In D. Reitter & F. E. Ritter (Eds.), Proceedings of the 14th International Conference on Cognitive Modeling (ICCM 2016). University Park, PA: Penn State.

Encoding and Accessing Linguistic Representations in a Dynamically Structured Holographic Memory System

Dan Parker (dparker@wm.edu)
Daniel Lantz (dflantz@email.wm.edu)
Computational & Experimental Linguistics Laboratory
College of William & Mary, Williamsburg, VA 23187, USA

Abstract

This paper presents a computational model that integrates a dynamically structured holographic memory system into the ACT-R cognitive architecture to explain how linguistic representations are encoded and accessed in memory. We show that a holographic memory system provides a cognitively plausible and principled explanation for the processing of sentences with negative polarity items (NPIs) like *ever* and *any*. The original ACT-R model fails to capture the full range of human reading times and judgments of grammaticality, whereas the integrated holographic memory model achieves good quantitative fits to human error rates and response latencies. These results provide proof-of-concept for the unification of two independent computational cognitive frameworks.

Keywords: Language processing; Memory; Holographic Reduced Representations; ACT-R

Introduction

A hallmark of human cognition is the ability to encode, access, and process compositional structures (Anderson, 1983; Fodor, 2001; Newell, 1990). A parade case involves language processing. For instance, understanding a sentence in a discourse requires mechanisms for encoding a structured representation of the sentence in memory and for accessing specific pieces of information in that representation later. However, it remains an open question how these mechanisms are neuro-computationally instantiated.

One model that has received much attention is the Lewis and Vasishth (2005) (henceforth LV05) model of sentence processing, realized in the Adaptive Control of Thought—Rational (ACT-R) architecture (Anderson, 1990; Anderson et al., 2004). In the LV05 model, sentence processing is construed as a series of cue-based memory retrievals, subject to similarity-based interference. The model is considered the most precise expression of the working memory retrievals and associated control structures that support language processing, and is commonly used to investigate the timing and accuracy of memory retrieval in sentence comprehension.

An initial success of the LV05 model was that it captured interference effects observed in the processing of linguistic dependencies, such as those involving negative polarity items (NPIs). NPIs are words like *ever* or *any*, which are generally acceptable only in sentences that contain a negative-like word in a syntactically higher position, e.g., ***No bills*** *that the senators supported will* ***ever*** *become law*. Previous work has shown that NPI licensing is highly susceptible to interference in sentences like *The bills that* ***no senators*** *supported will* ***ever*** *become law*, due to the presence of the negative distractor *no* that is in a syntactically irrelevant position (e.g., Drenhaus, Saddy, & Frisch, 2005). Interference manifests as decreased accuracy in judgments of grammaticality and decreased reading time disruptions at the NPI, relative to sentences that lack negation. Vasishth, Brüssow, Lewis, and Drenhaus (2008) argued that such effects are a natural consequence of the error-prone memory retrieval mechanisms embodied in ACT-R. Under this view, encountering an NPI triggers a retrieval for a negative licensor, but the wrong item can be retrieved if it matches some of the retrieval cues.

The LV05 model is able to capture many empirical effects, but there are cases where the model makes the wrong predictions. For instance, Parker and Phillips (2014; submitted) showed that NPI interference effects can be reliably switched on and off, depending on when the memory encoding is probed: interference is observed when the encoding of the licensing context is probed early in the sentence, but the effect disappears when the licensing context is probed from a later point in the sentence (see also Parker, 2014). These findings are unexpected under the ACT-R account, which predicts that interference effects should generalize across contexts, based on the assumption that there is a single set of principles that governs memory access.

Parker and Phillips suggested that the contrasting profiles observed for NPIs reflect untested assumptions about how sentences are encoded in memory. ACT-R assumes that the encoding remains fixed over time. However, the finding that interference can be switched on/off depending on when the encoding is probed suggests that the encoding is not fixed, but rather changes over time, such that the internal items become opaque as candidates for causing interference.

