Cognitive Science 32 (2008) 685–712

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1080/03640210802066865

Processing Polarity: How the Ungrammatical Intrudes on the Grammatical

Shravan Vasishth[a], Sven Brüssow[b], Richard L. Lewis[c], Heiner Drenhaus[a]

[a]*Institute for Linguistics, University of Potsdam*
[b]*Department of Psychology, University of Heidelberg*
[c]*Department of Psychology, University of Michigan*

Received 13 November 2006; received in revised form 14 July 2007; accepted 14 July 2007

Abstract

A central question in online human sentence comprehension is, "How are linguistic relations established between different parts of a sentence?" Previous work has shown that this dependency resolution process can be computationally expensive, but the underlying reasons for this are still unclear. This article argues that dependency resolution is mediated by cue-based retrieval, constrained by independently motivated working memory principles defined in a cognitive architecture. To demonstrate this, this article investigates an unusual instance of dependency resolution, the processing of negative and positive polarity items, and confirms a surprising prediction of the cue-based retrieval model: Partial-cue matches—which constitute a kind of similarity-based interference—can give rise to the intrusion of ungrammatical retrieval candidates, leading to both processing slow-downs and even errors of judgment that take the form of illusions of grammaticality in patently ungrammatical structures. A notable achievement is that good quantitative fits are achieved without adjusting the key model parameters.

Keywords: Computational modeling; ACT–R; Eye tracking; Reading; Sentence processing

1. Introduction

The act of comprehending a sentence triggers a complex set of rapid cognitive processes that engage multiple memory systems. Minimally, contact must be made with a long-term lexical memory, novel compositional structures incrementally created and maintained in a working memory, local and global ambiguities resolved at multiple levels of linguistic representation, and an interpretation of the sentence constructed that is integrated into a referential representation of the current discourse. A key process during all this is the integration of incoming lexical elements with the partial sentence-level structure built so far. Such integrations

Correspondence should be sent to Shravan Vasishth, Institute for Linguistics, University of Potsdam, Postfach 601553, D–14415, Potsdam, Germany. E-mail: vasishth@acm.org

are not instantaneous or cost free, and many different theories have been proposed to explain their psychological properties. The ultimate goal of these theories is to provide insight into the fundamental properties of the linguistic working memory systems that support the rich combinatorial capacity of human language.

For example, in syntactic ambiguity resolution, the motivation behind principles like minimal attachment and late closure is that they serve to systematically reduce the parser's working memory load (Frazier, 1979, p. 39). In fact, the costs of incremental integration can be characterized independently of classic ambiguity problems. Incremental integrations are necessary to create grammatically licensed linguistic relations, or *dependencies*. Dependencies are a pervasive property of language: Linguistic elements such as noun-phrase arguments depend on verbs, pronouns and reflexives depend on antecedents, gaps depend on their fillers. These dependencies must be resolved in order to build an interpretation of an event, or to resolve reference of pronominal and null elements. Here, a central question of interest is, "How are dependents integrated with each other?" Answering this question is fundamental to understanding the nature of working memory in human sentence comprehension.

Chomsky (1965, pp. 13–14) was among the first to propose that the reduced acceptability of sentences containing a "nesting of a long and complex element" arises from "decay of memory." In related work, Just and Carpenter (1980, 1992) directly addressed dependency resolution in sentence comprehension in terms of memory retrieval (similar early approaches are the production–system-based models of Anderson, Kline, & Lewis, 1977). Just and Carpenter developed a model of integration that involved activation decay (as a side-effect of capacity limitations) as a key determinant of processing difficulty. For example, under the rubric of distance effects, they described the constraints on dependency resolution as follows: "The greater the distance between the two constituents to be related, the larger the probability of error and the longer the duration of the integration process" (1992, p. 133).

The activation decay idea as a determinant of dependency resolution difficulty was taken a great deal further in the Syntactic Prediction Locality Theory (SPLT; for a historical overview of the connection between decay and distance, see Gibson, 1998, p. 9) and, more recently, the Dependency Locality Theory (DLT; Gibson, 2000). The DLT proposes (among other things) that the cognitive cost of assembling a dependent with a head is partly a function of the number of new intervening discourse referents that were introduced between them. Another related theory is Early Immediate Constituents (EIC; Hawkins, 1994), which assigns a greater processing cost when there is an increase in the number of words that make up a syntactic constituent. The SPLT and DLT, in particular, have yielded a rich body of experimental research that provides strong support for the existence of distance effects in English (e.g., Gibson & Thomas, 1999; Grodner & Gibson, 2005; Warren & Gibson, 2005).

In recent work, another approach to the dependency resolution issue has been proposed (Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006; Van Dyke & Lewis, 2003; Vasishth & Lewis, 2006). This theory differs from previous accounts in that instead of defining constraints on retrieval in terms of linguistic primitives such as the number of intervening new discourse referents (DLT) or the number of words per constituent (EIC), the cognitive costs of dependency resolution are derived from an independently motivated theory of working memory retrieval: Dependents are retrieved through a content-based retrieval process that relies on cues expressed as feature-value specifications, and retrieval difficulty emerges from the dynamic interaction of constraints on working memory processes, including especially interference and decay. This mechanism has been shown in previous work to account for a range of self-paced reading data from languages like English and Hindi (Lewis & Vasishth, 2005; Van Dyke & Lewis, 2003; Vasishth & Lewis, 2006).

In the work we present below, we confirm a surprising prediction of the cue-based retrieval model that distinguishes it from the theories mentioned above: Partial-cue matches, which constitute a kind of similarity-based interference, can give rise to the intrusion of ungrammatical retrieval candidates, leading to both processing slow-downs and even errors of judgment that take the form of illusions of grammaticality in patently ungrammatical structures (for related work on similarity-based interference, see Gordon, Hendrick, & Johnson, 2001, 2004; Gordon, Hendrick, Johnson, & Lee, 2006; Gordon, Hendrick, & Levine, 2002; Lewis, 1996; Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006). The specific dependency resolution problem that we focus on here is one that has received little psycholinguistic attention but is of considerable interest in linguistic theory: the processing of polarity items.[1]

1.1. Polarity licensing dependencies

Negative polarity items (NPIs), such as the adverb *ever,* are usually licensed only when they appear in some kind of "negative context" like *no man*; compare 1a and 1c in the following example. Specifically, in a structure such as 1b, mere linear precedence of the licensor is not good enough: The licensor must "c-command" the NPI *ever*. Formally, a node *A* c-commands another node *B* if, and only if, *A* does not dominate *B*; and every node *X* that dominates *A* also dominates *B* (Reinhart, 1981); a node *P* dominates another node *Q* if *P* occurs at a depth higher than *Q* and a path exists from *P* to *Q* (the depth of a node from the root is the number of vertices traversed exactly once from root to node). As an illustration, consider Example 1a; here, *No man* c-commands *ever,* but *a beard* does not:

1. a. [DP **No man** [who had a beard]] was **ever** thrifty.
   b. *[DP **A man** [who had **no** beard]] was **ever** thrifty.
   c. *[DP **A man** [who had a beard]] was **ever** thrifty.

The same constraint applies to German *jemals* "ever": In Example 2a, the licensor *Kein Pirat* c-commands the NPI; in 2b, the licensor *keinen Braten* occurs in a structural location (inside the relative clause modifying the determiner phrase [DP]) that does not c-command the NPI; and in 2c, there is no licensor at all:

2. a. Accessible NPI licensor:
   Kein Pirat, [der einen Braten gegessen hatte,] war **jemals** sparsam
   No pirate who a roast eaten had was ever thrifty
   "No pirate who had eaten roast (meat) was ever thrifty."
   b. Inaccessible NPI licensor:
   Ein Pirat, [der **keinen Braten** gegessen hatte,] war **jemals** sparsam
   A pirate who no roast eaten had was ever thrifty
   "A pirate who had eaten no roast (meat) was ever thrifty."

c. No NPI licensor:
Ein Pirat, [der einen Braten gegessen hatte,] war **jemals** sparsam
A pirate who a roast eaten had was ever thrifty
"A pirate who had eaten roast (meat) was ever thrifty."

Much controversy surrounds the precise constraints operating on negative polarity licensors (e.g., see Baker, 1970; Chierchia, 2006; Fauconnier, 1975a, 1975b; Giannakidou, 1998; Horn, 2001; Israel, 2006; Krifka, 1995; Ladusaw, 1980; Linebarger, 1987; Szabolcsi, 2004; van der Wouden, 1997). However, for the above examples, it can be argued that 1b, 1c, 2b, and 2c violate the c-command constraint on NPIs.[2] It is therefore surprising that a speeded grammaticality judgment task (Drenhaus, Saddy, & Frisch, 2005) showed an asymmetry in the judgments for the two ungrammatical sentences 2b versus 2c: Participants were significantly worse at judging 2b as ungrammatical (Drenhaus et al., 2005).

In the Drenhaus et al. (2005) experiment, participants saw the matrix DP, the embedded DP, and each of the other words in isolation for 300 msec each. A blank screen was presented for 100 msec between each presentation; 500 msec after the last word of the sentence was presented, participants had to judge the acceptability of the sentence within a maximum of 3,000 msec; 1,000 msec after their response, the next trial was presented. The essential finding was that a linearly preceding but structurally inaccessible licensor can sometimes result in an illusion of grammaticality. Drenhaus et al. referred to this as the "intrusion effect."[3]

1.2. The intrusion effect

Table 1 summarizes the percentage accuracies and reaction times of the Drenhaus et al. (2005) study. The percentage of correct grammaticality judgments for the inaccessible NPI licensor condition (2b) was significantly lower than for the felicitous condition (2a) and the ungrammatical condition (2c) that did not have an inaccessible licensor.

