A grammatically robust cognitive model of English and Korean sentence processing

Stephen Jones (s.m.jones@rug.nl)
Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence,
University of Groningen, 9700AB Groningen, Netherlands

Abstract

This paper presents a novel approach to the cognitive modelling of human sentence processing in ACT-R. The model assumes a cognitive distinction between cross-linguistic knowledge of the overall possibilities for combining elements of language structure, represented in procedural memory, and language-specific knowledge of the combinatorial constraints on structure-building, which are stored as part of the lexicon in declarative memory. Sentence structure is built incrementally using an extension of an established, computationally robust grammar theory, Lexical Functional Grammar (Bresnan, 1982). Using a single set of productions, together with a dual lexicon representing grammar fragments of English and Korean, the model is able to parse complex sentences in both languages, constructing syntactic representations that match human judgements. The model reproduces garden path phenomena reported by English and Korean native speakers, and introduces a cross-linguistic treatment of prosodic breaks to avoid garden-paths during processing. Limitations to the model are discussed, as well as questions that are currently under investigation.

Keywords: ACT-R; Lexical Functional Grammar; sentence processing; parsing; syntactic structure; English; Korean.

Introduction

ACT-R has been used to construct models of human language processing for more than twenty years. The aims of individual models have varied, including exploring the nature of memory for language processing (Anderson, Budiu, & Reder, 2001; Budiu & Anderson, 2004), and the extraction of meaning from text input in real time (e.g Ball, 2011). The preeminent model (R. L. Lewis & Vasishth, 2005, henceforth LV05) demonstrated that the retrieval of memory chunks representing syntactic structure could replicate the processing time-courses for English sentences with differing levels of complexity. The mathematical underpinning of LV05 has been used to extend the model to other languages without the need to create full structural representations; this subsequent work has raised questions about some of the assumptions in the original model (e.g. Jäger, Engelmann, & Vasishth, 2015).

All of the models were designed to address specific research questions, but these design choices reduce the models' generalisability from linguistic and cognitive modelling perspectives. With regard to cognitive modelling, all of the models assume extra working memory capacity in the form of overt or hidden buffers that are each linked to one a set of assumed phrasal categories. Thus in principle the size of additional buffer capacity is limited only by the chosen grammar theory rather than cognitive considerations. Left-corner parsing is a common assumption, but this requires either a stack or additional memory capacity in place of a stack. Assuming some additional capacity may be reasonable, but it is still an open question as to how capacity can be added parsimoniously.

The extent to which speakers' acquired knowledge of grammar affects real-time processing is very much live in psycholinguistics (e.g. S. Lewis & Phillips, 2015). Even theories such as Good Enough Processing (Ferreira, Bailey, & Ferraro, 2002) require specificity in their descriptions of syntactic and semantic processing. For cognitive models to address this question productively, their representations of syntax, and of the syntax-semantics interface, need to be theoretically grounded and generalisable.

From a linguistic perspective, existing models make questionable assumptions about structural representations and the relationship between syntax and meaning. Previous models have generated binary branching tree structures based on versions of X-bar theory (Jackendoff, 1977). However, non-configurational languages such as Wambaya (Nordlinger, 1998) provide scant evidence of a binary-branching structure. Models that rely on binary trees thus restrict themselves to phenomena from a subset of languages, rather than considering the general human language faculty. The relationship between syntactic structural position and meaning is either stipulated or falls outside the scope of the model, rather than being grounded in linguistic theory.

The pervasive ambiguity of language is a challenge for modellers, and the stimuli chosen for psycholinguistic experiments often complicate the task of modelling. Ambiguity is typically addressed by working with a fragment of a grammar, and by assuming that the model has knowledge unavailable to native speakers, e.g. predicting the structure that will follow a particular sentence opening, or distinguishing in advance between particular types of clause. Models developed in this way sidestep some of ACT-R's architectural constraints, but restrict their generalisability to other phenomena or languages, and weaken their link to linguistic theory.

In the light of these deficiencies, there is still a need for models of language processing that are based on robust theories of grammar beyond constituent structure, and which seek to generalise a processing model across different languages.

Grammar formalism

Lexical Functional Grammar (LFG) is a modular, constraint-based theory that permits accounts of language phenomena across syntax, semantics, information structure, discourse, and sentence prosody. Syntactic constraints are distributed across two levels of representation: c-structure, which represents surface constituent structure governed by language-specific phrase-structure rules; and f-structure, a universal representation of the functional relationships between meaning-bearing elements of the sentence. The sentence-based theory has been expanded to an incremental theory of sentence growth, in which c-structure and f-structure constraints interact to restrict the possibilities for new information to be added to an emerging structure. F-structure is not only language-independent, but is also the base from which semantic and discourse representations are projected. Thus the representation of syntactic structure in the model is based on f-structure.