This paper presents a computational model that integrates a holographic memory system (e.g., Plate, 2003) into the ACT-R framework to explain the empirically observed effects that the LV05 model fails to capture. Holographic memory systems assume that the atomic components of a compositional structure are periodically bound together in memory to create a single, unitized encoding for interpretation. A key prediction of our model is that interference effects during linguistic dependency formation should be selective, depending on when the encoding is probed. Modeling results show good quantitative fits to a variety of measures, providing proof-of-concept for the unification of two computational cognitive frameworks.

The research reported in this paper builds on previously published literature on holographic memory models and integrating holographic models with ACT-R. Rutledge-Taylor, Kelly, West, and Pyke (2014) and Kelly, Kwok, and West (2015) have shown that a holographic declarative memory system similar to the one proposed here can be integrated into ACT-R to capture decision-making tasks, the fan effect, and delayed learning. Our model demonstrates that this unified framework can capture more specialized cognitive abilities, such as language processing.

The ACT-R model of sentence processing

ACT-R is a cognitive architecture based on independently motivated principles of memory and cognitive skills, and has been used to study a wide range of cognitive phenomena (Anderson, 1990). The LV05 ACT-R model applies those principles to the specialized task of sentence processing.

In the LV05 ACT-R model, linguistic constituents are encoded as 'chunks' in content-addressable memory, and the syntactic representation of a sentence arises as the consequence of pointers that index the hierarchical relations between chunks. Chunks are encoded as bundles of feature-value pairs. Features include lexical content (e.g., morpho-syntactic and semantic features), syntactic information (e.g., category, case), and local hierarchical relations (e.g., sister, parent). Values for features include symbols (e.g., ±singular, ±animate) or pointers to other chunks (e.g., NP_1, VP_2).

Linguistic dependencies, such as those between an NPI and its licensor, are formed using a general retrieval mechanism that probes all task-relevant chunks in parallel for the left part of the dependency (the target), using a set of retrieval cues. Retrieval cues are derived from the current word, the linguistic context, and grammatical knowledge, and correspond to a subset of the features of the target (Lewis, Vasishth, & Van Dyke). Chunks are differentially activated based on their match to the retrieval cues. The probability of retrieving a chunk is proportional to the chunk's overall activation at the time of retrieval, modulated by decay and interference from other items that match the retrieval cues.

The activation of a chunk i (A_i) is defined as follows.[1]

$$A_i = B_i + \sum_{j=1}^{m} W_j S_{ji} - \sum_{k=1}^{p} PM_{ki} + \epsilon \tag{1}$$

The first term of Equation 1 describes the baseline activation of chunk i, which is calculated according to Equation 2. Equation 2 describes the usage history of chunk i as the summation of all n successful retrievals of i, where t_j is the time since the jth successful retrieval of i to the power of the negated decay parameter d. The output is passed through a logarithmic transformation to approximate the log odds that the chunk will be needed given its usage history. After a chunk has been retrieved, the chunk receives an activation boost, followed by decay.

$$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right) \tag{2}$$

The second term of Equation 1 reflects the degree of match between chunk i and the retrieval cues. W is the weight associated with each retrieval cue j, which defaults to the total amount of goal activation G available divided by the number of cues (i.e., G/j). Weights are assumed to be equal across all cues. The degree of match between chunk i and the retrieval cues is the sum of the (weighted) associative boost for each retrieval cue S_j that matches a feature value of chunk i. The associative boost that a cue contributes to a chunk that it matches is reduced as a function of the *fan* of that cue, i.e., the number of chunks in memory that match the cue (Anderson, 1974), according to Equation 3.

$$S_{ji} = S - \ln(\text{fan}_j) \tag{3}$$

The third term of Equation 1 reflects the penalty for a partial match between the cues of the retrieval probe and the feature values of chunk i. Partial matching makes it possible to retrieve a chunk that matches only some of the cues, creating the opportunity for retrieval interference (Anderson et al., 2004; Anderson & Matessa, 1997). Partial matching is calculated as the matching summation over the k feature values of the retrieval cues. P is a match scale, and M_{ki} reflects the similarity between the retrieval cue value k and the value of the corresponding feature of chunk i, expressed by maximum similarity and maximum difference.

