c-structure information can be captured in a representation of f-structure if:

(i) lexical specifications and procedural knowledge constrain the ordering of attributes that can receive new information; and

(ii) lexically-specified features TYPE and PRED.CAT are added to the representation

3.1 Korean

Korean is a head-final language where grammatical function information is generally provided lexically through particles[1]. The ordering of arguments and modifiers before a head is free, and arguments may be omitted if they are clear from the context. However, if an argument or a

[1]Where no particles are provided, default SOV word order is inferred. If all previous arguments have been lexically specified, the GF of the final argument may be inferred rather than explicitly provided by a particle (Kiaer, 2011)

modifier is expressed, its head must also be expressed. One constituent[2] may be scrambled to the front of the sentence and marked with the discourse particle *-un/-nun* (Sohn, 1994).

The phrase structure rules for Korean in (1) are derived from Cho and Sells (1995), who use the notion of types to account for morphosyntactic constraints on particles. There are three types: TYPE:NO, TYPE:V-SIS, and TYPE:N-SIS.

(1) a. $$S \rightarrow \left\{ \begin{array}{c} \Sigma' \\ \text{TYPE:NO} \\ \uparrow=\downarrow \end{array} \middle| \begin{array}{cc} X' & V' \\ \text{TYPE:V-SIS} & \text{TYPE:NO} \\ (\uparrow \text{DF}) =\downarrow & \uparrow=\downarrow \end{array} \right\}$$

b. $$\begin{array}{c} X' \\ \text{TYPE:}\alpha \end{array} \rightarrow \left\{ \begin{array}{c} X \\ \text{TYPE:}\alpha \\ \uparrow=\downarrow \end{array} \middle| \begin{array}{cc} \Sigma' & X' \\ \text{TYPE:X-SIS} & \text{TYPE:}\alpha \\ \uparrow=\downarrow & \uparrow=\downarrow \end{array} \right\}$$

From (1b), a phrase of arbitrary type X-SIS is constrained to be the right sister of a phrase whose head is of category X. The rightmost phrase of a sentence is constrained to be TYPE:NO. The type of a phrase and the category of its head can be included in f-structure as the features TYPE and PRED.CAT respectively.

Because the head of a phrase provides the PRED value for its f-structure, and because Korean is strictly head-final, the constraint derived from (1b) can be framed in terms of f-structure as shown in (2), where $\leftrightarrow$ represents the correspondence between a type and a category.

(2) f PRED.CAT $\leftrightarrow$ f GF TYPE

In other words, if a word's lexical entry includes TYPE = X-SIS, then the word can be a GF in relation to an f-structure where f PRED.CAT = X. F-structure is built incrementally using one of the strategies Complete, Add daughter, Add granddaugther, and Push down. The available strategies are determined by comparing the category and type of the word to be added with the PRED.CAT and TYPE values of the target f-structure that requires a PRED value, using the algorithm shown in Figure 2.

The algorithm decision points shown in blue with a single edge on the poster, and the strategies Complete, Add daughter, and Add granddaughter, are completely determined by the contents of the lexical entry and target f-structure. Table 1 shows the operation of these strategies.

The two choice points with double-lined edges (shown in green on the poster), and the Push down strategy allow optionality in the parsing process. The choice point *Embed (some) existing content?* allows Push down to be selected, if the word to be processed's TYPE value corresponds to its own category and to *f* PRED.CAT. If a word could provide the value of *f* PRED and so close the f-structure, the choice point *Early closure?* allows that word to be added as a daughter or granddaughter f-structure of *f*, allowing late closure to take place.

The Push down strategy results in one or more f-structure daughters of the target f-structure being embedded underneath the word that is added. If Push down is used, there is a free choice as to how many of the most recently-added daughter f-structures are embedded. Table 2 shows the outcome of applying Push down either fully, with all daughter f-structures embedded, or partially, with some daughter f-structures remaining unembedded.

The model can be used to generate predictions from hypotheses about circumstances under which particular free choices are made, including prior context, and the presence of particular prosodic boundaries. These can then be tested empirically, which is an area for future exploration.