In other words, participants made significantly more errors in judging the inaccessible licensor condition ungrammatical. The mean reaction times in the inaccessible licensor condition (2b) was also slower than in other conditions.

Similar results were found in a replication of the speeded judgment task. This replication was conducted as part of an event-related potentials (ERP) study (Drenhaus et al., 2005). The main finding was that, compared to the grammatical condition (2a), both the inaccessible licensor (2b) and no licensor (2c) conditions showed N400 and P600 components at the

Table 1
Judgment accuracies and reaction times in the (Drenhaus, Saddy, & Frisch, 2005) experiment

Condition	Accuracy (% Correct)	Speed (in Milliseconds)
(2a) Accessible licensor	85	540
(2b) Inaccessible licensor	70	712
(2c) No licensor	83	554

NPI jemals. Because the N400 component, in general, reflects semantic integration problems and violations of selectional restrictions and implausibility (Kutas & Petten, 1994); and because the P600 component reflects syntactic reanalysis and repair (Friederici, 1995, 2002), increased syntactic complexity and ambiguity (Friederici, Hahne, & Saddy, 2002; Frisch, Schlesewsky, Saddy, & Alpermann, 2002; Kaan, Harris, Gibson, & Holcomb, 2000), the results suggested that an NPI occurring in an illegal environment results in both semantic and syntactic processing problems compared to their licensed counterparts (also see Drenhaus, beim Graben, Saddy, & Frisch, 2006).

1.3. Explaining the intrusion effect

It is likely that the intrusion effect is due to a processing problem: It does not appear to have an explanation in linguistic theory, which, in general, can only provide categorial (and deterministic) predictions about the ungrammaticality of both the inaccessible licensor condition (2b) and the no licensor condition (2c). To our knowledge, there does not exist any competence theory of polarity licensing (nor any implemented computational model thereof) that can generate probabilistic, non-deterministic grammaticality decisions, which is a prerequisite for explaining the Drenhaus et al. (2005) accuracy patterns.

Notice that the processing of structures like Example 2 are an instance of the dependency resolution problem: The licensor and NPI need to be integrated in order for the sentence to be comprehended and judged grammatical. However, providing a processing explanation of the effect is a challenge; a complete account would have to (a) explain why errors occur in speeded judgments, (b) provide an interpretation of the N400 and P600 components, and (c) deliver quantitative predictions about moment-by-moment processing costs.

The contribution of this article is twofold. First, we demonstrate the occurrence of the intrusion effect in an eye-tracking reading study. Second, we show that the cue-based retrieval model, which is an independently motivated computational model of sentence processing (Lewis & Vasishth, 2005; Lewis et al., 2006; Vasishth & Lewis, 2006), can account for the grammaticality–judgment accuracy patterns in the intrusion effect, as well as eye-tracking dependent measures at the polarity item.

One noteworthy fact about the model is that previously fixed numerical parameters are used to fit an entirely new set of behavioral data, demonstrating the model's robustness. Another important point is that the underlying behavior of the model emerges from independently developed and empirically motivated principles of working memory realized within a cognitive architecture (ACT–R; Anderson et al., 2004). The model therefore demonstrates the central role of domain-independent working memory principles involved in a highly specialized and skilled information processing activity—sentence comprehension. By tightly specifying the relation between memory processes and parsing, a detailed picture emerges of human sentence comprehension grounded in the cognitive system.

In the remainder of the article, we present the theory and then its application to the intrusion effect. Then, an eye-tracking experiment is described that further demonstrates the robustness of the intrusion effect in a more natural experimental setting than speeded judgment tasks. Finally, we discuss the effectiveness of the model in explaining the intrusion effect and compare the model with other theories of sentence processing.

2. Cue-based retrieval in parsing

The computational model and its current empirical coverage are described in detail elsewhere (Lewis & Vasishth, 2005; Lewis et al., 2006; Vasishth & Lewis, 2006). Here we present its major features before turning to the model of the intrusion effect. The complete source code of the model will be made available online upon publication of this article. The model has two parts, a symbolic and a subsymbolic component, which we discuss next.

2.1. The symbolic component: chunks and productions

Long-term lexical information is encoded in a long-term declarative system, and grammatical knowledge is held in procedural form as a set of specific condition–action associations (production rules in ACT–R). This procedural memory simultaneously represents the grammar and the knowledge of how to apply it to incrementally parse sentences.

The declarative memory also maintains the unfolding representation of the novel structure of the sentence. This working memory system consists of a sharply limited focus of attention (represented as a limited set of buffers in ACT–R), along with declarative memory elements that are in a high state of activation as a result of being recently created or processed. Critically, processing is driven only by those memory elements in the focus of attention; in ACT–R, this constraint corresponds to the limitation that production rules only match against chunks in the limited set of buffers (described briefly below). This basic architecture for working memory is consistent with recent proposals in cognitive psychology that distinguish a severely limited focus from a penumbra of memory elements that are highly active but must nevertheless be retrieved into a focused state in order to affect processing (Cowan, 2001; McElree, 2006; Oberauer, 2002).

Each lexical item is assumed to be available in long-term declarative memory as a set of *chunks* (Miller, 1956), which are represented in the model as feature-value specifications not unlike those used in head-driven phrase structure grammar (Pollard & Sag, 1994). Elements in both long-term declarative memory and working memory are chunks. Thus, apart from lexical items, non-terminal nodes are also chunks; and, through these feature descriptions, the sub-parts of a tree are assembled into a parse tree. As shown in Fig. 1, for example, a parse of

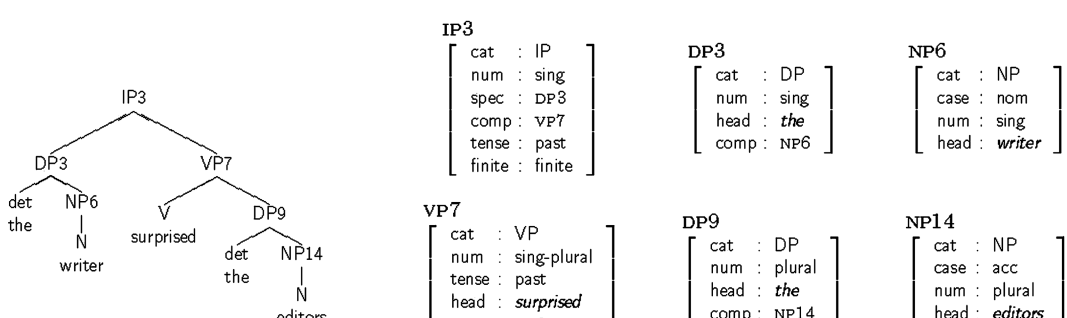

Fig. 1. Chunks in memory corresponding to the sentence, “The writer surprised the editors.” *Note:* IP = inflectional phrase; DP = determiner phrase; VP = verb phrase; NP = noun phrase; V = verb; N = noun.

the sentence, "The writer surprised the editors," is simply a collection of chunks representing terminals and non-terminals that are interlinked by feature-value specifications: The chunk DP3 *the writers* is the value of the specifier feature of sentence level node IP3, and so on. Chunks corresponding to lexical items are stored permanently in memory. In addition, the model creates temporary chunks at runtime that encode non-terminal nodes interconnected as described above.

The production rules effectively implement a parser that drives the retrieval, integration, and construction of the chunks that eventually constitute a parse tree such as Fig. 1. The parsing steps are defined by the production rules (Anderson et al., 1977; Anderson & Lebiere, 1998, p. 6; Newell, 1973): If certain conditions hold, a production fires and certain actions are triggered. The conditions that trigger production-rule firing are defined in terms of patterns of buffer contents, and the actions are defined in terms of changes to those buffer states. These changes in state can, in turn, lead to the firing of other productions; this goes on until the processing task is completed.

We turn now to the buffers defined in ACT–R; we discuss only those relevant to the sentence processing model. In general, buffers in ACT–R have the property that they may hold only a single chunk. The buffers relevant for the model are the *goal buffer* and the *retrieval buffer*. The former serves to represent the current control state information, whereas the latter serves as an interface to declarative memory. A retrieval is carried out when a production fires that sets retrieval cues in the retrieval buffer; a retrieval occurs if these cues sufficiently match a chunk in declarative memory that has sufficient activation (described below). As an example, consider the situation where a transitive verb like *drank* is being processed. Because this verb requires an animate subject noun phrase, integration of the verb with the appropriate DP would require a retrieval that asks for a noun-phrase chunk with those properties ([+ animate, + nominative]). As described in Lewis and Vasishth (2005), the parsing actions encoded in the production rules emulate a left-corner parsing strategy (Johnson-Laird, 1983; Resnik, 1992).

In summary, sentence processing consists of an iterative sequence of retrievals, all guided by the grammatical knowledge encoded in the production rules. We now focus on the properties of the model that govern the nature of working memory retrieval.

2.2. *The subsymbolic component*

Apart from the symbolic system (i.e., the memory structures and the procedural rules constituting the parser), which is responsible for structure building and representation, the model's behavior depends crucially on constraints imposed on the retrieval of chunks in memory. These constraints are defined in terms of a set of subsymbolic computations that affect the activation of chunks.

2.2.1. *Activation and the base-level learning equation*

A key assumption is that retrieval probabilities and latencies are governed by activation levels, which fluctuate as a function of frequency, recency, and pattern of prior exposure. Anderson and Schooler (1991) originally explored this issue with respect to the pattern of past information presentation (prior exposure to an item), and provided a rational, functional motivation for a certain class of decaying activation functions. Across a range of task domains, these activation functions corresponded well to the empirical probability that a past item would actually be needed at some point in time.

