frontiers in
PSYCHOLOGY

Language Sciences

Distinctiveness and encoding effects in online sentence comprehension

Philip Hofmeister and Shravan Vasishth

Journal Name:	Frontiers in Psychology
ISSN:	1664-1078
Article type:	Original Research Article
Received on:	10 May 2014
Accepted on:	11 Oct 2014
Provisional PDF published on:	11 Oct 2014
www.frontiersin.org:	www.frontiersin.org
Citation:	Hofmeister P and Vasishth S(2014) Distinctiveness and encoding effects in online sentence comprehension. *Front. Psychol.* 5:1237. doi:10.3389/fpsyg.2014.01237

Frontiers in Language Sciences

Research Article
10 October 2014

Distinctiveness and encoding effects in online sentence comprehension

Philip Hofmeister [1,*] **and Shravan Vasishth** [2]

[1] *Department of Language and Linguistics, University of Essex, Colchester, United Kingdom*
[2] *Department of Linguistics, University of Potsdam, Potsdam, Brandenburg, Germany*
School of Mathematics and Statistics, University of Sheffield, Sheffield, South Yorkshire, United Kingdom

Correspondence*:
Philip Hofmeister
University of Essex, Department of Language and Linguistics, Colchester, United Kingdom, phofme@essex.ac.uk

ABSTRACT

In explicit memory recall and recognition tasks, elaboration and contextual isolation both facilitate memory performance. Here, we investigate these effects in the context of sentence processing: targets for retrieval during online sentence processing of English object relative clause constructions differ in the amount of elaboration associated with the target noun phrase, or the homogeneity of superficial features (text color). Experiment 1 shows that greater elaboration for targets during the encoding phase reduces reading times at retrieval sites, but elaboration of non-targets has considerably weaker effects. Experiment 2 illustrates that processing isolated superficial features of target noun phrases — here, a green word in a sentence with words colored white — does not lead to enhanced memory performance, despite triggering longer encoding times. These results are interpreted in the light of the memory models of Nairne 1990, 2001, 2006, which state that encoding remnants contribute to the set of retrieval cues that provide the basis for similarity-based interference effects.

Keywords: encoding, retrieval, similarity, distinctiveness, sentence processing

In everyday life and in laboratory experiments, people remember the unusual better than the usual. Von Restorff's classic findings illustrate this in terms of superior memory for isolated items, such as a bright green word in the context of a list of words colored black (**von Restorff**, 1933). More generally, a background of homogeneous stimuli favors the recall and recognition of contextually isolated stimuli. These so-called isolation effects share certain key characteristics with another set of memory effects tied to meaning-related processing. The latter include findings that people recall random trivia facts better if they subsequently hear causally-related information (**Bradshaw and Anderson**, 1982). Word recall and recognition benefits, too, from meaning-related processing (e.g., assessing the pleasantness of word meanings) compared with the processing of superficial features (e.g., identifying whether the word contains the letter 'e'), at least under conditions where the memory retrieval phase taps word meaning (**Craik and Lockhart**, 1972; **Hyde and Jenkins**, 1973; **Craik and Tulving**, 1975; **Stein et al.**, 1978).

Although clearly different in some respects (meaning-related processing is not typically taken to be 'unusual' or 'bizarre'), these two sets of effects can be thought of as being parallel in light of their relationship to both encoding and retrieval. In particular, elaboration and isolation each tend to give rise to longer encoding or study times. Elaboration, like isolation, also raises the probability of contextually unique features that serve to differentiate study items at retrieval, because elaboration typically yields highly diagnostic, meaning-related units of information. Thus, these two memory phenomena both potentially reflect a common set of core principles on the encoding-retrieval relationship and the dynamics of retrieval interference.

Correspondingly, mechanistic explanations for both kinds of effects have hinged on processes operative at either the encoding or the retrieval stage. From one view, the mnemonic benefits may arise from increased processing or attention during the encoding phase (**Hirshman et al.**, 1989; **Watkins et al.**, 2000; **Shiffrin**, 2003), leading to higher fidelity representations, more highly activated representations, or simply a richer set of self-generated features that form a partly redundant network with the core memory representation. This implies a type of investment-reward strategy; by paying for the cognitive costs of "enhanced" representational encoding, the costs of memory retrieval are lessened.