The model

The model assumes that language-specific grammatical knowledge is encoded in the lexicon and stored in declarative memory. Additional working memory capacity is assumed in the form of three additional buffers, each loosely associated with events (verbs), things (nouns) or qualities (adjectives/prepositions). Incomplete phrases are maintained in working memory and a multibuffer (Salvucci & Taatgen, 2008) allows the processing of embedded clauses. Each new word is attached into structure before the next is read. The combinatorial possibilities for attachment are encoded in procedural memory, where a single production set is used for both English and Korean sentences.

The generated sentence structure is a graph composed of chunks that represent f-structure. Each chunk other than the root has one mother chunk, and a chunk's relationship to its mother is specified as a grammatical function. The structure is not binary-branching: because of this and ACT-R's assumptions on the growth of information held in a chunk, it was not possible to follow LV05 in modelling retrieval effects on processing time-courses. I return to this in the Discussion.

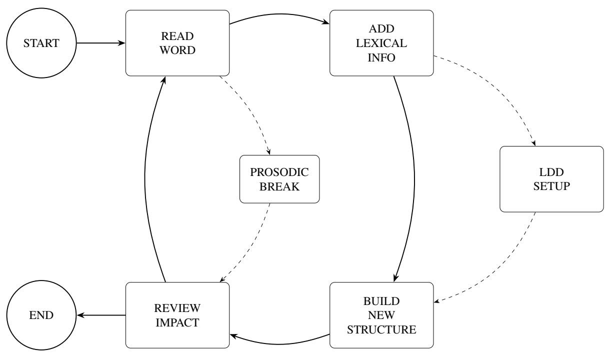

Figure 1: The model parsing cycle

Figure 1 shows the parsing cycle assumed in the model. The model assumes full incremental processing, with only one structural representation maintained. Each word is integrated into the structure before the next is processed. The representation of grammatical functions is based on decomposition into features. An attachment site is chosen by comparing the feature set of the word to be processed is with the unfilled grammatical functions of structural chunks in the buffers. This has two consequences that reflect human behaviour in processing. One is that attachment is context dependent: identical strings are processed differently depending on prior content. The second is that linguistic ambiguities or alternations can be expressed by means of partial feature specifications, allowing flexible combination of words but still generating fully specified output.

The model further assumes that prosodic breaks act to signal the end of a syntactic phrase and force a structural chunk to be cleared from a buffer. This allows the modelling of prosodically-modulated garden path effects.

Results

Results are presented here to illustrate four properties of the model: the ability to analyse complex sentences in English and Korean; the role of feature underspecification in processing structural alternations; context-dependent parsing of identical strings; and the representation of prosodic breaks in guiding the path of structure-building. All figures showing structural representations are given at the end of the paper.

Comparing outputs between languages The model provides an equivalent representation of sentences that have the same meaning in English and Korean, despite the significant variation in constituent structure between the two languages. Korean constituents are strongly head-final, with no fixed ordering of arguments to a verb. Relative clauses precede their head and are marked morphologically at the end of the clause. English has a fundamental SVO word order, with relative clauses appearing after their heads and optionally marked by a complementiser.

Sentences (1) and (2), taken from Kwon, Gordon, Lee, Kluender, and Polinsky (2010), carry broadly equivalent meanings in the two languages (the English sentence lacks an adverb), but different word orders, as can be seen from the gloss in (2). In example sentences, relative clauses are indicated by square brackets.

(1) The conductor [who the famous vocalist invited to the festival] insulted the senator.

(2) [*yumyenghan sengakkaka chwukceney chotayhan*]
[famous vocalist.SUBJ festival.LOC invited]
cihwuycaka uywonul kongkongyenhi
conductor.SUBJ senator.OBJ publicly
moyokhayssta
insulted

"The conductor [who the famous vocalist invited to the festival] publicly insulted the senator."

Figures 2 and 3 show the outputs for the English and Korean sentences respectively. The chunk indices reflect the different word order, and the difference between the chunks 'pro' and 'PRO' reflects the presence or absence of a lexical complementiser in a relative clause[1].