Lastly, random noise is added to the activation level of chunk i, generated from a logistic distribution with a mean of 0, controlled by the noise parameter s, which is related to the variance of the distribution, according to Equations 4 and 5.

$$\epsilon \sim \text{logistic}(0, \sigma^2) \tag{4}$$

$$\sigma^2 = \frac{\pi^2}{3} s^2 \tag{5}$$

Activation A_i determines the probability of retrieving a chunk, according to Equation 6. The probability of retrieving chunk i is a logistic function of its activation with gain $1/s$ and threshold τ. Chunks with higher activation are more likely to be retrieved.

$$P(\text{recall}) = \frac{1}{1 + e^{(-A_i - \tau)/s}} \tag{6}$$

[1] Readers familiar with ACT-R may notice the non-standard presentation of Equation 1: the sign on the partial match component has been flipped to indicate its penalizing nature.

Activation A_i also determines the retrieval latency T_i of a chunk, according to Equation 7. F is a scaling factor that sets predictions on an appropriate time scale. Chunks with a higher activation value have a faster retrieval latency.

$$T_i = Fe_i^{-A_i} \quad (7)$$

Predictions of the ACT-R model

The LV05 ACT-R model predicts that retrieval for linguistic dependency formation should be subject to interference from non-target or syntactically irrelevant items that match some of the retrieval cues (partial match interference). This prediction is based on the assumptions that retrieval accesses all chunks in parallel and that a partial match between the retrieval cues and a chunk can result in erroneous retrieval of that chunk (see Equation 1). Many studies have shown that this prediction is borne out for a range of dependencies, including subject-verb agreement (Dillon et al., 2013; Wagers et al., 2009; Tanner et al., 2014), anaphora (Parker et al., 2015), case licensing (Sloggett, 2013), and ellipsis (Martin, 2015).

For instance, the LV05 model has been used to explain interference effects observed in the processing of negative polarity items (NPIs). NPIs are words like *ever*, *any*, or *yet*, that can be licensed by a negative-like word in a syntactically higher position. The NPI *ever* in (2a) is licensed because it appears in the scope of the negative phrase *no students*. When negation is absent, (2b), or is in a syntactically irrelevant position, (2c), the NPI is not licensed.

(2) a. *No students* have *ever* passed the test.
  b. The students have *ever* passed the test.
  c. The students that *no teachers* liked *ever* passed the test.

Previous research has shown that NPI licensing is highly susceptible to interference in sentences like (2c), due to the presence of the negative distractor, e.g., *no teachers*, that is in a syntactically irrelevant position for the purpose of NPI licensing. This effect manifests as decreased accuracy in judgment tasks and decreased reading time disruptions when processing the unlicensed NPI, relative to sentences that lack negation, like (2b).

Vasishth et al. (2008) argued that such effects are a natural consequence of the error-prone retrieval mechanisms embodied in ACT-R. Under this account, NPI licensing is implemented as an item-to-item dependency by retrieving a negative licensor from memory using syntactic and semantic cues, e.g., [+scope], [+negative]. In (2a), retrieval finds an item that matches both cues. In (2b), retrieval fails to find a match to either cue. In (2c), retrieval finds a partially matched item, i.e., a semantically appropriate item in a syntactically irrelevant position. The activation boost from this partial match, combined with stochastic noise, can cause the syntactically irrelevant licensor to be retrieved, spuriously licensing the NPI. Vasishth et al. showed that Equations 1-6 achieve good quantitative fits to both human reading times and judgements of grammaticality.