[2] In exceptional circumstances two arguments to the same head may be fronted, with prosodic support (Kempson and Kiaer, 2010)

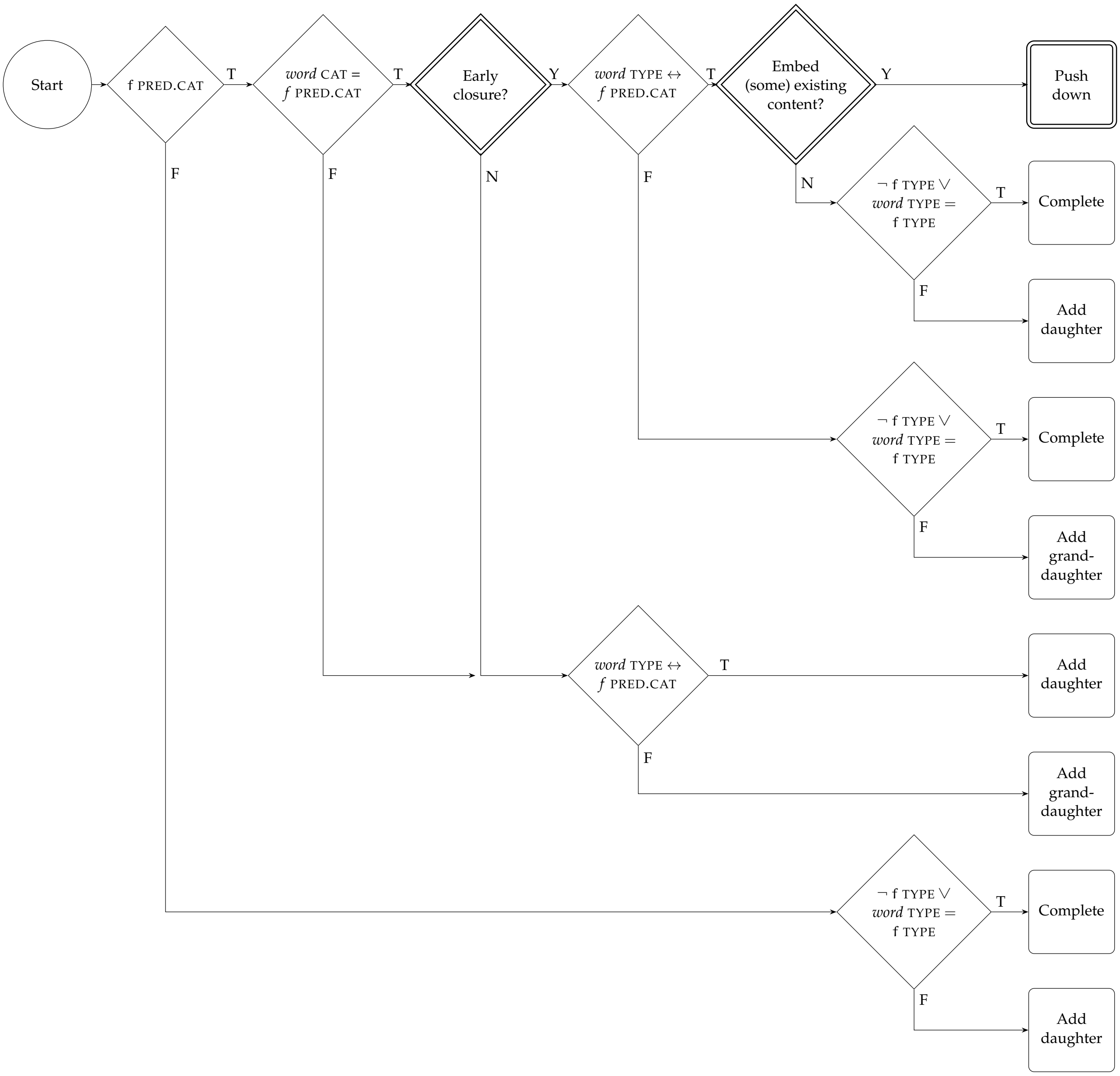

Figure 2: Korean: Algorithm for adding information from *word* to f-structure f

3.1.1 Worked example

(3) *mina-nun minho-ka coh-un yenghwa-lul bwass-tako sayngkakha-nta*
Mina-DF Minho-SBJ good-COMP film-OBJ saw-QUOT think-PRES.PLAIN
"As for Mina, she thinks that Minho saw a good film."