Equation 1 is the current formalization of activation in ACT–R. It determines the base-level activation of a chunk i where t_j is the time since the j th retrieval of that chunk. The summation over all n retrievals results in the current activation of chunk i:

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

This equation essentially describes the past usage history of some chunk i in terms of the number of n successful retrievals (presentations) of i. Here, t_j is the time since the jth successful retrieval of i. The decay parameter d, in general, has the default value of 0.5. The summation for all n presentations of t_j to the power of the negated decay parameter is passed through a logarithmic transformation to yield the base-level activation value. Equation 1 thus describes an asymptotic function that in case of frequent presentations of a chunk results in an increase of its otherwise decaying activation. We refer to this hereafter as *activation boost*. After a chunk has been retrieved, it experiences an activation boost and then decay immediately sets in.

Activation A_i affects both the probability of the chunk's i retrieval and its retrieval latency. The higher the activation, the faster a chunk can be retrieved from declarative memory and placed into working memory. The mapping from activation to retrieval latency is accomplished by Equation 2, where the retrieval latency T_i of a chunk i is the time it takes to retrieve that chunk from declarative memory and make it available in the retrieval buffer. F is a scaling constant that varies across ACT–R models, although typically within a limited range. In the present model, the parameter was adjusted by visual inspection to 0.46 in order to fit the dependent measures (in previous simulations—Lewis & Vasishth, 2005; Vasishth & Lewis, 2006—its value was 0.14):

$$T_{\mathrm{i}} = F e_i^{-A} \tag{2}$$

In addition, a retrieval threshold value is defined for chunks. This value determines the minimum activation each chunk has to bear in order to be, in principle, retrievable. If a chunk was not retrieved for a period of time, and the available retrieval cues are insufficient to boost its activation past threshold, the chunk will not be retrieved. In the simulations presented here, the retrieval threshold is kept fixed at -2, the same value that was used in Lewis and Vasishth (2005) and Vasishth and Lewis (2006).

2.2.2. *Associative retrieval and similarity-based interference*

Apart from the computation of the base-level activation, other factors contribute to a chunk's overall activation. For any set of retrieval cues, all chunks that have feature values (hereafter, we use the term *slot values* for consistency with the ACT–R literature; Anderson et al., 2004) corresponding to the retrieval cues receive activation via the second term of Equation 3:

$$A_i = B_i + \sum_{j=1}^{m} W_j S_{ji} \tag{3}$$

In this equation, B_i is the base-level activation as determined by Equation 1. W_j reflects the weights associated with the j elements (slot values) in the goal buffer. The weights W_j default to G/j where G is the total amount of goal activation available. In ACT–R, by default, G is set to 1, and we do not change this value.

Finally, S_{ji} are the strengths of association from elements j to chunk i. *Associative retrieval interference* arises because the strength of association from a cue is reduced as a function of the "fan," which is the number of items associated with the cue (Lewis & Vasishth, 2005, p. 5). Equation 4 calculates the strength of association S_{ji}. Anderson and Reder (1999) adopted a value of 1.5 for S for modeling the fan effect across a range of verbal memory experiments; the present model also takes this value:

$$S_{ji} = S - \ln(\text{fan}_j) \tag{4}$$

Equations 1 through 4 determine the activation of chunks in memory, and Equation 2 maps activation to retrieval time.

2.2.3. Partial matching

As mentioned above, a retrieval request for a chunk having some specific slot values may not lead to successful retrieval. One reason could be that its activation falls beneath the retrieval threshold. Another possibility is that retrieval cues do not perfectly match the slot values of available chunks. However, a partial match between a retrieval specification and a chunk's feature values can result in a successful retrieval of a chunk as long as its activation remains above the threshold.

The notion of partial matching is a core component of the ACT–R architecture, and plays a central role in the model to be presented in this work. In the context of the intrusion effect, this mechanism gives rise to the retrieval of chunks representing structurally inaccessible licensors of polarity elements. The important point to note here is that the partial matching component is independently motivated in the architecture and is based on previous empirical research on human memory processes (Anderson et al., 2004; Anderson & Matessa, 1997). It can be seen as a simple abstraction over the kind of partial matching routinely observed in neural network models of memory, and often included in other mathematical models of memory retrieval.

The extended and final equation for the computation of activation in sentence processing takes partial matching into account and is defined as follows:

$$A_i = B_i + \sum_{j=1}^{m} W_j S_{ji} + \sum_{k=1}^{p} P M_{ki} + \epsilon \tag{5}$$

Partial matching is implemented as a matching summation over the k slot values of the retrieval specification in the condition part of a production. The variable P refers to the match scale—that is, the amount of weighting given to the similarity in slot k; the ACT–R default value 1 is retained in the model. The term M_{ki} refers to the similarity between the value k in the retrieval specification and the value in the corresponding slot of chunk i. This similarity is expressed by maximum similarity and maximum difference. The similarity between anything and itself is set to maximum similarity (the default is 0), and the similarity between any non-identical values is the maximum difference (default is -1). In the present model, we set

Table 2
A comparison of previous and current model parameters

Parameter	Previous Models[a]	Current Model
Decay	0.50	0.50
Maximum associative strength	1.50	1.50
Retrieval threshold	–1.50	–1.50
Maximum difference	–0.60	–0.60
Latency factor	0.14	0.46
Noise	0, 0.15	0.15,0.30, 0.45

[a]Lewis and Vasishth (2005), Vasishth and Lewis (2006).

the maximum difference to −0.6 because this was the value used in the earlier simulations (Lewis & Vasishth, 2005; Vasishth & Lewis, 2006). Nothing hinges on this particular value; default values for maximum difference could have equally been used.

In the simulations discussed in this article, the maximum difference affects the mismatch penalty for DPs that do not perfectly match the retrieval cues; the possible mismatches involve the slot values nominative versus accusative case, and positive versus negative polarity of the matrix and embedded DPs. This becomes clearer when we present the model's actions in relation to the polarity sentences. Other chunk–pair similarities are the same as in Lewis and Vasishth (2005), and are available with the source code of the model.

The model outlined above can explain a range of empirical results in English (Lewis & Vasishth, 2005) and Hindi (Vasishth & Lewis, 2006), some of which pose a challenge for other theories of sentence processing. The details are discussed in these and other articles (for a general overview, see Lewis et al., 2006). The numerical parameters relevant in the sentence processing model are shown in Table 2; there are several other parameters in ACT–R, but these are not relevant for the present discussion and were kept at their default values.

This completes the description of the core ACT–R-based architecture that is relevant to the present model. We discuss next how the intrusion effect is modeled.

3. A model of Drenhaus et al.'s (2005) intrusion effect

As discussed earlier with reference to Example 2, NPIs like *jemals* have the property that they require a c-commanding licensor. In other words, a dependency must be established between the NPI and a licensor. In order to complete this dependency, the NPI initiates a search for an item with two properties: a c-commanding element that is also a negative polarity licensor. This search is driven by an attempt to retrieve an item that has the feature specification "c-commander of NPI" and "NPI licensor." Note that in the constructions considered in Example 2, and repeated below, the licensor c-commands the NPI just in case it is the subject of the sentence (Example 3a); when the licensor occurs inside the relative clause (Example 3c), it does not have the c-commanding property:

1. a. Accessible NPI licensor:
   Kein Pirat, [der einen Braten gegessen hatte,] war **jemals** sparsam
   No pirate who a roast eaten had was ever thrifty
   "No pirate who had eaten roast (meat) was ever thrifty."
   b. Inaccessible NPI licensor:
   Ein Pirat, [der **keinen Braten** gegessen hatte,] war **jemals** sparsam
   A pirate who no roast eaten had was ever thrifty
   "A pirate who had eaten no roast (meat) was ever thrifty."
   c. No NPI licensor:
   Ein Pirat, [der einen Braten gegessen hatte,] war **jemals** sparsam
   A pirate who a roast eaten had was ever thrifty
   "A pirate who had eaten roast (meat) was ever thrifty."

As illustrated in Fig. 2, partial matching plays a crucial role during the resolution of the licensor–NPI dependency. In the grammatical condition (3a), both the retrieval cues at the NPI (c-commanding element and NPI licensor, represented in the figure by the feature + negative) perfectly match the licensor *Kein Pirat,* which is then successfully retrieved. In the intrusion condition (3b), the cue "c-commander" matches the subject DP *ein Pirat,* but the cue "NPI-licensor" (+ negative) matches the embedded DP *keinen Braten.*

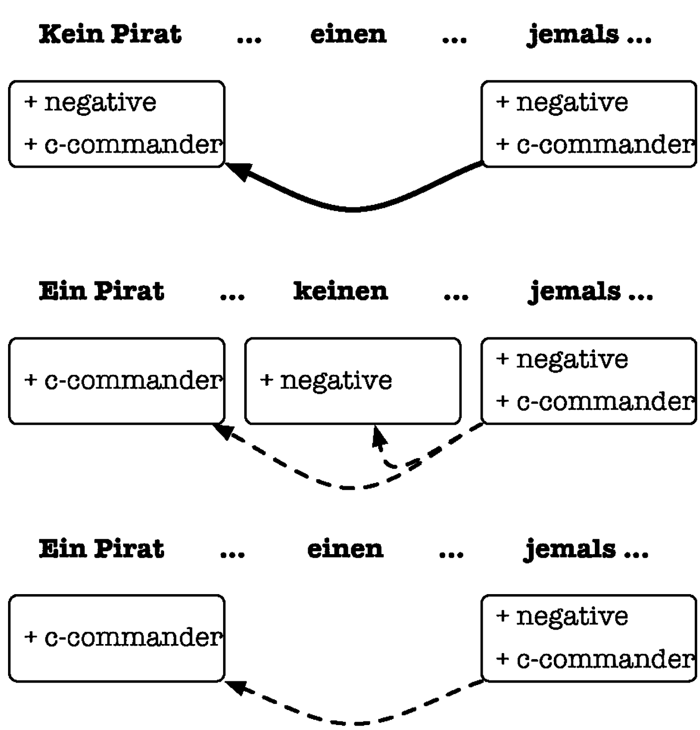

Fig. 2. Schematic illustration of retrieval cues at the negative polarity item. *Note:* The solid-line arrows indicate situations where both retrieval cues match with a target's feature specification, and dashed lines indicate partial-cue matches.