Challenges for the ACT-R model

The LV05 ACT-R model predicts that interference during NPI licensing should generalize across syntactic environments, since the effect is attributed to error-prone retrieval mechanisms that are engaged whenever an NPI is encountered. However, this prediction is not borne out. Parker and Phillips (2014; submitted) showed that interference effects for NPIs can be reliably switched on/off, depending on when the memory encoding of the licensing context is probed. They manipulated the position of the NPI relative to the potential licensors in sentences like (3), and found contrasting profiles: interference was observed when the NPI appeared early in the sentence, i.e., in the main clause (position 1), replicating previous findings, but not when it appeared later in the sentence, i.e., in the embedded clause (position 2). These effects were shown using both reading time measures and speeded acceptability judgments.

(3) The journalists that no editors recommended (ever_1) thought that readers would (ever_2) understand physics.

These findings suggest that the interference effects observed for NPIs cannot simply be due to noisy retrieval mechanisms that are engaged whenever an NPI is encountered, as assumed in ACT-R. Furthermore, the effects cannot reflect decay or faulty encoding of the licensing context, since that would predict difficulty in the grammatical conditions, contrary to fact.

Parker and Phillips argued that the contrasting profiles observed for NPIs reflect untested assumptions about how sentence representations are encoded in memory. ACT-R assumes that the encoding of the sentence remains fixed over time. However, the finding that interference effects can be switched on/off depending on when the encoding is probed suggests that the encoding is not fixed, but rather changes over time: at one moment, irrelevant items are transparently accessible via partial matching; but then at a later point in time, those same irrelevant items become opaque as candidates for causing interference.

In the next section, we discuss how such effects are predicted in an alternative, dynamically structured holographic memory system.

Multiple-stage encoding schemes

The LV05 ACT-R model assumes that the encoding of a sentence remains fixed over time. However, this is not a widespread assumption. Many cognitive models, including the entire class of Vector Symbolic Architectures (VSAs), e.g., Tensor Product Models (Smolensky, 1990), Holographic Memory (Plate, 2003), Binary Spatter Codes (Kanerva, 1994), assume that there is a qualitative shift over time in the format of an encoding in memory.

In VSAs, compositional structures are encoded in two stages. When a representation is first encoded, it is equivalent to its subparts, such that the individual features of the representation can be evaluated independently from their position in a structured representation, creating the opportunity for partial match interference at retrieval. Then, at a later point, those same features may be bound together, creating a single, unitized encoding that is dissimilar to its sub-parts to conserve memory resources. In this state, individual features are no longer independently evaluable, and the representation must exhibit an all-or-none match to the cues of the retrieval probe in order to be recovered, preventing the possibility of partial match interference. This idea of "recoding" is based on Miller's (1956) principle of chunking, which provides a central explanation for how human memory works.

Proposal

An implicit assumption of VSAs is that compositional structures are encoded in multiple stages. VSAs make a distinction between "atomic" representations that are typically randomly generated versus "complex" compositional representations that are constructed from atomic representations. We propose that these two representational stages may be mapped to distinct cognitive processing stages as a principled explanation of the contrasting profiles observed for NPI licensing. Previously, VSA-based cognitive models have not assumed that particular cognitive processing stages are associated with the two representational schemes. However, if the format of the encoding changes over time, as implicitly assumed in VSAs, then we should expect different behaviors at different points in time, depending on when the encoding is probed, as suggested for NPI licensing.

Encoding linguistic structure in multiple stages

In VSAs, the feature-values of a linguistic representation may be encoded as high-dimensional vectors that are recursively bound together by compressing their outer product into a single vector. For instance, in a tensor-product scheme (e.g., Smolensky, 1990), features are bound together in memory by taking the outer product of the vector representations of the features, as shown in (4).

(4) a. Feature vectors
$[+\text{scope}] = [123]$; $[+\text{negation}] = [\text{abc}]$

b. Tensor-product feature binding

$$\begin{pmatrix}1\\2\\3\end{pmatrix} \otimes \begin{pmatrix}a\\b\\c\end{pmatrix} = \begin{pmatrix}1a & 1b & 1c\\2a & 2b & 2c\\3a & 3b & 3c\end{pmatrix}$$