Next word	Available operations	Starting f-structure
		$f\begin{bmatrix}\text{PRED} & \ldots① \\ \text{PRED.CAT} & ? \\ \text{TYPE} & ?\end{bmatrix}$

N *mina-nun* TYPE:V-SIS (↑ GF = DF)	Complete	$g\left[\begin{array}{ll} \text{PRED} & \ldots\textcircled{1} \\ \text{PRED.CAT} & \text{V} \\ \text{TYPE} & ? \\ \text{DF} & f\left[\begin{array}{ll} \text{PRED} & \text{'Mina'} \\ \text{PRED.CAT} & \text{N} \\ \text{TYPE} & \text{V-SIS} \end{array}\right] \end{array}\right]$
N *minho-ka* TYPE:V-SIS (↑ GF = SUBJ)	Add daughter	$g\left[\begin{array}{ll} \text{PRED} & \ldots\textcircled{1} \\ \text{PRED.CAT} & \text{V} \\ \text{TYPE} & ? \\ \text{DF} & f\left[\begin{array}{ll} \text{PRED} & \text{'Mina'} \end{array}\right] \\ \text{SUBJ} & h\left[\begin{array}{ll} \text{PRED} & \text{'Minho'} \\ \text{TYPE} & \text{V-SIS} \end{array}\right] \end{array}\right]$
V *coh-un* TYPE:N-SIS (↑ GF = ADJ)	Push down **Add granddaughter**	$g\left[\begin{array}{ll} \text{PRED} & \ldots\textcircled{2} \\ \text{PRED.CAT} & \text{V} \\ \text{TYPE} & ? \\ \text{DF} & f\left[\begin{array}{ll} \text{PRED} & \text{'Mina'} \end{array}\right] \\ \text{SUBJ} & h\left[\begin{array}{ll} \text{PRED} & \text{'Minho'} \end{array}\right] \\ \text{GF} & j\left[\begin{array}{ll} \text{PRED} & \ldots\textcircled{1} \\ \text{PRED.CAT} & \text{N} \\ \text{TYPE} & \text{V-SIS} \\ \text{ADJ} & \left\{ i\left[\begin{array}{ll} \text{PRED} & \text{'good'} \\ \text{PRED.CAT} & \text{N} \\ \text{TYPE} & \text{N-SIS} \end{array}\right] \right\} \end{array}\right] \end{array}\right]$
N *yenghwa-lul* TYPE:V-SIS (↑ GF = OBJ)	Complete	$g\left[\begin{array}{ll} \text{PRED} & \ldots\textcircled{1} \\ \text{PRED.CAT} & \text{V} \\ \text{TYPE} & ? \\ \text{DF} & f\left[\begin{array}{ll} \text{PRED} & \text{'Mina'} \end{array}\right] \\ \text{SUBJ} & h\left[\begin{array}{ll} \text{PRED} & \text{'Minho'} \end{array}\right] \\ \text{GF} & j\left[\begin{array}{ll} \text{PRED} & \text{'film'} \\ \text{PRED.CAT} & \text{N} \\ \text{TYPE} & \text{V-SIS} \\ \text{ADJ} & i\left\{\left[\begin{array}{ll} \text{PRED} & \text{'good'} \end{array}\right]\right\} \end{array}\right] \end{array}\right]$

V *bwass-tako* TYPE:V-SIS (↑ GF = COMP)	**Push-down** Complete	*g* [PRED …① PRED.CAT V TYPE ? DF *f* [PRED 'Mina'] COMP *k* [PRED 'think' PRED.CAT V TYPE V-SIS SUBJ *h* [PRED 'Minho'] OBJ *j* [PRED 'film' ADJ { *i* [PRED 'good'] }]]]
V *sayngkahna-nta* TYPE:V-SIS	Complete	*g* [PRED 'think' PRED.CAT V TYPE NO DF *f* [PRED 'Mina'] COMP *k* [PRED 'think' PRED.CAT V TYPE V-SIS SUBJ *h* [PRED 'Minho'] OBJ *j* [PRED 'film' adj { *i* [PRED

Table 1: Available strategies for incremental growth of structure in Korean

$$\text{Adding} \quad \begin{array}{c}\text{Y}\\ \textit{word}\\ \text{TYPE:X-SIS}\end{array} \quad \text{in the context of} \quad f\begin{bmatrix}\text{PRED} & \ldots\text{①}\\ \text{PRED.CAT} & ?\\ \text{TYPE} & ?\end{bmatrix}$$