Underspecification The model uses underspecification of grammatical functions to handle lexical ambiguity, without assuming knowledge that is not available to human subjects. The English verb *give* freely alternates between a form that takes a subject, plus two noun phrases as objects, and a form in which the third argument is a prepositional phrase. The choice of argument structure determines the distribution of semantic roles, and until the third argument is processed, its grammatical function, and the resulting semantic roles, are ambiguous. The sentences in (3) show alternate argument structures for the same event. In each case the second and third arguments of *give* are shown in *italics* and **bold** respsectively.

(3) a. The conductor gave *the senator* **a gift**.
    b. The conductor gave *a gift* **to the senator**.

In sentence (3a), the grammatical functions of *give* are $\langle$[SUBJ, OBJ, OBJ$_\theta$]$\rangle$, with OBJ providing the semantic Goal. However in (3b), the grammatical functions are $\langle$[SUBJ, OBJ, OBL$_\theta$]$\rangle$ and OBJ provides the semantic Theme.

Table 1: Underspecification and argument alternation

	−R	+R
−O		OBL$_\theta$
+O	OBJ	OBJ$_\theta$

give	SUBJ	*unrestricted*	
	OBJ	+O	−R
	GF$_3$		+R
gift		+O	
senator		+O	
to		−O	

Table 1 shows the feature system used by the model: the left-hand table gives the full specification for the different grammatical functions using the features ±R 'restricted' and ±O 'object'. The specification on the right shows the features associated with grammatical functions and with words that might fulfil a grammatical function. The third grammatical function of *give* is underspecified as +R, meaning that it could be either OBL$_\theta$ or OBJ$_\theta$. When the word *gift* is encountered in (3a), its feature +O combines with +R to fully specify the grammatical function as [+O +R], giving OBJ$_\theta$. Conversely when *to* is encountered in (3b), the resulting feature set is [−O +R], giving the grammatical function OBL$_\theta$.

Context-dependent parsing The parser successfully uses context to distinguish between different meanings of identical strings. Where a required argument of a verb is missing (a governed grammatical function), the parser preferentially attaches there, rather than provide an adjunct. Consider the English string *the boy the dog bit*, which has two sequential noun phrases *the boy* and *the dog* followed by a verb *bit*. Figure 4 shows that in subject position, *the dog bit* is interpreted as a reduced relative clause. The verb of the main sentence is not yet known, and so there are no required grammatical functions which can be assigned to the second noun phrase *the dog*. Thus *the dog bit* attaches as an adjunct to *the boy*.

In Figure 5, the context contains the main verb *give*, which requires two subsequent arguments. Thus *the dog* is attached as the second object of *give*. This results in an incoherent structure after *bit* is processed, representing the garden path effect observed in human subjects.

Prosodic disambiguation

Sentence (4) produces a garden path effect for Korean speakers because it initially appears that the relative clause includes the first two words of the sentence, *yumyenghan cihwuycaka*, 'famous conductor'.

(4) *yumyenghan cihwuycaka* [*sengakkalul*
famous conductor.SUBJ [vocalist.SUBJ
chwukceney chotayhan] *uywonul kongkongyenhi*
festival.LOC invited] senator.OBJ publicly
moyokhayssta
insulted

"The famous conductor publicly insulted the senator [who invited the vocalist to the festival]."

Figure 6 shows the structure generated by the model for this sentence: although the representation is structurally grammatical, it is semantically incoherent and requires reanalysis to derive the intended meaning.

If a prosodic break is inserted between the second and third words, the garden path effect disappears and the desired meaning is easily accessible to native speakers. The model simulates this by clearing the active buffer, indicating the right edge of a phrase. The output of the model with prosodic support is shown in Figure 7.

Discussion

While previously published ACT-R models of sentence processing have been successful in developing accounts of the role of memory in human sentence processing, there are aspects of them that are theoretically problematic, including unconstrained additional working memory capacity, assumptions of knowledge unavailable to a human subject, and grammar theories and structural representations that are non-standard and not generalisable.

The model presented addresses many of the criticisms of previous models. Its three additional buffers are a limited amount of additional working memory capacity. It can process ambiguous structures without requiring specificity unavailable in human language, and it is based on standard assumptions of phrase structure without relying on distinct representations for relative vs. main clauses (cf. LV05). It can

[1]In Korean *yumyenghan* 'famous' is a verb and morphosyntactically forms a relative clause.

derive similar structures from two typologically different languages, thus separating the processing of a specific language from general cognitive capability. It is also in principle extensible to other languages with different degrees of configurationality. However, it lacks the ability to reproduce processing time-courses and does not use the structural retrieval mechanism successfully proposed by LV05.