However, as the structure grows, the size of the code grows exponentially, which is undesirable given the stringent limits on the amount of information that can concurrently occupy working memory (Cowan, 2001). Plate (2003) proposed a solution using Holographic Reduced Representations (HRRs), which rely on circular convolution to bind features together, according to Equation 8.[2] Importantly, the size of the code does not grow as more features are added, since the circular convolution of two n-dimensional vectors using modulo subscripts produces a vector with dimensionality n.

$$t_j = \sum_{k=0}^{n=1} c_k x_{j-k} \tag{8}$$

for $j = 0$ to $n - 1$
(subscripts are modulo-n)

$$\textbf{Binding}_{\boldsymbol{t}} = [+\textbf{scope}]_{\boldsymbol{c}} \circledast [+\textbf{negation}]_{x}$$

$$t_0 = c_0x_0 + c_2x_1 + c_1x_2$$
$$t_1 = c_1x_0 + c_0x_1 + c_2x_2$$
$$t_2 = c_2x_0 + c_1x_1 + c_0x_2$$

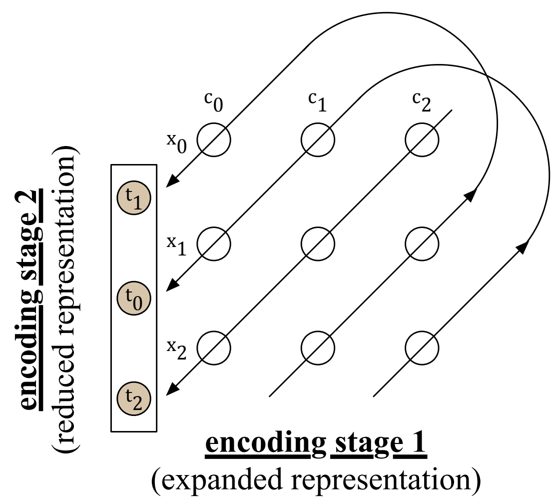

Figure 1. Circular convolution represented as the compressed outer product *t* of the feature vectors *c* and *x*.

Figure 1 shows circular convolution as the ('reduced') outer product *t* of the feature vectors *c* and *x,* corresponding to the linguistic features [+scope] and [+negation] for n=3. Convolution is calculated as the summation of the outer product values along the paths of the lines.

In the uncompressed form (encoding stage 1), individual features *c* and *x* are independently evaluable, making the representation susceptible to partial matching. In the 'reduced' form (encoding stage 2), the individual features *c* and *x* are no longer independently evaluable, preventing the possibility of partial matching. In this state, the representation must be recovered holistically with an all-or-none match to the cues of the retrieval probe.

Similarity between the retrieval probe *p* and a memory *m* measured by their normalized dot product, i.e., cosine similarity, according to Equation 9.

[2] Convolution is the core mathematical operation behind holography, hence the term "holographic".

$$\text{sim}(p,m) = \frac{p \cdot m}{\| p \| \| m \|} = \frac{\sum_{i=0}^{n-1} p_i m_i}{\sqrt{\sum_{i=0}^{n-1} p_i^2}\sqrt{\sum_{i=0}^{n-1} m_i^2}} \quad (9)$$

One concern is that encoding n-dimensional bindings using circular convolution can be slow, since convolution calculates the sum of products (convolution with modulo subscripts takes $O(n^2)$ time). Processing can be sped up by performing convolution in the frequency domain with the Fast Fourier Transform, which involves element-wise multiplication, as shown in Equation 10. This process implements circular convolution in $O(n \log n)$ time.

$$[+\text{scope}]_c \circledast [+\text{negation}]_x = f'(f(c) \odot f(x)) \quad (10)$$

The most important property of HRRs, for present purposes, is that the encoding changes such that the internal items become opaque for partial matching with the passage of time. This property could provide a principled explanation for the contrasting profiles observed for NPIs. If the format of the encoding changes over time, as assumed in a holographic memory system, then we should see different behaviors at different points in time, depending on when the encoding is probed.

In the next section, we show how a holographic memory system can be integrated into the LV05 ACT-R model to simulate human reading times and judgments of grammaticality.