Strategy	Resulting f-structure
Complete	$g\begin{bmatrix}\text{PRED} & \ldots\text{①}\\ \text{PRED.CAT} & \text{X}\\ \text{TYPE} & ?\\ \text{GF} & f\begin{bmatrix}\text{PRED} & \text{‘word’}\\ \text{PRED.CAT} & \text{Y}\\ \text{TYPE} & \text{X-SIS}\end{bmatrix}\end{bmatrix}$
Add daughter	$f\begin{bmatrix}\text{PRED} & \ldots\text{①}\\ \text{PRED.CAT} & \text{X}\\ \text{TYPE} & ?\\ \text{GF} & g\begin{bmatrix}\text{PRED} & \text{‘word’}\\ \text{PRED.CAT} & \text{Y}\\ \text{TYPE} & \text{X-SIS}\end{bmatrix}\end{bmatrix}$
Add granddaughter	$f\begin{bmatrix}\text{PRED} & \ldots\text{②}\\ \text{PRED.CAT} & ?\\ \text{TYPE} & ?\\ \text{GF} & h\begin{bmatrix}\text{PRED} & \ldots\text{①}\\ \text{PRED.CAT} & \text{X}\\ \text{TYPE} & ?\\ \text{GF} & g\begin{bmatrix}\text{PRED} & \text{‘word’}\\ \text{PRED.CAT} & \text{Y}\\ \text{TYPE} & \text{X-SIS}\end{bmatrix}\end{bmatrix}\end{bmatrix}$

3.2 English

The constraints on building f-structure derived in this section are based on the mini-grammar of English provided in Falk (2001). In English, GF information is provided principally through structural information, rather than lexically. In general, a functional head precedes the head of the lexical phrase that depends on it, and specifiers precede functional heads. The PRED value of the f-structure projected from a functional phrase is provided by the head of its lexical dependent. Complements follow the phrasal head and, for heads that select multiple GF com-

Table 2: Full and partial Push down

Adding $\begin{array}{c}\text{X}\\ \textit{word}_3\\ \text{TYPE:X-SIS}\end{array}$ in the context of $f\left[\begin{array}{ll} \text{PRED} & \ldots \textcircled{1} \\ \text{PRED.CAT} & \text{X} \\ \text{GF}_1 & g\left[\begin{array}{ll}\text{PRED} & \text{`word}_1\text{'} \\ \text{TYPE} & \text{X-SIS}\end{array}\right] \\ \text{GF}_2 & h\left[\begin{array}{ll}\text{PRED} & \text{`word}_2\text{'} \\ \text{TYPE} & \text{X-SIS}\end{array}\right]\end{array}\right]$

Strategy	Resulting f-structure
Full Push-down	$f\left[\begin{array}{ll} \text{PRED} & \ldots \textcircled{1} \\ \text{PRED.CAT} & \text{X} \\ \text{GF}_4 & i\left[\begin{array}{ll} \text{PRED} & \text{`word}_3\text{'} \\ \text{PRED.CAT} & \text{X} \\ \text{TYPE} & \text{X-SIS} \\ \text{GF}_1 & g\left[\begin{array}{ll}\text{PRED} & \text{`word}_1\text{'} \\ \text{TYPE} & \text{X-SIS}\end{array}\right] \\ \text{GF}_2 & h\left[\begin{array}{ll}\text{PRED} & \text{`word}_2\text{'} \\ \text{TYPE} & \text{X-SIS}\end{array}\right]\end{array}\right]\end{array}\right]$
Partial Push-down	$f\left[\begin{array}{ll} \text{PRED} & \ldots \textcircled{1} \\ \text{PRED.CAT} & \text{X} \\ \text{GF}_1 & g\left[\begin{array}{ll}\text{PRED} & \text{`word}_1\text{'} \\ \text{TYPE} & \text{X-SIS}\end{array}\right] \\ \text{GF}_4 & i\left[\begin{array}{ll} \text{PRED} & \text{`word}_3\text{'} \\ \text{PRED.CAT} & \text{X} \\ \text{TYPE} & \text{X-SIS} \\ \text{GF}_2 & h\left[\begin{array}{ll}\text{PRED} & \text{`word}_2\text{'} \\ \text{TYPE} & \text{X-SIS}\end{array}\right]\end{array}\right]\end{array}\right]$

plements, there are constraints on the order that those GFs appear. These are shown in Table 4.

The lexical entries for words from functional categories C, D, and I carry the specifications in (4).