An important implementation issue relates to the manner in which the retrieval cue ”c-commander” is specified. Our ACT–R implementation simply looks for the matrix subject DP, which in the present stimuli is distinguished from the embedded DP by having nominative case marking. In the experiment items, there is an isomorphism in the present example sentences between the case marking of the DPs and their c-commanding status with respect to the polarity item. A full implementation of the c-command relation would have to mark the relationships between all non-terminal nodes and the polarity item. We did not build such a full implementation because of the isomorphic relationship between case marking and c-command (relative to the polarity item). Clearly, a general theory of c-command as a retrieval cue would require considerably more detail in the model.

The partial matching term in the activation (Equation 5) penalizes the activations of the target DPs, reducing their activation; the DP with a higher final activation gets retrieved, but the probability of the embedded DP being retrieved is higher. Specifically, whenever a mismatch occurs between the retrieval cues at the polarity items and the DPs’ corresponding slot values, the maximum difference penalty (–.60) reduces the activation of the mismatching chunk, as discussed earlier with reference to Equation 5.

Finally, in the no licensor condition (3c), only one retrieval cue (“c-commander”) matches the subject DP, resulting in a partial matching penalty and, therefore, slower retrieval; but, in contrast to the intrusive condition, the probability of the embedded DP being retrieved is low because it does not match either retrieval cue.

In sum, the fastest retrieval will occur in the grammatical condition because both the retrieval cues succeed in finding the correct (main) DP for retrieval. In the intrusion condition, the matrix DP matches the c-command cue, but the embedded DP matches the NPI-licensor cue; in any given run of the model, both DPs will get a mismatch penalty resulting in lower activation, and whichever has higher activation will be retrieved. This results in greater proportions of retrieval errors and longer retrieval time compared to the grammatical condition. The no licensor condition (3c) will also involve relatively slow retrieval due to partial matching.

Partial-cue matching is thus a major component of the explanation for the intrusion effect: The embedded DP occasionally ends up incorrectly licensing the NPI, giving an illusion of grammaticality. We will see below that partial matching is responsible only for making it possible to retrieve an element that matches a partial description; the predictions of the model fall out of an interaction with other components of the theory such as interference, decay, and stochastic noise. This interaction is non-obvious and can only be explored by simulation and parametric variation. This is discussed in more detail in conjunction with the eye-tracking experiment further on.

3.1. Modeling results

As discussed earlier, the first goal of the modeling task was to explain the pattern of correct-response proportions that Drenhaus et al. (2005) found. Here, it is necessary to first lay out our assumptions regarding the connection between speeded grammaticality judgments and online processing complexity. Making a grammaticality judgment is not an activity that humans normally engage in while comprehending a sentence outside of experimental settings. The source of the judgment itself is presumably a decision process that takes as input the products of (possibly partially) completed online processing. Some relation is assumed to exist between the cost of online processing and the proportion of correct judgments that follow from the decision process (Fanselow & Frisch, 2006). If this assumption is correct, then it is reasonable to assume that the product of a correct or incorrect retrieval during parsing will affect the grammaticality judgment, especially under time pressure. The grammaticality judgment is probably also affected by other factors that are related to the decision process *per se* and not to the processing cost, but the judgment at least bears some relation to the product of the retrieval. By contrast, the *latency* of the judgment may or may not bear any relation to retrieval latency—its source could be any of the factors that come into play in the decision-making process. For example, the speed of the judgment could depend on the ease with which the products of online processing are accessed, which (although an interesting question per se) is orthogonal to the main issue of interest: the reflection of processing difficulty in the grammaticality judgment.

Given the above discussion, we model only the proportion of correct responses in each case, not the latency of these responses. We consider a grammaticality judgment as being correct when the matrix DP is successfully retrieved. This assumption derives from two facts about the stimulus sentences. The first is that the embedded DP is incompatible with the adjective. The second is that the NPI *jemals* requires that an NPI licensor c-command it; in the present structures, the only c-commanding DP is the matrix one. In the accessible licensor condition, a retrieval of the matrix DP (*Keinen Pirat* "no pirate") corresponds to a judgment that the sentence is grammatical. In the inaccessible licensor condition, a retrieval of the matrix DP (*Ein Pirat* "a pirate") corresponds to a correct judgment that the sentence is ungrammatical. By contrast, in the inaccessible licensor condition, a retrieval of the embedded DP (*einen Braten* "a roast") results in an incorrect judgment that the sentence is grammatical. In the no licensor condition, the retrieval of either the matrix or embedded DP would result in a correct judgment that the sentence is ungrammatical; however, retrieving the embedded DP should signal an ungrammaticality due to its incompatibility with the adjective.

The model thus yields the *proportion* of correct retrievals (of the matrix DP) over many trials, which, under our assumptions, is related to the process of grammaticality judgment decisions. The results of 500 runs of the model are presented in Table 3; they show a pattern of matrix–DP retrieval consistent with the judgment data. In the grammatical (accessible licensor) condition the model performs extremely well and retrieves the correct matrix DP 88.5% of the time; this is somewhat better than the participants' performance. In the inaccessible licensor condition, it retrieves the matrix DP only in about 58.5% of cases, and in the no licensor condition the matrix DP is retrieved 76.6% of the time. Although the percentages of correct

Table 3
Percentage of correct judgments and of correct retrievals in the model

Condition	Data	Model
(2a) Accessible licensor	85	88.5
(2b) Inaccessible licensor	70	58.5
(2c) No licensor	83	76.6

retrievals do not match the data perfectly, the pattern is qualitatively similar to the behavioral data. It may be possible to find the right combination of parameter values in the model to approximate the percentages in the data; but this was not the goal of the modeling exercise. The goal is rather to build a predictive (rather than *post hoc* model) in the sense of Anderson et al. (2004, p. 1046); in other words, the goal is to determine whether the data can be fit using parameters that have been previously fixed. Because none of the parameters were varied, except the latency factor (which defines the mapping between activation and latency) and activation noise (which, at 0.45, is close to the default value of 0.40 in ACT–R; Anderson et al., 2004, p. 1048), the exercise can be considered at least a partial success.

It is worth repeating here that we do not model the *latency* of making grammaticality judgments because doing so would require building a theory of the underlying the decision processes that result in grammaticality judgments. Although such a theory would be of inherent value, it lies beyond the goals of this article.

A possible criticism of the empirical basis of the intrusion effect is that the data come from only two experiments involving the speeded grammaticality judgment task (Drenhaus et al., 2005, carried out the experiment twice, once combined with an ERP experiment). However, the intrusion effect has been replicated and extended for English by another laboratory using the rapid serial visual presentation task (Xiang, Dillon, & Phillips, 2006). Xiang et al. replicated the effect using the licensor *no,* which is a stronger NPI licensor (van der Wouden, 1997), and frequently co-occurs with the NPI *ever* (12.5%) and with other weaker and less frequently occurring licensors: *few* and *only* (2.4% and 7.2%, respectively).

Although judging sentences under time pressure may be indirectly and partly related to online parsing processes, it is important to establish whether the intrusion effect can be found in a different task that involves automatic processing rather than the task of providing a grammaticality judgment under time pressure. Eye tracking during reading is an ideal method for addressing this question because it yields highly articulated measures of moment-by-moment comprehension difficulty (for a comprehensive review, see Rayner, 1998).

Therefore, we conducted an eye-tracking study of the intrusion effect; but, in addition to the intrusion effect with NPIs, we also considered the effect of an intrusive *negative* polarity licensor on *positive* polarity items (PPIs). We discuss the details of this experiment next.

4. An eye-tracking investigation of the intrusion effect

4.1. Method

4.1.1. Participants

Forty-eight native German speakers (undergraduates at the University of Potsdam) took part in this study, each receiving 7€ for participating. Participants were tested in individual sessions, and took approximately 30 min to complete the experiment.

4.1.2. Stimuli and procedure

Both filler and target materials were presented as whole texts on a single line.

Participants were seated 55 cm from a 17-in. color monitor with 1,024 × 768 pixel resolution. Participants were asked to sit comfortably in front of an IView-X eye tracker (SensoMotoric Instruments) running at 240 Hz sampling rate, 0.025° tracking resolution, and <0.5° gaze position accuracy. They were asked to place their head in a frame and to position their chin on a chinrest for stability. Viewing was binocular, but only the participant's right eye was tracked. The angle per character was 0.26° (3.84 characters per degree of visual angle).

Participants were asked to avoid large head movements throughout the experiment. A standard three-button mouse was used to record button responses. The presentation of the materials and the recording of responses was controlled by two PCs running proprietary software (the software used was Presentation and SensoMotoric Instruments' own software for eye-tracker control).

Each participant was randomly assigned one of six different stimulus files that comprised different item–condition combinations according to a Latin Square. There were 86 filler sentences and 36 stimulus sentences in each list, and each list was pseudo-randomly reordered. The trials per session were randomized once for each file, subject to the constraint that each session started with at least three fillers.