There are two main reasons for this. The first arises from interactions between LFG and ACT-R. Functional structure in LFG is not binary branching: all of the grammatical functions associated with a particular word are contained in a single structural unit. Thus any retrieval-based mechanism needs to make multiple retrievals to build a complete structure, adding new information to a chunk at each retrieval. However, adding information to a retrieved and copied chunk means that on release back into declarative memory, it cannot merge with the chunk from which it was copied. There is a proliferation of the chunks containing the information of one f-structure, each representing a different stage of the emerging structure, and these chunks interfere with subsequent retrievals, causing non-human-like errors.

The second reason is more general, arising from the ambiguity inherent in language. It is often unclear whether or not a word attaches into an existing phrase or starts a new phrase. The model presented manages this by holding incomplete phrases in a buffer. Thus if attachment to the chunk in the buffer is possible, the new word attaches there, and if not, or if the buffer is empty, a new structural chunk is created to allow attachment. In a retrieval-based model, chunks would not be maintained in buffers, and so it will only become apparent after a retrieval failure that a new chunk must be created. In ACT-R, this serial process will result in time-course effects that are not seen in human processing data.

Research in progress

Both of the barriers mentioned require changes to architectural assumptions in ACT-R. Work is in progress to develop a language-specific module based on amended assumptions in two areas.

The first area relates to the behaviour of language structure chunks on release into declarative memory. The aim is to allow chunk merger not only for identical chunks, but also where a chunk has added information monotonically compared to the chunk from which it was copied. This addresses the problem of chunk proliferation and the consequent non-human errors in structural analysis. The second area of work is to allow a buffer request *retrieve-or-create*, that either retrieves a chunk against a specification, or creates a new chunk in the case of retrieval failure, without requiring two separate productions. This increases the capacity of models to process ambiguous structures without assuming prescient knowledge, without adding unnecessary processing steps that do not reflect human data.

Once complete, the model will be in a position to generate time-courses that are testable against human data. However, it was not possible to include outputs in this paper.

Conclusion

Cognitive modelling has a role to play in addressing a live question in psycholinguistics: the extent to which grammatical knowledge is accessed on-line during language processing. To engage effectively in the debate, models must be both grammatically and cognitively robust, and generalisable beyond specific phenomena or specific languages. The model presented here is grammatically robust, and is not restricted to a single language. It has cognitively plausible elements in that it does not include a stack, and it assumes only limited additional cognitive capacity. It is not yet able to model time-courses, which requires the development and testing of a model with different architectural assumptions to core ACT-R. However, it offers a step towards a model of language processing that addresses the deficiencies of previous work.

References

Anderson, J. R., Budiu, R., & 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.

Ball, J. T. (2011). A pseudo-deterministic model of human language processing. In *Proceedings of the cognitive science society* (Vol. 33, pp. 495–500).

Bresnan, J. (1982). *The mental representation of grammatical relations*. Cambridge, Mass: MIT Press.

Budiu, R., & Anderson, J. R. (2004). Interpretation-based processing: a unified theory of semantic sentence comprehension. *Cognitive Science*, *28*, 1-44.

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11-15.

Jackendoff, R. S. (1977). *$\bar{X}$ syntax: A study of phrase structure*. Cambridge, MA.: MIT Press.

Jäger, L. A., Engelmann, F., & Vasishth, S. (2015). Retrieval interference in reflexive processing: experimental evidence from Mandarin, and computational modeling. *Frontiers in Psychology*, *6*(617), 1-24.

Kwon, N., Gordon, P. C., Lee, Y., Kluender, R., & Polinsky, M. (2010). Cognitive and linguistic factors affective subject/object asymmetry: An eye-tracking study of prenominal relative clauses in Korean. *Language*, *86*(3), 546-582.

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

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

Nordlinger, R. (1998). *Constructive case: Evidence from Australian languages*. Stanford, CA.: CSLI Publications.

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review*, *115*(1), 101-130.