Integrating HRRs into ACT-R

A new memory module for the LV05 ACT-R model was developed using HRRs replacing traditional ACT-R chunks with holographic vectors. Holographic vectors retain the same expressive power of the chunks used in the LV05 model, but allow for dynamic changes in the format of the encoding.

To implement HRRs in the ACT-R system, we made the following changes to the original LV05 ACT-R model. First, linguistic feature-value specifications and retrieval cues were encoded as vectors (one dimensional arrays) of n numbers, randomly sampled from a normal distribution. For our simulations, n = 10,000. In this format, different feature-value specifications and the corresponding retrieval cues are represented by different patterns in a continuous, high-dimensional space.

In encoding stage 1 (expanded representation), feature-value pairs are superimposed by adding the vectors together to create linguistic chunks (bundles of feature-value pairs, as defined in the original LV05 ACT-R model). Retrieval probe vectors are constructed in the same manner. In this state, the individual features of a chunk are independently evaluable at retrieval and hence susceptible to partial matching, as assumed in the original LV05 model.

In encoding stage 2 (reduced representation), convolution as computed according to Equation 10 is used to bind the vectors representing the feature-value pairs within a chunk. To enable successful retrieval of a chunk, the cues of the retrieval probe must be combined in the same way. In this state, a chunk represents a single, unitized encoding that must exhibit an all-or-none match to the retrieval probe to be recovered, i.e., partial matching is not possible. For present purposes, we assumed that feature binding was triggered upon encountering the main clause verb of a sentence during comprehension. According to Parker and Phillips (submitted), encountering a main clause verb may force the parser to 'wrap-up' and consolidate the encoding of the previous context to conserve memory resources.

Second, we modified the standard ACT-R equation for activation values (Equation 1) to accommodate HRR vectors. Specifically, we substituted cosine similarity, as computed according to Equation 9, for the third term of the standard ACT-R equation for activation value, i.e., the term that computes the penalty for a partial match between the cues of the retrieval probe and the feature values of chunk i.

Simulations

We investigated whether the contrasting profiles observed for NPIs would be best captured by the original LV05 ACT-R model or the integrated HRR/ACT-R model. To achieve this, we conducted a side-by-side comparison of the LV05 model with the integrated model, without adjusting key model parameters.

Procedure

Previous implementations of the ACT-R model of sentence processing have included a wide range of modules, including modules for visual information processing, lexical access, memory retrieval, and syntactic parsing (e.g., Lewis & Vasishth, 2005; Vasishth et al., 2008). However, the simulations reported here focus solely on the module for retrieval, and abstract away from the contribution of the peripheral modules by stipulating the chunks in memory and retrievals required to parse a sentence. There are additional processes associated with sentence comprehension that contribute to behavioral measures, but for current purposes, we adopt the standard assumption that the dynamics and output of memory retrieval map monotonically to the behavioral measures of interest (Anderson & Milson, 1989).

We simulated the hypothesized retrievals involved in the key manipulations reported in Parker and Phillips (submitted). Three conditions were simulated, manipulating the presence and location of an NPI licensor (appropriate licensor, irrelevant licensor, no licensor) and the position of the NPI (main clause, embedded clause), based on the sentence structures in (3). For each condition, a schedule of constituent creation times and retrievals was estimated from the reading times reported in Parker and Phillips (submitted). Differences between conditions were modeled only as differences in NPI position and the feature composition of the licensors (±scope, ±negation).

To ensure that the modeling results for the LV05 and integrated HRR/ACT-R model would be directly comparable, all models used the same default parameter settings, following Lewis and Vasishth (2005) and Vasishth

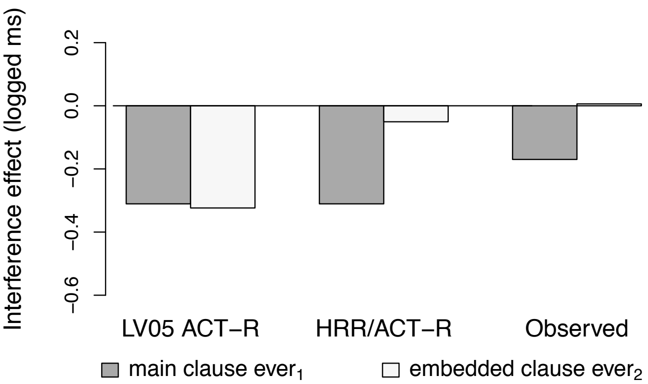

Figure 1. Comparison of predicted and observed interference effects for reading time measures of main clause *ever*$_1$ vs. embedded clause *ever*$_2$.