At the start of the experiment, the experimenter performed a standard calibration procedure, which involves participants looking at a grid of 13 fixation targets in random succession in order to validate their gazes. Calibration and validation were repeated after every 10 to 15 trials throughout the experiment, or if the experimenter noticed that measurement accuracy was poor (e.g., after large head movements or a change in the participant's posture).

Each trial was structured as follows: First, a fixation target in the same position as the first character of the text display was presented; a fixation on this target triggered the presentation of the sentence (this ensured that participants always started reading in the left-most character position). Participants were instructed to read the sentence at a normal pace and to move their gaze to a dot at the bottom right of the screen after finishing the sentence. This triggered the presentation of a simple comprehension question, which the participant answered by clicking one of two boxes on the screen. Responding to the question triggered the presentation of the next trial. The comprehension questions were included in order to ensure that the sentences were read for comprehension.

As discussed below, three of the six conditions in the experiment consisted of ungrammatical sentences, which implies that participants had to occasionally answer questions about ungrammatical structures. For this reason, we do not attempt to interpret the response accuracies, although we report them in the results below for completeness.

The six conditions in the experiment are illustrated below. The reader may wonder why we do not have a condition with an NPI licensor in both the main and embedded clauses (*Kein Pirat, den keinen Braten gegessen hatte, . . .*). The reason is that in the previous experiments by Drenhaus et al. (2005), participants were unable to process sentences with two NPI licensors, rendering the results difficult to interpret:

1. a. Accessible NPI licensor NPI:
Kein Pirat, [der einen Braten gegessen hatte,] war **jemals** sparsam
No pirate who a roast eaten had was ever thrifty
"No pirate who had eaten roast (meat) was ever thrifty."

b. Inaccessible NPI licensor, NPI:
Ein Pirat, [der **keinen Braten** gegessen hatte,] war **jemals** sparsam
A pirate who no roast eaten had was ever thrifty
"A pirate who had eaten no roast (meat) was ever thrifty."
c. No NPI licensor, NPI:
Ein Pirat, [der einen Braten gegessen hatte,] war **jemals** sparsam
A pirate who a roast eaten had was ever thrifty
"A pirate who had eaten roast (meat) was ever thrifty."
d. Accessible NPI licensor, PPI:
Kein Pirat, [der einen Braten gegessen hatte,] war **durchaus** sparsam
No pirate who a roast eaten had was certainly thrifty
"No pirate who had eaten roast (meat) was certainly thrifty."
e. Inaccessible NPI licensor, PPI:
Ein Pirat, [der **keinen Braten** gegessen hatte,] war **durchaus** sparsam
A pirate who no roast eaten had was certainly thrifty
"A pirate who had eaten no roast (meat) was certainly thrifty."
f. No NPI licensor, PPI:
Ein Pirat, [der einen Braten gegessen hatte,] war **durchaus** sparsam
A pirate who a roast eaten had was certainly thrifty
"A pirate who had eaten roast (meat) was certainly thrifty."

The first three NPI conditions need no further explanation. The PPI conditions were included in order to explore the model's behavior with a different kind of polarity item. PPIs have the property that they *cannot* occur in the scope of a negative element. Thus, in Example 4d, the PPI *durchaus* "certainly" is not licensed because of the presence of a c-commanding negative polarity licensor; in Examples 4e and 4f, it is licensed because the c-commanding element is not a negative polarity licensor. The cue matches and mismatches for PPIs are illustrated in Fig. 3. In the ungrammatical condition (4d), the retrieval cue "c-commander" matches the subject DP, whereas the cue "PPI licensor" (+ positive) matches the embedded DP; this results in a partial match penalty on both DPs (i.e., a lowered activation of the DPs), and the DP with the higher activation is retrieved. The probability of a mis-retrieval here is higher than in the other PPI conditions; in the embedded NPI-licensor condition (4e), there is a perfect match with the subject DP, resulting in fast and accurate retrievals; and in the no NPI-licensor condition (4f), there is also a perfect match, although the embedded DP also has a partial match and, therefore, a reduced activation. This is a description of the qualitative behavior of the model; only by running the model can we determine its quantitative predictions.

The question of interest was, "Can the model explain any of the patterns in the dependent measures at the two types of polarity items?" We turn next to the predictions of the model for these six conditions.

4.1.3. Reading time predictions of the model and their mapping to dependent measures

In the model, after lexical access succeeds and syntactic integration processes are completed, the NPI triggers an attempt to retrieve a licensor that c-commands it, and the PPI similarly

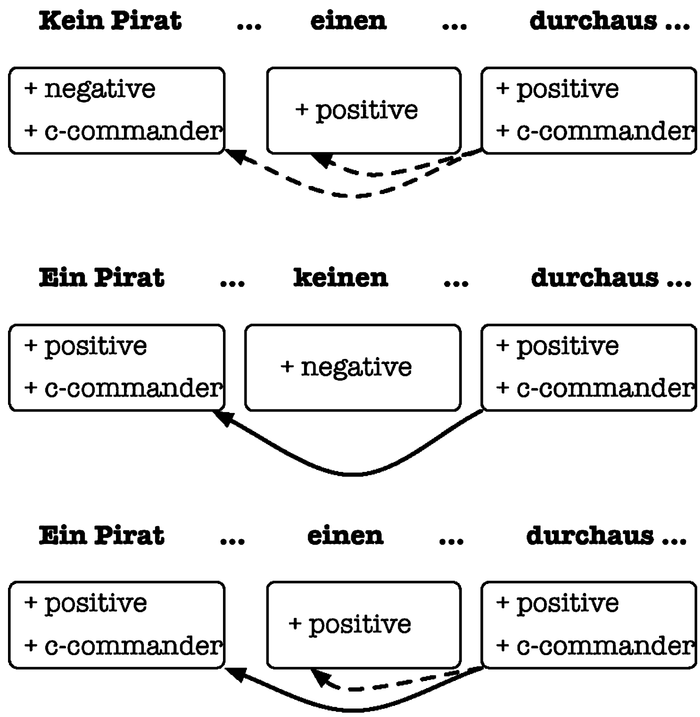

Fig. 3. Schematic illustration of retrieval cues at the positive polarity item and the relevant slot values at the determiner phrases. *Note:* The solid-line arrows indicate situations where both retrieval cues match with a target's feature specification, and dashed lines indicate partial-cue matches.

attempts to retrieve a licensor that does not have the NPI-licensing property. As discussed earlier (Figs. 2 and 3), this retrieval process is a content-addressable search for a previously processed element with certain properties.

Table 4 shows the model's quantitative reading time predictions for the six conditions; four activation noise values are used in order to illustrate the impact of noise on the dynamics of retrieval latency. Without any activation noise and with partial matching switched off, the model simply fails to process the ungrammatical conditions. If partial matching is on but noise switched off, the model can retrieve a DP, but its behavior is deterministic: In the grammatical conditions, the DP is retrieved quickly (375 msec); and in the ungrammatical conditions, retrieval is slow (601 msec). Once noise is switched on, the model display an interesting interaction with partial matching (and other numerical variables such as decay and interference), and results in the non-determinism that yields a gradient response. The table also shows values when the latency factor is left unchanged from previous simulations at the value 0.14 (Lewis & Vasishth, 2005; Vasishth & Lewis, 2006).

In order to map the model's predictions to eye-tracking dependent measures, it is necessary to arrive at an understanding of the mapping between eye-tracking dependent measures and human parsing processes. The most common dependent measures and their interpretation in terms of reading processes are as follows: *First fixation duration* (FFD) is the first fixation during the first pass, and has been argued to reflect lexical access costs (Inhoff, 1984). *Gaze* *duration* or *first pass reading time* (FPRT) is the summed duration of all the contiguous fixations in a region before it is exited to a preceding or subsequent word; Inhoff suggested that FPRT reflects text integration processes, although Rayner and Pollatsek (1987) argued that FFD and FPRT may reflect similar processes and could depend on the speed of the cognitive process. *Right-bounded reading time* (RBRT) is the summed duration of all the fixations that fall within a region of interest before it is exited to a word downstream; it includes fixations occurring after regressive eye movements from the region, but does not include any regressive fixations on regions outside the region of interest. RBRT may reflect a mix of late and early processes because it subsumes FFDs. *Re-reading time* (RRT) is the sum of all fixations at a word that occurred after first pass; RRT has been assumed to reflect the costs of late processes (Gordon et al., 2006, p. 1308). Another measure that may be related to late processing is *regression path duration* (RPD), which is the sum of all fixations from the first fixation on the region of interest up to, but excluding, the first fixation downstream from the region of interest. Finally, *total reading time* (TRT) is the sum of all fixations on a word.

Which of these measures should map onto the retrieval times generated by the model? As we discuss in the conclusion, a detailed answer to this question demands a highly articulated model of the link between eye-movement control and the kind of higher level linguistic processes examined here. Such models currently do not exist, and developing one is beyond the scope of the present article. Nevertheless, in advance of such developments, we can bring to bear a number of empirical and theoretical considerations to narrow the set of plausible measures to align with predictions of the present model.

Arguably, the first major distinction is between early and late measures (Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). Although the reported effects in the empirical literature are somewhat mixed (Clifton, Staub, & Rayner, 2007), post-lexical effects such as similarity-based interference are more reliably reflected in measures such as RRT than in measures such as FFD or FPRT (Gordon et al., 2006).