From a different but not mutually exclusive perspective, semantic processing increases the *distinctiveness* of the stimuli at the time of retrieval: "additional conceptual or semantic features help to differentiate the studied words from each other, making these memories less susceptible to interference and/or providing more features that can be cued on a typical recall or recognition memory test" (**Gallo et al.**, 2008, p. 1096; see also **Moscovitch and Craik**, 1976; **Fisher and Craik**, 1977; **Jacoby and Craik**, 1979; **Hunt and Worthen**, 2006). In other words, semantic processing of words trumps superficial processing because processing a word's meaning generates more contextually unique features than focusing on its sound or orthographic features. For instance, many words in memory may have the sound [aʊ] or the letter sequence 'ch'. But relatively few items in memory may be associated with features like 'sandy' and 'next to the ocean'. Consequently, such accounts predict more than a simple contrast between meaning-related and non-meaning related processing. If semantic processing increases the chances of conceptual distinctiveness, then as semantic processing increases, the chances for successful retrieval from memory should improve up to some arbitrary limit. One implication for at least some such distinctiveness accounts is that a memory target will contrast more with other stimuli, and hence be remembered better, if those *competing* representations elicit more semantic processing. That is, differentiation of two study items may in principle be modulated by the presence/absence of unique semantic features of *either* item, as adding contextually unique features to a competitor cuts down on potential overlap between a competitor and memory target.

Much of this prior research deals with explicit memory for language stimuli, particularly word lists. How linguistic representations are recovered in their most natural setting – online sentence processing — as a function of either elaboration or isolation has not played a significant part in this line of research. This is no doubt due to the implicit nature of memory retrieval during comprehension. Yet comprehending sentences perpetually requires reaccessing some previously perceived information, such as when a pronoun must be interpreted or when the subject of a verb needs to be remembered, and this prior content may vary considerably in the requisite amount of syntactic and semantic processing. Another context in which retrieval from memory happens is in so-called long-distance dependencies (a.k.a. filler-gap dependencies), as in (1):

(1) I finally gave up reading the novel that James Joyce wrote ___ in the 1930s.

To understand this sentence, "the novel" must be retrieved at the embedded verb "wrote" to be properly interpreted as the thematic patient. Evidence that memory retrieval of the argument takes place at the verb comes from reading time data, cross-modal priming tasks, neurophysiological studies, and speed-accuracy tradeoff data (**Tanenhaus et al.**, 1985; **Nicol and Swinney**, 1989; **Kluender and Kutas**, 1993; **Osterhout and Swinney**, 1993; **McElree**, 2000).

The purpose of the present investigation is to identify whether elaboration and isolation effects occur in online sentence processing and the extent to which such effects might be explained by relating encoding times to retrieval times. The working hypothesis, therefore, is that factors that predict the *success of explict recall* also contribute to the *efficiency of implicit retrieval*. While extant sentence processing models generally ignore variation in the encoding stage as a potential source of processing variation at retrieval sites, cue-based models of retrieval do predict that unique features in a memory target can facilitate retrieval (**McElree**, 2000; **Van Dyke and Lewis**, 2003; **Lewis and Vasishth**, 2005; **Lewis et al.**, 2006; **Van Dyke and McElree**, 2006, 2011). However, such theories do not make across-the-board predictions that targets with more semantic features, or contextually unique features, ought to be easier to retrieve. This is due to the fact that only those features cued by the retrieval trigger bear on assessments of similarity. For instance, in (2) below, "was complaining" initiates a retrieval probe targeting the animate subject NP "the resident":

(2) a. The worker was surprised that the resident who was living near the dangerous warehouse *was complaining* about the investigation. [= LOW INTERFERENCE]
b. The worker was surprised that the resident who said that the neighbor was dangerous *was complaining* about the investigation. [= HIGH INTERFERENCE]

In the high interference condition, the head and dependent are separated by an NP ("the neighbor") which is a type of semantic object that "can complain" and is also subject marked, similar to the retrieval target. The intervening NP in the low interference condition, in contrast, is inanimate and the object of a preposition, thus mismatching the target semantically and syntactically. Using such materials, **Van Dyke** (2007) observed evidence of a processing disruption in the high interference condition beginning at the key verbal cluster, which she interpreted in terms of the mechanics of cue-based retrieval. On such an account, features not in the retrieval probe triggered by the verb should have little bearing on memory interference. Whether a target is the only word to begin with an 'r' or appears in an unusual font should be immaterial to retrieval efficacy, for example, if verbs do not trigger retrieval probes containing such cues.