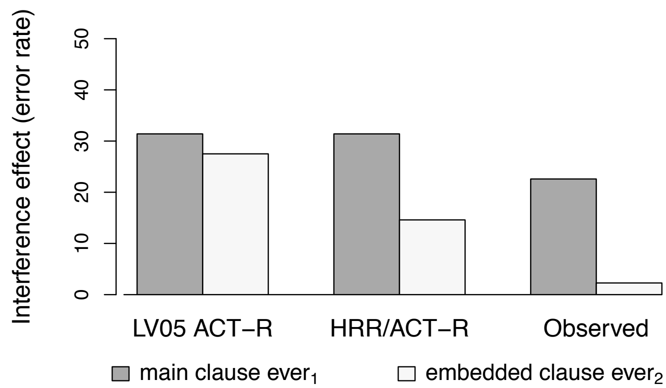

Figure 2. Comparison of predicted and observed interference effects in judgment accuracy for main clause *ever*$_1$ and embedded clause *ever*$_2$.

et al. (2008). The only exception was the scaling parameter *F*, which was optimized to fit the behavioral time scale (in all models, $F = 0.6$). 5,000 Monte Carlo simulations were run for each condition.

We report two measures of interest: (i) Retrieval error rate reflects the percentage of runs for which the distractor, rather than the target was retrieved. This measure maps monotonically to speeded acceptability judgments, with higher retrieval error rates corresponding to increased rates of judgment errors. (ii) Retrieval latencies reflect the average amount of time it took to retrieve the most probable item, and map monotonically to reading times, with higher latencies corresponding to longer reading times. These measures were used to calculate the predicted interference effect as the difference in predicted error rates and retrieval latencies between the ungrammatical conditions with and without a negative distractor (NPI interference is observed only in ungrammatical conditions). Thus, for predicted error rates, a larger positive value corresponds to a higher rate of interference, reflecting increased rates of acceptance for sentences with a distractor relative to sentences with no distractor. For predicted retrieval latencies, a smaller negative value corresponds to a higher rate of interference, reflecting facilitated processing for sentences with a distractor relative to sentences with no distractor.

We compared the observed interference effects with those predicted by the LV05 model and the HRR/ACT-R model for the reading time measures (Figure 1) and judgment data (Figure 2) reported in Parker and Phillips (2014; submitted).

Simulation results

Across both behavioral measures, the integrated HRR/ACT-R model provided a better fit to the observed data, without adjusting the key model parameters (fit with the HRR/ACT-R model was adjusted R^2 = 0.79; fit with the LV05 model was adjusted R^2 = 0.28). The LV05 model failed to capture the observed on/off behavior, predicting similar rates of interference across NPI positions. The integrated model, on the other hand, captured the basic contrast between NPI positions, with significantly less interference for embedded clause NPIs (*ever*$_2$).

Although the values predicted by the integrated HRR/ACT-R model did not match the observed data perfectly, the predicted profiles were qualitatively similar to the observed data. We could explore different parameter values to achieve an even better fit with the observed data, but this was not our goal. Rather, our goal was to determine whether the ACT-R model enhanced with a holographic declarative memory system would predict the basic contrasts without adjusting previously fixed parameter values.

The contrasting profiles predicted by the HRR/ACT-R model are consistent with the hypothesis that the contrasting profiles observed for NPIs reflect changes over time in the encoding of compositional representations in memory. After the features of the representation are bound together, the representation must exhibit an all-or-none match to the cues of the retrieval probe, preventing partial match interference.