Table 4
The model's predicted retrieval latencies for the six conditions (500 runs)

Parameters			Conditions					
F	PM	Noise	a	b	c	d	e	f
0.46	Off	0.00	375	Fail	Fail	Fail	375	375
	On	0.00	375	601	601	601	375	375
	On	0.15	420	644	688	636	424	421
	On	0.30	463	620	699	630	469	430
	On	**0.45**	**482**	**602**	**679**	**580**	**491**	**441**
0.14	On	0.15	181	235	250	234	182	181
	On	0.30	185	232	250	228	152	182
	On	0.45	192	224	260	228	197	187

Note. The latency factor (F) is 0.46 (the value used in the simulations shown in Fig. 7); or 0.14 (the value used in earlier simulation, Lewis & Vasishth, 2005), partial matching (PM) switched off or on, and (when partial matching is on) with different noise levels. In this article, in addition to latency factor 0.46, we use a noise level of 0.45. The relevant row of parameter settings and retrieval latencies is shown in bold.

The DP retrievals in the present model are processes that occur after lexical retrieval is complete; they must follow the initial lexical access and integration with the verb—the DP retrieval is contingent upon information generated by these processes. Therefore, it makes sense that the retrieval durations would affect dependent measures that tend to reflect post-lexical processes. We can therefore narrow the set of candidate measures to RPD, RBRT, and RRT, and restrict our subsequent analyses to these three.

First consider RPD. Clifton et al. (2007) suggested that RPD may reflect, among other things, the overcoming of processing difficulty at a word—which is comparable to the retrieval latencies at the polarity item. As they put it, "The occurrence of a regression reflects difficulty in integrating a word when it is fixated, arguably an early effect. The [RPD] measure reflects this effect, *but also reflects the cost of overcoming this difficulty, which may well occur late in processing*" (p. 349, italics added). Thus, although RPD is a mix of reading times at the critical word and any number of words preceding the critical word, it may also include a component that reflects retrieval difficulty.

Apart from RPD, RBRT and RRT may also be good candidate measures because they are restricted to reading times at the critical word. Of these two, RRT provides the purest measure of late processing; RBRT includes both early and late measures, as discussed earlier.

On the basis of this analysis, we therefore suggest the following plausible mapping: The model's retrieval time predictions should align most closely with RRT, followed by RBRT, and possibly RPD—the last measure less closely, given the additional inherent variability introduced by reading times from other regions. We now show that the model's predictions match the empirical results reasonably well under this mapping.

4.2. Results

4.2.1. Dependent measures

Five orthogonal contrasts were carried out in the polarity-item region: (1) effect of polarity type: NPI versus PPI, (2) grammaticality effect on NPI: a versus b and c, (3) intrusion effect on NPI: b versus c, (4) grammaticality effect on PPI: d versus e and f, (5) intrusion effect on PPI: e versus f. All reading times less than 50 msec were removed and treated as missing values. The alpha value was set at 0.05; the Bonferroni correction was not necessary because we based our inferences on Bayesian (highest posterior density) confidence intervals for the multilevel linear model's coefficient estimates; these are more conservative than standard least squares estimates (Gelman & Hill, 2007; Gelman & Tuerlinckx, 2000). Analyses were carried out using raw as well as log-transformed values; the latter are more appropriate when additivity and linearity are not reasonable assumptions (Gelman & Hill, 2007, p. 59). The results were comparable, except in one case: The NPI intrusion effect in RBRT was no longer statistically significant (the sign of the estimated coefficient did not change). Table 5 summarizes the results for the comparisons using raw reading times because this is the convention in psycholinguistics. Data and R code accompanying this article (http://www.ling.uni-potsdom.de/~vasishth/VBLDCCogSci08/) allow the reader to generate all results themselves.

NPIs were read slower than PPIs in RBRT, RPD, and RRT; in these measures, the grammatical NPI condition (a) was also read faster than ungrammatical conditions (b and c). The intrusive NPI condition (b) was faster than Condition c in RBRT and RPD, but was not significant

Table 5
Results of multilevel data analysis

Measure	Comparison	Estimate	SE	Lower	Upper	t
RRT						
	c1	78.83	15.13	49.23	108.62	5.21
	c2	62.57	14.50	34.33	91.36	4.32
	c3	−11.49	24.17	−61.19	34.06	< 1.00
	c4	51.00	15.61	20.03	81.31	3.27
	c5	30.40	28.16	−25.69	85.28	< 1.00
RBRT						
	c1	38.76	5.60	27.55	49.64	6.93
	c2	22.94	5.53	12.21	33.99	4.15
	c3	−21.66	9.55	−40.60	−3.20	−2.27
	c4	1.33	5.69	−9.84	12.45	< 1.00
	c5	4.88	9.78	−13.53	24.58	< 1.00
RPD						
	c1	127.37	23.78	80.87	174.08	5.35
	c2	80.66	23.54	34.36	125.81	3.43
	c3	−89.02	40.59	−167.04	−8.11	−2.19
	c4	−14.44	24.19	−62.56	32.02	< 1.00
	c5	5.14	41.61	−77.91	86.20	< 1.00

Note. The five orthogonal contrasts for re-reading time (RRT), right-bounded reading time (RBRT), and regression path duration (RPD). The contrast cl is the effect of polarity type, c2 the effect of negative polarity item (NPI) grammaticality, c3 the intrusion effect in NPIs, c4 the grammaticality effect in positive polarity items (PPIs), and c5 the intrusion effect in PPIs. T scores with absolute values greater than 2 are statistically significant.

in RRT. The grammatical PPI conditions (e and f) were read significantly faster than the ungrammatical one (d) in RRT; none of the other measures showed a significant difference. The intrusive condition (e) was not significantly different from the grammatical one (f).

4.3. Discussion and comparison with the model's predictions

4.3.1. NPIs

Increased processing difficulty is experienced when the NPI is not licensed, and RPDs show an ascending-steps pattern (Conditions a–c): The grammatical condition is fastest, the intrusive licensor condition is intermediate, and the no licensor condition is slowest. In RRTs and TRTs, the difference between the two ungrammatical conditions (intrusive and no licensor) disappears. These results suggest that, compared to the grammatical condition, increased processing difficulty occurs in the intrusive and no licensor conditions, and this difficulty is possibly greater in the no licensor condition than the intrusive licensor condition. Our explanation for this difference is that the parsing mechanism sometimes mis-retrieves the intrusive (illicit) licensor due to partial matching, and the reader therefore assumes that the sentence is grammatical. (Note that the speeded judgment study of Drenhaus et al., 2005, also yielded judgment latencies, and these do not match the pattern found in the eye-tracking data. However, as discussed earlier, we do not assume that the time required for making a grammaticality judgment is related to the online processing cost at the polarity item as expressed in eye-tracking measures.)

Here, one may question the evidence for the intrusion effect in the NPI conditions; after all, the intrusion-effect contrast (c3) is significant in RBRT and RPD, but not in RRT. Notice, however, that the coefficient estimates are negative for this contrast in all three dependent measures. This stability of the coefficient estimate across the three measures is a better decision criterion than p values (Gelman & Hill, 2007, pp. 73–74).

4.3.2. PPIs

The PPIs (Conditions d–f) show a tendency toward a descending-step pattern in RRTs. This pattern suggests that the greatest difficulty occurs in the ungrammatical sentence and the least in the grammatical sentences. In the ungrammatical condition (d), slower processing would occur due to partial matches with the matrix DP; whereas in the intrusive NPI condition (e) and the no NPI-licensor condition (f), there is a perfect and, therefore, fast match to the matrix DP (see Fig. 3).

4.4. Comparing the model's predictions with the dependent measures

The next question of interest is, "How well do the reading times match the model discussed earlier?" The Drenhaus et al. (2005) experiment yielded percentages of judgment accuracy, which the model is able to fit adequately (Table 3).

As mentioned earlier, we did not model the latencies of grammaticality judgments because they may reflect the time course of processes underlying the meta-linguistic task of providing judgments, and are not necessarily a measure of difficulty experienced during *automatic* processing—after all, humans do not read sentences in order to judge them grammatical or not, but rather to comprehend the content. It follows that we do not expect any correspondence between the latencies in the speeded judgment task and the eye-tracking dependent measures. Modeling the eye-tracking dependent measures is a greater challenge because our goal was not to merely fit the data but to explore the predictions of the model while holding the numerical parameters at previously fixed values.

Fig. 4 shows the results of the comparisons between dependent measures and the retrieval latencies from the model. As discussed earlier, with reference to Table 4, the only parameter that is different from earlier simulations (Lewis & Vasishth, 2005) is the scaling factor F, which was set at 0.46. The previously used value 0.14 shows identical patterns, except that retrieval latencies are obviously faster. Note that the retrieval latencies from the model only reflect the difficulty at the polarity item of retrieving and integrating a targeted licensor. Thus, the model's predictions provide a necessarily incomplete picture of the factors that determine the reading times.