In the present experiments, key targets for implicit retrieval in long-distance dependencies differ in the amount of elaboration or "complexity" associated with them (Experiment 1), or with respect to the homogeneity of their text color with the surrounding text (Experiment 2). In both cases, the key features — prenominal modifiers and text color — are unlikely to be directly cued by the retrieval triggering verbs, i.e., verbs don't normally select arguments on the basis of color or the number of modifiers. If elaboration and isolation effects pattern in implicit memory retrieval tasks as they do in explicit memory tasks, then we should expect to see retrieval-related benefits in sentence processing given elaboration or isolation.

Recent reading time data provide some initial evidence that memory retrieval in sentence processing is sensitive to a memory target's representational complexity (**Hofmeister**, 2011). The term "complexity" is shorthand for the idea that discourse references can differ in semantic complexity via category hierarchy differences, e.g. "a thing" vs. "a stethoscope", as well as syntactic complexity. For instance, "the landmark on the bluff" encodes both syntactic and semantic features absent in "the landmark". In clefted constructions like those in (3), participants spent longer reading the head noun of the clefted element as the number of modifiers increased. At the words immediately following the subcategorizing verb (underlined below), however, reading times were faster given more features associated with the target. It is at this subcategorizing verb and the immediately following regions that we expect to observe signs of reactivation and retrieval of the representation in the cleft. Notably, the faster reading times for elaborated conditions do not appear until the subcategorizing verb or shortly thereafter:

(3) a. It was a communist that the members of the club banned <u>from ever entering</u> the premises.
b. It was an alleged communist that the members of the club banned <u>from ever entering</u> the premises.
c. It was an alleged Venezuelan communist that the members of the club banned <u>from ever entering</u> the premises.

Further experiments showed this same pattern even when holding the number of words and syntactic complexity constant, e.g. “which person” vs. “which soldier”. At least in some contexts, therefore, syntactic and semantic processing of linguistic representations facilitates their retrieval from memory. It further suggests that recoverability increases gradiently with semantic processing—something that the list memory literature has so far not shown.

The present self-paced reading studies expand upon these findings in several ways. In Experiment 1, not only the target noun phrase, but also a preceding non-target noun phrase varies in syntactic and semantic complexity. In (4), for example, the matrix object noun phrase is the target for retrieval at “encouraged” and appears in either elaborated or non-elaborated form:

(4) The (senior foreign) diplomat contacted the (ruthless military) dictator who the activist from the United Kingdom encouraged to preserve natural habitats and resources

In addition, the preceding matrix subject noun phrase also varies between an elaborated and non-elaborated form. This manipulation of a competitor’s complexity serves two purposes (note: “the activist” serves as a second potential competitor). First, it addresses the previously discussed question of whether elaborative processing linked to non-targets/competitors may facilitate differentiation at retrieval points. Such an idea is plausible from the perspective that providing more detail about any discourse referent or event lowers the chances that it will be confused with some other candidate for memory retrieval. Second, in (3) above, the key retrieval region (“banned”) appears later in the complex sentences than in the simpler ones, opening the door to an explanation based on word position effects. Due to the manipulation of the complexity of multiple phrases in Experiment 1, it will be possible to directly assess whether the effects observable at retrieval sites are reducible to word position effects.

In Experiment 2, the essential components of von Restorff’s design are carried over to the domain of sentence processing. Key words in the test sentences are systematically manipulated to make them superficially homogeneous or isolated with the expectation that this will give rise to longer encoding times. The question is whether superficial isolation or differentiation of words in sentences produces retrieval effects that are qualitatively similar to the effects of elaboration in online sentence processing. If they do, then we have evidence of a tight correspondence between implicit and explicit retrieval processes targeting linguistic stimuli.