Conclusion

We presented a computational model that integrates a holographic memory system into the ACT-R model of sentence processing to explain how compositional linguistic structures are encoded and accessed in memory. Modeling results showed that the integrated system is better suited to capture contrasting profiles of interference effects in sentence comprehension, relative to existing models, yeilding a good quantitative fit to data from a variety of behavioral tasks. These results provide proof-of-concept for the unification of two independently developed computational cognitive frameworks, and offer new insights into how humans encode and access compositional representations in memory.

Acknowledgments

The code for the LV05 ACT-R model of sentence processing was generously provided by Rick Lewis. We thank Alan Du for his revisions and additions to this code. We thank Brian Dillon, Dave Kush, Colin Phillips, and Matt Wagers for helpful discussion related to this work.

References

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, *6*, 451-474.

Anderson, J. R. (1983). *The architecture of cognition*. Harvard University Press.

Anderson, J. R., & Milson, R. (1989). Human memory: an adaptive perspective. *Psychological Review*, *96*, 703-719.

Anderson, J. R. (1990). *The Adaptive Character of Though*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036-60.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87-185.

Dillon, B., Mishler, A, Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*, 85-103.

Drenhaus, H., Saddy, D., & Frisch, S. (2005). Processing negative polarity items: When negation comes through the backdoor. In S. Kepser, & M. Reis (Eds.), *Gradience in grammar: Generative perspectives*. NY: Oxford University Press.

Fodor, J. A. (2001). Language, Thought, and Compositionality. *Mind & Language*, *16*, 1-15.

Kanerva, P. (1994). The binary spatter code for encoding concepts at many levels. In M. Marinaro, & P, Morasso (Eds.), *ICANN '94: Proceedings of the International Conference on Artificial Neural Networks*, London, Springer-Verlag.

Kelly, M. A., Kwock, K., & West, R. L. (2015). Holographic Declarative Memory and the Fan Effect: A Test Case for A New Memory Module for ACT-R. In the Proceedings for the 2015 International Conference on Cognitive Modeling (ICCM).

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

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Science*, *10*, 447-454.

Martin, A. E. (2015). Cue-based interference from illicit attractors: ERP evidence from VP ellipsis. Poster at the 2015 Architectures and Mechanisms for Language Processing (AMLAP) Conference.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

Newell, A. (1990). *Unified theories of cognition: The William James lectures*. Harvard University Press.

Parker, D. (2014). The cognitive basis for encoding and accessing linguistic structure. Unpublished doctoral dissertation.

Parker, D., & Phillips, C. (2014). Time heals semantic illusions, but not syntactic illusions. Talk at the 27th CUNY Conference on Human Sentence Processing.

Parker, D., Lago, M., & Phillips, C. (2015). Interference in the processing of adjunct control. *Frontiers in Psychology*, *6*, 1-13.

Parker, D., & Phillips, C. (submitted). Linguistic illusions and the encoding of compositional representations.

Plate, T. (2003). *Holographic Reduced Representation: Distributed Representation for Cognitive Structures*. CA: CSLI Publications.

Rutledge-Taylor, M. F., Kelly, M. A., West, R. L., & Pyke, A. A. (2014). Dynamically structured holographic memory. *Biologically inspired Cognitive Architectures*, *9*, 9-32.

Sloggett, S. (2013) Case licensing in processing: Evidence from German. Poster at the 26th CUNY Conference on Human Sentence Processing.

Smolensky, P. (1990). Tensor Product Variable Binding and the Representation of Symbolic Structures in Connectionist Systems. *Artificial Intelligence*, *46*, 159-216.

Tanner, D., Nicol, J., & Brehm, L. (2014). The time-course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction. *Journal of Memory and Language*, *76*, 195-215.

Vasishth, S., Brüssow, S., Lewis, R. L., Drenhaus, H. (2008). Processing Polarity: How the Ungrammatical Intrudes on the Grammatical. *Cognitive Science*, *32*, 685-712.

Wagers, M., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*, 206-237.