Figures

Complex sentences in English and Korean

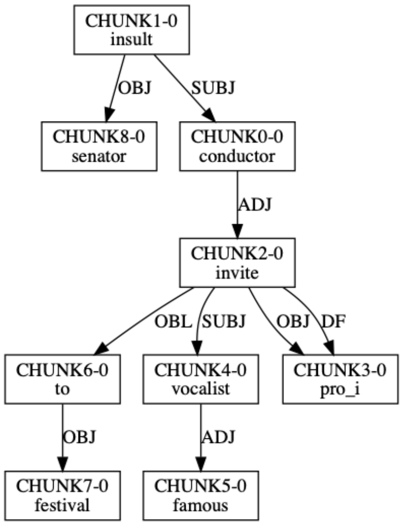

Figure 2: Output from processing sentence (1)

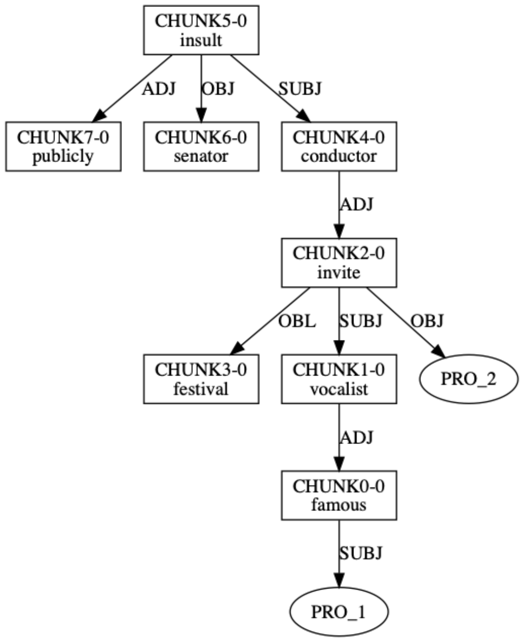

Figure 3: Output from processing sentence (2)

Context-dependent parsing

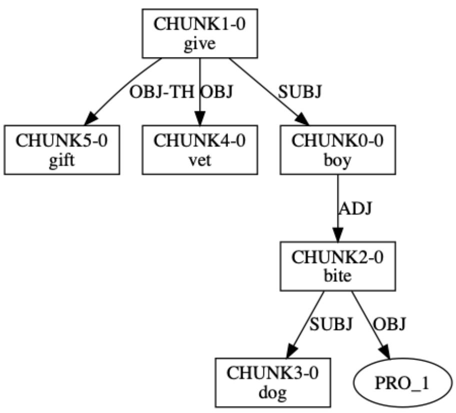

Figure 4: *The boy the dog bit* gave the vet a gift.

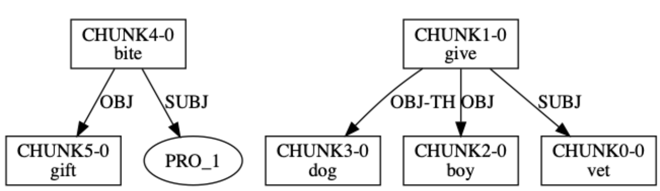

Figure 5: The vet gave *the boy the dog bit* a gift.

Prosodic disambiguation

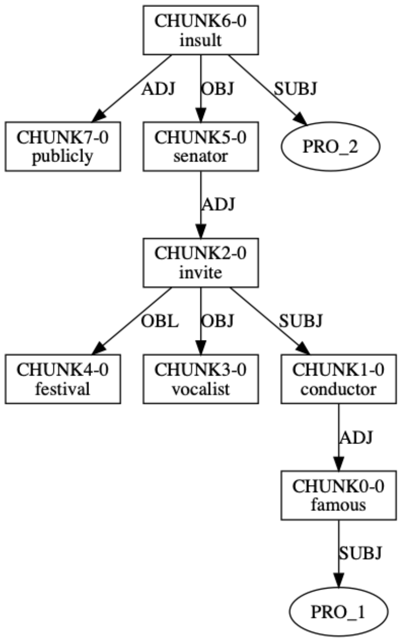

Figure 6: Output from processing sentence (4) *without* prosodic support

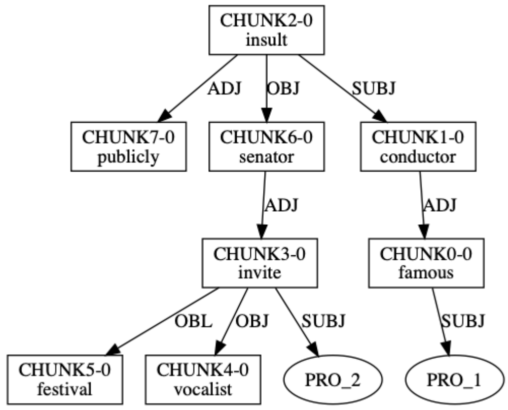

Figure 7: Output from processing sentence (4) *with* prosodic support