(4) a. C: ($\uparrow$ PRED.CAT) $=_c$ V

b. D: ($\uparrow$ PRED.CAT) $=_c$ N

c. I: ($\uparrow$ PRED.CAT) $=_c$ V

Table 4: Ordering constraints on building f-structure incrementally in English

<table>
<tr><th>PRED.CAT</th><th>Order in which attributes provided</th><th>Other constraints</th></tr>
<tr><td>V</td><td>PRED ④
PRED.CAT V
{DF | COMPFORM} ❶
SUBJ ②
OBJ ⑤
OBJ_{θ} ⑥
XCOMP ⑦
OBL_{θ} ⑧
COMP ⑨
ADJ (❶) ❸ ❿</td><td>❸ ❿ PRED.CAT $=_c$ {ADV|P}
⑤⑥⑦⑧⑨ lexically specified
⑤⑥ PRED.CAT $=_c$ N
⑦⑨ PRED.CAT $=_c$ V
⑧ PRED.CAT $=_c$ P</td></tr>
<tr><td>P</td><td>PRED ②
PRED.CAT P
{OBJ | OBL_{θ} | COMP} ③
ADJ ❶ ❹</td><td>❶ PRED.CAT $=_c$ ADV
❹ PRED.CAT $=_c$ P</td></tr>
<tr><td>N</td><td>PRED ③
PRED.CAT N
{DEF | POSS} ①
COMPLEMENT ④
ADJ ❷ ❺</td><td>❷ PRED.CAT $=_c$ A
④ lexically specified
❺ PRED.CAT $=_c$ {P|V}</td></tr>
<tr><td>A</td><td>PRED ②
PRED.CAT {A | ADV}
{OBL_{θ} | COMP} ③
ADJ ❶ ❹</td><td>❶ PRED.CAT $=_c$ ADV
③ lexically specified
❹ PRED.CAT $=_c$ P</td></tr>
</table>

Although the structural rules for English are very different to those for Korean, it is still possible to build f-structure incrementally by comparing the lexical specification from the current word with the current f-structure including its constraints. The constraints operate as follows:

(i) Attributes shown black-on-white ① are required (if lexically specified).

(ii) Attributes shown white-on-black ❶ are optional.

(iii) Ordering constraints on sisters in the same f-structure must be followed.

(iv) Introducing a daughter f-structure may introduce the requirement for additional information within that daughter. This takes precedence over other outstanding requirements.

(v) Only when an f-structure is complete can information be added to its parent.

In general, new structure is added using the "add daughter" strategy described for Korean. However, the "push-down" strategy is also used.

(English worked example available separately from `stephen.jones@ling-phil.ox.ac.uk` *)*

4 Getting through Christiansen & Chater's *Now-or-Never* bottleneck

The Now-or-Never bottleneck demands that ambiguities are managed immediately. Possible strategies to do this include predictive assignment and underspecification, which make differing, testable predictions about the need for subsequent reanalysis. These are the subject of ongoing investigation using the model.

5 Reanalysis

Empirically, some reanalyses are easier than others. In the model the ends of GF paths are easier to change than intermediate links. Sometimes GFs can be reassigned during processing: does this support feature-based rather than atomic representation of GFs?

This is illustrated by comparing the sentences in (5), taken from Fodor and Ferreira (1998) where no garden path is observed for (5a), but a garden path is observed at *was* in (5b).

(5) a. The guests saw the cake *was* still being decorated.
    b. While the guests ate the cake *was* still being decorated.

Under the model, reanalysis is required in both sentences to add the copula *was*, but this has a different impact in the two cases. The reanalysis is presented below in Figures 3 and 4, where the reanalysed grammatical functions before and after reanalysis are shown in a box, and text in ~~strikeout~~ shows elements from the prior f-structure that no longer appear after reanalysis.

Figure 3 shows the reanalysis for (5a), where no garden path is observed, and Figure 4 shows the reanalysis for (5b).

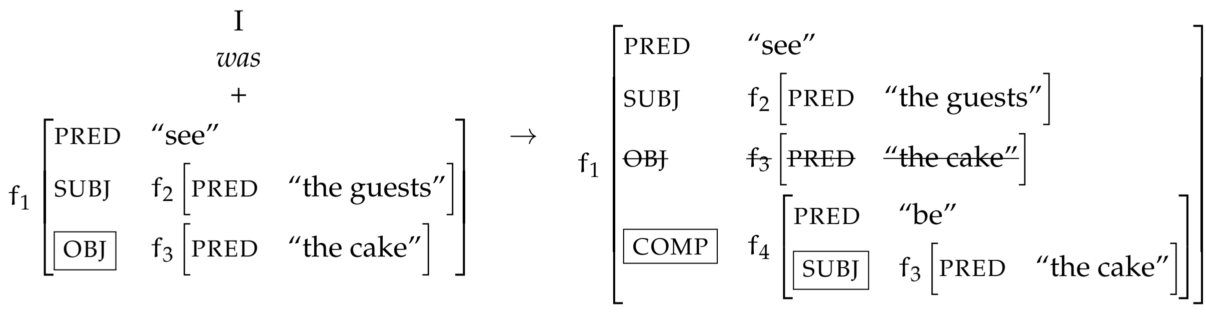

Figure 3: Reanalysis associated with no garden path at 'The guests saw the cake *was* ...'