Overall, the only pattern that fits well with retrieval latencies are RRTs, adjusted $R^2 = 0.88$. The fit with RBRT was $R^2 = 0.62$ and with RPD was $R^2 = 0.52$. The divergence between model and data in these last two measures could be due to the fact that in the grammatical conditions (e and f) of the PPIs, the retrieval target (the main clause DP) matches perfectly with the retrieval cues (see Fig. 3). This is not the case in the two ungrammatical conditions

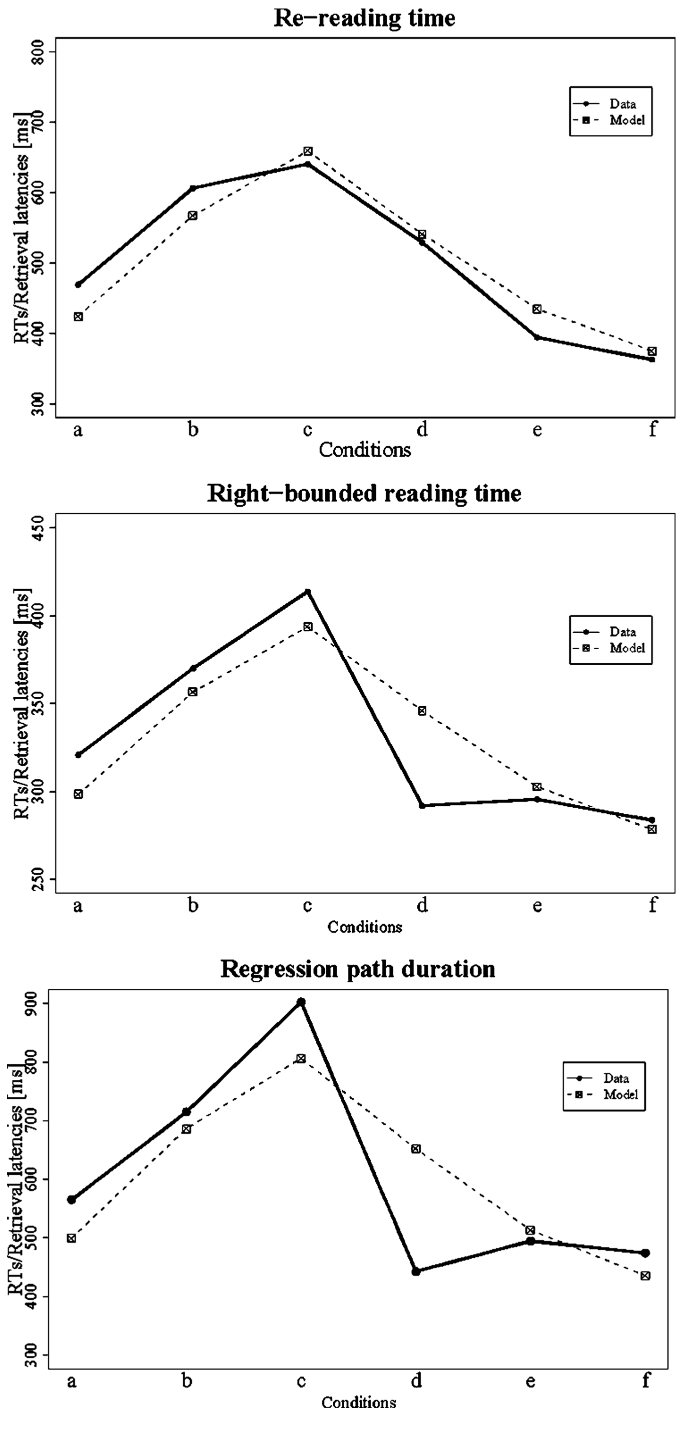

Fig. 4. A comparison of the model's predictions for dependent measures at the negative and positive polarity items. *Note:* In these fits, noise is 0.45, and the scaling parameter is $F = 0.46$. See Table 2 for other parameter values, and see Table 4 for a summary of the effects of varying noise and the scaling parameter. RT = reading time.

(a and b) for the NPIs; there, a partial match occurs in each case. It is possible that these differences have an impact on the retrieval patterns (Fig. 2) in a manner not captured by the model.

The model–data comparison thus suggests that RRT may reflect difficulties associated with the cue-based integration process. Indeed, eye-tracking research by Gordon et al. (2006) has also found evidence for similarity-based interference effects in RRT, a result that is consistent with our linking hypothesis here.

5. Concluding remarks

We have argued that dependency resolution in sentence processing is driven by cue-based retrieval processes (for a related proposal, see Van Dyke & Lewis, 2003), and that retrieval latency is subject to several general constraints on activation. We demonstrated this by modeling an otherwise difficult-to-explain set of results involving polarity licensing. The intrusion effect, we argue, can be explained in terms of constraints defined in an existing ACT–R, coupled with a sentence processing model implemented within this architecture. A notable result is that the model's retrieval latencies are fitted to the data without any adjustment of the key numerical parameters in the model. To the extent that the model can account for the observed reading times at the polarity items, the present results provide new support for the model.

Of course, there is much that the model currently does not achieve. First, it makes no predictions about N400 and P600 effects found at the polarity item. Second, the model does not include a general theory of polarity licensing, and so it is has nothing to say about the rich array of constraints that affect polarity items. Third, although the model addresses eye-tracking data, it does not include a specification of the interaction of linguistic versus eye-movement control.

Regarding the first issue (absence of an explanation for the N400 and P600 components), the relationship between cue-based retrieval mismatches and the N400 and P600 components can be qualitatively (and very speculatively) examined. In both the intrusive licensor condition (b) and the no licensor condition (c) for NPIs, the increased processing difficulty due to cue mismatches could express itself in the ERP components. In principle, it is possible to transform this hypothesis into an ACT–R-based model that delivers predictions of ERP effects, and we intend to address this in future work.

Second, regarding the issue that the model has no general theory of polarity processing, we stress that this was not a goal of the modeling task. The goal was rather to explain a surprising empirical result using an existing computational model of sentence processing, and to extend the result with a different experimental paradigm (eye tracking). The remarkable result in this article is that the model can fit RRT patterns in NPIs and PPIs without modifying the parameters for decay, interference, and partial matching. To our knowledge, there exists no other model of sentence processing (implemented computationally or verbally stated) that could, without making additional, post-hoc assumptions, explain the subtle polarity licensing facts presented in this and earlier work. In addition, although there are several theories of polarity licensing in linguistics, currently there exists no competence theory-based explanation that would predict the judgment patterns and reading time patterns.

A third shortcoming of the model is that a fuller specification of sentence processing that depends on eye-tracking data should ideally be tightly coupled to a computational model of eye-movement control. However, in principle, this shortcoming does not prevent us from pursuing the central question we address in this and other articles: How are dependencies established? We have argued that this process is mediated by cue-based retrieval, which, critically, is subject to several independently motivated constraints on human working memory (as opposed to arbitrarily defined ones). We have tried to show that the interactions between these constraints result in a surprising pattern of retrievals and latencies that are also observed in the behavioral data.

Notes

1. The term polarity *item* may strike members of the non-linguistic audience as misleading or confusing; a better term might be polarity *element*. However, in this article, we follow the linguistic convention of referring to such elements as polarity items
2. For the purpose of this article, we restrict ourselves only to the negative quantifier as the licensing environment for *jemals*. However, this characterization of the licensing contexts for negative polarity items (NPIs) is incomplete. NPIs can be licensed in other contexts such as yes–no questions ("Did you see anyone?"), wh- questions ("Who saw anyone at all?"), antecedents of conditionals ("If you see anyone, let me know"), S-conditionals ("She ran faster than anyone expected."), the restrictor of the universal quantifier ("Every student who had read anything about Einstein passed the exam."), before- clauses ("Before John had a chance to talk to any student, the class started."), and quantifiers like *few* ("Very few professors read anything."; cf. Giannakidou, 1998, 2001). In addition, there are contexts in which a polarity item is licensed even if it is not overtly c-commanded by negation ("A doctor who knew anything about acupuncture was not available.").
3. Tabor, Galantucci, and Richardson (2004) demonstrated other kinds of intrusion effects in sentence processing, where a part of a sentence is incoherent in the global syntactic context but locally grammatical and coherent; their research showed that the ungrammatical substring intrudes on the processing of the sentence. The phenomenon we address is also an intrusion effect, but it should not be confused with Tabor et al.'s use of the term.

Acknowledgments

This article has greatly benefited from comments by the associate editor, John Anderson; and the reviewers, Bruno Emond, Ted Gibson, and Erik Reichle; as well as from discussions with several colleagues—principally, Gisbert Fanselow, Sam Featherston, Reinhold Kliegl, Pavel Logačev, Don Mitchell, Stefan Müller, Colin Phillips, Douglas Saddy, Esther Sommerfeld, Julie Van Dyke, and Sashank Varma. Shravan Vasishth is particularly grateful to Reinhold Kliegl for many extended discussions and advice regarding multilevel modeling; meetings with Andrew Gelman and Douglas Bates were also extremely helpful. Sabine Kern, Pavel Logačev, Kathleen Raphael, and Kai Sippel provided considerable technical assistance. All statistical analyses were carried out with the *R* system for statistical computing and graphics (R Development Core Team, 2006), Version 2.4.1 (2006–12–18), which is available under the GNU General Public License Version 2. Within R, the following packages were used: *lme4* (Bates & Sarkar, 2007), *coda* (Plummer, Best, Cowles, & Vines, 2006), *lattice* (Sarkar, 2006), and *Hmisc* (Harrell, 2006). In addition, the dependent measures for eye-tracking data were computed with the R package *em* (Logačev & Vasishth, 2006), available from Shravan Vasishth. The source code of the ACT–R model, the eye-tracking data analyzed in this article, and the R code used for analysis are provided as accompanying material with this article (see http://www.cogsci.rpi.edu/CSJarchive/Supplemental/index.html).

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111,* 1036–1060.

Anderson, J. R., Kline, P., & Lewis, C. (1977). A production system model for language processing. In M. Just & P. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 271–312). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Anderson, J. R., & Lebiere, C. (Eds.). (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Anderson, J. R., & Matessa, M. (1997). A production system theory of serial memory. *Psychological Review, 104,* 728–748.

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General, 128,* 186–197.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2,* 396–408.

Baker, C. (1970). Double negatives. *Linguistic Inquiry, 1,* 169–186.