As we shall see, both elaboration and isolation give rise to longer encoding times, but only the former yields strong evidence for faster reading times at sentence-internal retrieval sites. Moreover, while the elaboration associated with a non-target has striking downstream effects on encoding processes for other discourse referents, the evidence for an effect of non-target complexity on the retrieval of target representations is considerably weaker.

1 EXPERIMENT 1: TARGET & NON-TARGET COMPLEXITY

1.1 PARTICIPANTS

Fifty-two University of Essex undergraduates participated in this study for course credit or payment. All participants identified themselves as native English speakers without significant exposure to a second language before the age of five. No participant data was removed on the basis of accuracy, as all participants scored above 67% correct.

1.2 METHODOLOGY & MATERIALS

In this 2×2 self-paced, moving window experiment, 28 items varied in terms of the complexity of a target noun phrase and a non-target noun phrase in the same sentence. Specifically, all sentences contained a transitive matrix clause of the form [NP V NP], where the object noun phrase was modified by an object relative clause. The matrix subject (NP1) appeared with either 0 or 2 modifying words, as did the matrix object NP (NP2), as illustrated below:

(5) a. The congressman interrogated the general who a lawyer for the White House advised to not comment on the prisoners. (= SIMPLE SIMPLE)
b. The conservative U.S. congressman interrogated the general who a lawyer for the White House advised to not comment on the prisoners. (= COMPLEX SIMPLE)
c. The congressman interrogated the victorious four-star general who a lawyer for the White House advised to not comment on the prisoners. (= SIMPLE COMPLEX)
d. The conservative U.S. congressman interrogated the victorious four-star general who a lawyer for the White House advised to not comment on the prisoners. (= COMPLEX COMPLEX)

The subject of the object relative clause (NP3) was always of the form [DET NOUN]. At the critical embedded verb (*advised* in the example above), proper interpretation of the sentence requires retrieval of the representation referred to by NP2. It is also at such sentence internal retrieval sites that prior psycholinguistic evidence has repeatedly identified signs of similarity-based memory retrieval interference from competing representations (**Gordon et al.**, 2001, 2002, 2006; **Van Dyke and McElree**, 2006).

Each participant saw only one condition of each item. All sentences were followed by a yes/no comprehension question, and participants received feedback if they answered incorrectly. The comprehension questions targeted information about one of the three referents introduced in the sentence, e.g., *Was the general advised not to comment on the prisoners?* with numerous questions asking about the relationship between two referents, e.g, *Did a photographer embarrass a celebrity*? In Experiment 1, mean comprehension accuracy across all trials, including fillers, was 84% (min = 70%, max = 97%). 70 fillers accompanied the main experimental items for this experiment. 28 of these were from an unrelated experiment.

Materials were presented and randomized with the reading time software LINGER v. 2.94, developed by Doug Rohde (available at `http://tedlab.mit.edu/~dr/Linger/`). The experimental items were randomized by the experimental software, and at least one filler separated each critical item. At the beginning of each trial, a fixation cross at the left of the screen appeared on the same line where the target sentence subsequently appeared. On pressing a key, the cross disappeared and the first word of the sentence was shown. Words not currently being read were not presented on screen and were not masked with dashes, i.e., the screen was blank except for the word currently being read. We opted for this method to prevent participants from using end-of-sentence information to modulate their reading rate, since the target sentences differed in overall length.

Prior to statistical analysis, raw reading times greater than 5000 ms or less than 100 ms were removed, affecting a total of .001% of the data. No additional outlier removal processes were performed. All data were analyzed regardless of comprehension accuracy in order to capture any reading time differences that may reflect memory retrieval failures. In other words, as we are investigating not only retrieval efficiency but also success, excluding trials that were incorrectly responded to would eliminate an important and relevant subset of the data on which retrieval of the target NP potentially failed. However, in the Appendix, we also present secondary analyses using only data from correctly answered trials.

Reading times were log-transformed to normalize the residuals and reduce the effect of extreme data points. Then, the log reading times for all stimuli (fillers included) were regressed against several predictors known to affect reading times in self-paced reading tasks: word length and log list position (**Ferreira and Clifton**, 1986; **Hofmeister**, 2011). Specifically, longer words predict longer reading times and later list positions predict faster reading times as participants progress through the experiment. The model estimating these effects included a random effects term for participants, i.e., by-participant random intercept adjustments. We used data from fillers in this process to produce maximally general estimates of word length and list position. The residuals of this model – RESIDUAL LOG READING TIMES – are the dependent variable analyzed here (Figure 1 shows raw reading times to provide a more interpretable scale for the effects). All categorical predictors variables were sum coded to reduce effects of collinearity.