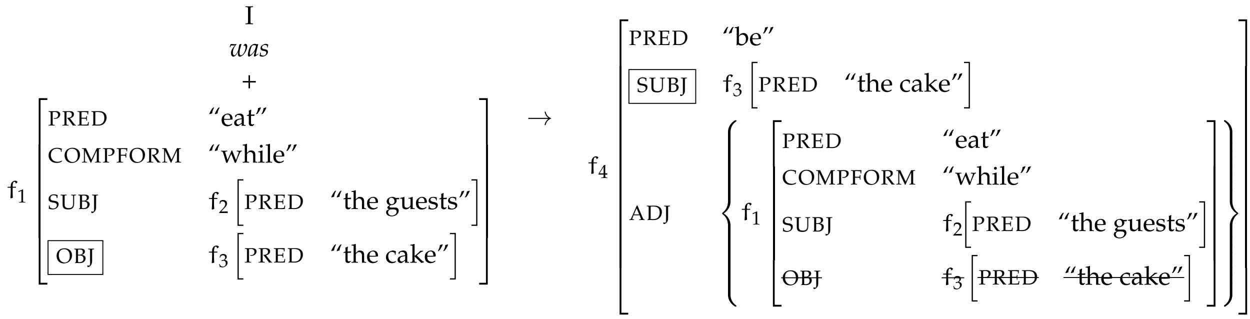

Figure 4: Reanalysis associated with garden path at 'While the guests ate the cake *was* ...'

Table 5: F-structure containment relationships before and after reanalysis

	No garden path (5a)	Garden path (5b)
Before reanalysis	f_1 contains f_2, f_3	f_1 contains f_2, f_3
	↓	↓
After reanalysis	f_1 contains f_2, f_3, f_4	f_1 contains f_2
	f_4 contains f_3	f_4 contains f_1, f_2, f_3

Table 5 shows the f-structure containment relationships that are present before and after reanalysis. Comparing the two cases, the reanalysis without garden path does not disrupt any containment relationships: the set grows monotonically. However, for the garden path case, containment relationships are disrupted.

In both reanalyses, the GF of f_3 is reassigned from OBJ to SUBJ. In the non-garden-path case, the valency of *see* changes from ⟨SUBJ,OBJ⟩ to ⟨SUBJ,COMP⟩. That these reassignments seem to be unproblematic raises questions about how grammatical functions should be encoded in the model, in particular whether a disjunction can be assigned initially, or whether in fact grammatical functions should be considered as feature sets as previously suggested by various authors including Butt (1995) and Findlay (2016). This requires further exploration.

References

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.

Butt, M. (1995). *The structure of complex predicates in Urdu*. CSLI Publications, Stanford, CA.

Cho, Y.-M. Y. and Sells, P. (1995). A lexical account of inflectional suffixes in Korean. *Journal of East Asian Linguistics*, 4(2):119–174.

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.

Falk, Y. N. (2001). *Lexical-Functional Grammar: An introduction to parallel constraint-based syntax*. CSLI Publications, Stanford, CA.

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

Fodor, J. D. and Ferreira, F., editors (1998). *Reanalysis in sentence processing*. Kluwer, Dordrecht.

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

Jones, S. (2017). Developing a computational model of language processing using LFG. Poster presented at the LFG17 Conference, Universität Konstanz, Germany.

Kempson, R. and Kiaer, J. (2010). Multiple long-distance scrambling: Syntax as reflections of processing. *Journal of Linguistics*, 46:127–192.

Kiaer, J. (2011). Word-order variation in Korean and the architecture of grammar. In Kempson, R., Gregoromichelaki, E., and Howes, C., editors, *The dynamics of lexical interfaces*, chapter 6, pages 165–203. 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.

Mycock, L. and Lowe, J. (2013). The prosodic marking of discourse functions. In Butt, M. and King, T. H., editors, *Proceedings of LFG13*, Stanford, CA. CSLI Publications.

Sohn, H.-M. (1994). *Korean*. London : Routledge, London.

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.