Bates, D., & Sarkar, D. (2007). *lme4: Linear mixed-effects models usings 4 classes* [R package Version 0.9975–11].

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the "logicality" of language. *Linguistic Inquiry, 37,* 535–590.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Clifton, C., Jr., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. V. Gompel, M. Fisher, W. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 341–372). New York: Elsevier.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–114.

Drenhaus, H., beim Graben, P., Saddy, D., & Frisch, S. (2006). Diagnosis and repair of negative polarity constructions in the light of symbolic resonance analysis. *Brain and Language, 96,* 255–268.

Drenhaus, H., Saddy, D., & Frisch, S. (2005). Processing negative polarity items: When negation comes through the backdoor. In S. Kepser & M. Reis (Eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives* (pp. 145–165). Berlin: de Gruyter.

Fanselow, G., & Frisch, S. (2006). Effects of processing difficulty on judgments of acceptability. In G. Fanselow, C. Fery, M. Schlesewsky, & R. Vogel (Eds.), *Gradience in grammar: Generative perspectives* (pp. 291–316). New York: Oxford University Press.

Fauconnier, G. (1975a). Polarity and the scale principle. *Chicago Linguistic Society, 11,* 188–199.

Fauconnier, G. (1975b). Pragmatic scales and logical structure. *Linguistic Inquiry, 6,* 353–375.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Friederici, A. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain & Language, 50,* 259–281.

Friederici, A. (2002). Towards a neural basis of auditory language processing. *Trends in Cognitive Science, 6,* 78–84.

Friederici, A., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research, 31,* 45–63.

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition, 85,* 83–92.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

Gelman, A., & Tuerlinckx, A. F. (2000). Type s error rates for classical and Bayesian single and multiple comparison procedures. *Computational Statistics, 15,* 373–390.

Giannakidou, A. (1998). *Polarity sensitivity as (non) veridical dependency*. Amsterdam: Benjamins.

Giannakidou, A. (2001). The meaning of free choice. *Linguistics and Philosophy, 24,* 659–735.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68,* 1–76.

Gibson, E. (2000). Dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain: Papers from the first mind articulation project symposium*. Cambridge, MA: MIT Press.

Gibson, E., & Thomas, J. (1999). Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes, 14,* 225–248.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27,* 1411–1423.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language, 51,* 97–104.

Gordon, P. C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory and Coginition, 32,* 1304–1321.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science, 13,* 425–430.

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input. *Cognitive Science, 29,* 261–290.

Harrell, F. E., Jr. (2006). *Hmisc: Harrell miscellaneous* [R package Version 3.1–2].

Hawkins, J. A. (1994). *A performance theory of order and constituency*. New York: Cambridge University Press.

Horn, L. (2001). *A natural history of negation*. Stanford, CT: Center for Study of Language and Information.

Inhoff, A. W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior, 23,* 612–624.

Israel, M. (2006). The pragmatics of polarity. In L. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 701–723). Oxford, England: Blackwell.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Just, M., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329–354.

Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122–149.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15,* 159–201.

Krifka, M. (1995). The semantics and pragmatics of polarity items. *Linguistic Analysis, 25,* 209–257.

Kutas, M., & Petten, C. K. V. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). San Diego: Academic.

Ladusaw, W. (1980). *Polarity sensitivity as inherent scope relations*. New York: Garland.

Lewis, R. L. (1996). Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research, 25,* 93–115.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29,* 1–45.

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

Linebarger, M. (1987). Negative polarity and grammatical representation. *Linguistics and Philosophy, 10,* 325–387.

Logačev, P., & Vasishth, S. (2006). *The em package for computing eyetracking measures.* Potsdam, Germany.

McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 46, pp. 155–200). San Diego: Academic.

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. (1973). Production systems: Models of control structures. In W. G. Chase (Ed.), *Visual information processing* (pp. 463–526). New York: Academic.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28,* 411–421.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). *Coda: Output analysis and diagnostics for mcmc* [R package Version 0.10–7].

Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422.

Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: A tutorial review. *Attention and Performance XII: The Psychology of Reading,* 327–362.

Rayner, K., Sereno, S., Morris, R., Schmauder, A., & Clifton, J., C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4*(3), 21–49.

R Development Core Team. (2006). *R: A language and environment for statistical computing.* Vienna, Austria.

Reinhart, T. (1981). Definite NP anaphora and c-command domains. *Linguistic Inquiry, 12,* 605–635.

Resnik, P. (1992). Left-corner parsing and psychological plausibility. In *Proceedings of COLING* (pp. 191–197).

Sarkar, D. (2006). *Lattice: Lattice graphics* [R package Version 0.14–16].

Szabolcsi, A. (2004). Negative polarity—Positive polarity. *Natural Language and Linguistic Theory, 22,* 409–452.

Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language, 50,* 355–370.

van der Wouden, T. (1997). *Negative contexts: Collocation, polarity, and multiple negation* (Vol. 1). London: Routledge.

Van Dyke, J., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language, 49,* 285–316.

Van Dyke, J., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language, 55,* 157–166.

Vasishth, S., & Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language, 82,* 767–794.

Warren, T. C., & Gibson, E. (2005). Effects of NP-type on reading English clefts. *Language and Cognitive Processes, 20*(6), 89–104.

Xiang, M., Dillon, B. W., & Phillips, C. (2006). Testing the strength of the spurious licensing effect for negative polarity items. In *Proceedings of the CUNY Sentence Processing Conference* (p. 148). New York.

Appendix

Stimuli used in the eye-tracking study

(5) (K)ein Chemiker, der (k)einen Kuchen gebacken hatte, war jemals/durchaus dumm.
(6) (K)ein Bettler, der (k)einen Geist gesehen hatte, war jemals/durchaus nüchtern.
(7) (K)ein Solist, der (k)eine Sonate gespielt hatte, war jemals/durchaus pünktlich.
(8) (K)ein Juwelier, der (k)einen Ring gefälscht hatte, war jemals/durchaus ängstlich.
(9) (K)ein Biologe, der (k)eine Brille aufgesetzt hatte, war jemals/durchaus gesprächig.
(10) (K)ein Polizist, der (k)einen Diebstahl beobachtet hatte, war jemals/durchaus taktvoll.
(11) (K)ein Junge, der (k)einen Kampf verloren hatte, war jemals/durchaus ordentlich.
(12) (K)ein Schüler, der (k)einen Baum gefällt hatte, war jemals/durchaus fleissig.
(13) (K)ein Elektriker, der (k)einen Stecker geprüft hatte, war jemals/durchaus verlässlich.
(14) (K)ein Säugling, der (k)eine Flasche getrunken hatte, war jemals/durchaus hungrig.
(15) (K)ein Professor, der (k)einen Fehler begangen hatte, war jemals/durchaus unterhaltsam.
(16) (K)ein Pirat, der (k)einen Braten gegessen hatte, war jemals/durchaus sparsam.
(17) (K)ein Künstler, der (k)eine Statue geschaffen hatte, war jemals/durchaus arrogant.
(18) (K)ein Kritiker, der (k)einen Vortrag gehalten hatte, war jemals/durchaus begeistert.
(19) (K)ein Angler, der (k)eine Fee erblickt hatte, war jemals/durchaus beschei.
(20) (K)ein Forscher, der (k)einen Schatz gefunden hatte, war jemals/durchaus faul.
(21) (K)ein Gärtner, der (k)eine Rechnung geschrieben hatte, war jemals/durchaus schwatzhaft.
(22) (K)ein Tourist, der (k)einen Anzug anprobiert hatte, war jemals/durchaus zufrieden.
(23) (K)ein Fleischer, der (k)einen Ochsen geschlachtet hatte, war jemals/durchaus kultiviert.
(24) (K)ein Wächter, der (k)eine Prügelei angezettelt hatte, war jemals/durchaus schläfrig.
(25) (K)ein König, der keinen Narren gehabt hatte, war jemals/durchaus beliebt.
(26) (K)ein Senator, der (k)einen Artikel verfasst hatte, war jemals/durchaus freundlich.
(27) (K)ein Leutnant, der (k)eine Taube geschossen hatte, war jemals/durchaus geduldig.
(28) (K)ein Pfarrer, der (k)einen Fisch gefangen hatte, war jemals/durchaus schweigsam.
(29) (K)ein Rentner, der (k)einen Nachbarn geärgert hatte, war jemals/durchaus tapfer.
(30) (K)ein Lehrling, der (k)einen Witz gemacht hatte, war durchaus aufgeregt.
(31) (K)ein Detektiv, der (k)einen Dieb gefasst hatte, war jemals/durchaus vorsichtig.
(32) (K)ein Artist, der (k)einen Trick geübt hatte, war jemals/durchaus tolpatschig.
(33) (K)ein Portier, der (k)eine Kabine gebucht hatte, war jemals/durchaus hässlich.
(34) (K)ein Jäger, der (k)einen Hochsitz gebaut hatte, war jemals/durchaus intelligent.
(35) (K)ein Archäologe, der (k)einen Krug vergraben hatte, war jemals/durchaus hastig.
(36) (K)ein Pianist, der (k)einen Auftritt erwartet hatte, war jemals/durchaus erfolgreich.
(37) (K)ein Sportler, der (k)einen Preis gewonnen hatte, war jemals/durchaus belesen.
(38) (K)ein Schaffner, der (k)eine Mütze getragen hatte, war jemals/durchaus nett.
(39) (K)ein Architekt, der (k)eine Skizze gezeichnet hatte, war jemals/durchaus sensibel.
(40) (K)ein Koch, der (k)einen Lutscher gekauft hatte, war jemals/durchaus schlank.