All analyses were conducted with Bayesian hierarchical models, fit with Stan and the `R` package `rstan`. We employed these models because they allow us to fit complex hierarchical models with maximal random effect structures that often do not converge using other popular linear regression packages such as `lme4`. Moreover, as noted in **Husain et al.** (2014), using Bayesian models allows us to assess and compare the weights of evidence for particular hypotheses. This means that we avoid categorizing effects as significant or non-significant, eschewing traditional statistical inference based on p-values. Instead, we make statistical inferences for particular hypotheses by computing the posterior probabilities for relevant parameters θ_i by sampling from their posterior distribution.

Each word region model used 4 chains, 5,000 samples per chain, a warm-up of 2,500 samples, and no thinning, resulting in 10,000 samples for each parameter estimate. All models contained fixed effect parameters for NP1 complexity, NP2 complexity, and their interaction. They also included by-participant random intercept adjustments and random slopes for all fixed effect terms (3 parameters), and by-item random intercept adjustments and random slopes for NP1 complexity, NP2 complexity, and their interaction (3 parameters). We utilized weak, uninformative priors for all key parameters, including participant and item adjustments. For each model, $P(\theta \mid data)$ indicates the probability that the parameter estimate is negative, i.e., speeding up occurs. For instance, an estimate that $P(\theta_{complex} < 0) = .99$ signifies that we can be 99% certain that complexity speeds up reading; in contrast, if $P(\theta_{complex} < 0) = .01$, we can infer with 99% certainty that complexity slows down reading. These probabilities were obtained by calculating the percentage of posterior samples above or below zero. To improve readability we will write $P(\theta < 0)$ for $P(\theta < 0 \mid data)$.

Three regions are analyzed in Experiment 1: the head noun of NP2, the head noun of NP3, and the verb that subcategorizes for NP2. As reading time effects in self-paced reading experiments often spill over onto subsequent words, results for the word regions immediately after the relevant sites are also reported. No significant effects of the experimental manipulations on comprehension accuracy were found so they are not discussed here (see data in Appendix).

1.3 RESULTS

NP2 head noun

As shown in Table 1, greater syntactic and semantic complexity of NP2 leads to longer reading times at this region. Greater complexity of NP1, however, has a weaker effect in the opposite direction. That is, reading times at the NP2 head noun were somewhat faster when NP1 was complex, compared to when it was syntactically and semantically simple. There is no compelling evidence for an interaction at this word region.

NP3 head noun + spillover

Complexity of NP2 also has an effect on reading times at the head noun of NP3 (e.g., “lawyer”): reading times are faster when NP2 is relatively complex. At the word immediately following the head noun (“for” in (5)), an interaction of NP1 & NP2 complexity arises, along with main effects of NP1 & NP2 complexity. This interaction stems from the fact that NP1 complexity leads to faster reading times only when NP2 is simple.

Relative clause verb + spillover

A main effect of NP2 complexity is evident at the critical relative clause verb: when NP2 is complex, reading times are faster than when NP2 is simple. Alongside this main effect, the results provide weak support of an interaction due to the fact that the complexity of NP1 affects reading times more when NP2 is simple. Put differently, there is no added processing facilitation due to the complexity of NP1 when NP2 is itself complex. The NP2 complexity effect also carries over onto the word immediately after the verb.

Table 1. Model summary for Experiment 1 for each region and fixed effect factor. Summary includes the posterior 95% Credible Interval (CrI), i.e., the lower CrI refers to the 2.5% bound and the upper CrI refers to the 97.5% bound. P($\beta < 0$) indicates the probability that complexity slows reading times effects, i.e., values closer to 0 indicate slowing down and values closer to 1 indicate speeding up due to complexity.

Region	*Factor*	*Mean*	*CrI lower*	*CrI upper*	*P($\beta < 0$)*
Head noun	NP1 complexity	−0.012	−0.033	0.009	.866
	NP2 complexity	0.038	0.011	0.065	.002
	NP1 × NP2 complexity	−0.003	−0.024	0.017	.621
RC subject head noun	NP1 complexity	−0.005	−0.025	0.014	.702
	NP2 complexity	−0.022	−0.039	−0.005	.996
	NP1 × NP2 complexity	0.002	−0.016	0.020	.397
RC subject head noun + 1	NP1 complexity	−0.014	−0.032	0.002	.955
	NP2 complexity	−0.012	−0.027	0.002	.951
	NP1 × NP2 complexity	0.013	−0.002	0.027	.043
RC verb	NP1 complexity	−0.005	−0.019	0.010	.757
	NP2 complexity	−0.014	−0.027	−0.001	.979
	NP1 × NP2 complexity	0.011	−0.005	0.025	.079
RC verb + 1	NP1 complexity	−0.004	−0.009	0.017	.281
	NP2 complexity	−0.019	−0.035	−0.004	.991
	NP1 × NP2 complexity	0.006	−0.007	0.019	.200

In fact, the effect is even more pronounced at this region. Here, signs of an interaction are considerably weaker, as illustrated in Figure 1.

Correctly answered trials only

We conducted secondary, post-hoc analyses using only data from correctly answered trials to determine whether the observed complexity effects were tied to trials where participants answered incorrectly. As depicted in Figure 2, all main findings persist in this data subset with NP2 complexity effects at the NP3 and the relative clause verb slightly increasing in magnitude.

1.4 DISCUSSION

When readers encode additional syntactic and semantic features, they read faster at sentence-internal retrieval sites. This pattern holds, however, primarily for NP2 — the downstream retrieval target. At the relative clause subject, reading times are faster when NP2 is syntactically and semantically complex, and this effect re-emerges at the retrieval triggering verb, continuing on into the spillover region.

Effects tied to NP1 — the preceding non-target — are comparatively weaker and tied to the status of NP2. Whereas the effects of the complexity of NP2 show up at the head noun of NP3, the impact of NP1 complexity does not emerge until the head noun's spillover region. More tellingly, NP1 complexity affects reading rates selectively: only when NP2 is simple, and hence syntactically similar to NP3, does greater NP1 complexity reduce reading times. At the retrieval region, too, effects of the complexity of NP1 are weak compared to those of NP2. While there are hints at the relative clause verb that NP1 complexity has some facilitatory effects, such effects (1) do not have the duration of those tied to NP2, (2) are statistically weaker, and (3) only appear when NP2 is simple. In essence, differences in the feature-based complexity of a competitor do not weigh as significantly on retrieval in sentence comprehension as differences in target complexity. This suggests rather specific constraints on the dynamics of encoding and retrieval with respect to the computation of similarity-based interference in sentence processing that are dealt with in the General Discussion.

Two notable conclusions can be drawn from these results. First, word position alone cannot account for the reading time differences at the retrieval sites. Inside the relative clause, the COMPLEX-SIMPLE and SIMPLE-COMPLEX conditions match each other with respect to word position, yet display different profiles at the word following the subcategorizing verb. Moreover, if elaboration effects at the retrieval region owe their existence to a basic linkage between word position and reading rate, then we would expect the reading times for the conditions to be ordered according to word position. However, the COMPLEX-COMPLEX condition proved to be no faster than the SIMPLE-COMPLEX, despite the retrieval region appearing two words later in the sentence. Second, the lack of a main effect of NP1 complexity at the retrieval region argues against a general preference for maximal descriptiveness. Indeed, nowhere in the sentence does there seem to be a notable advantage for modifying both NPs in the matrix clause. As noted above, however, NP1 complexity does impact the processing of NP3 when NP2 is simple. We take this to mean that encoding interference arises at NP3 when all the NPs match in form, but altering the form of either of the preceding NPs mitigates these interference effects.

A valid concern with respect to these data concerns the relationship between the effects at NP3 and the verb. Are these separate effects, or do the effects at the verb simply reflect extended spillover effects that originate with processing NP3 in the above stimuli? This concern is amplified by signs of NP2 complexity effects at the region before the retrieval-triggering verb. Several arguments, however, speak against the interpretation that the differences at the verb and its spillover region reflect a continuation of previously initiated processes. First, a separate analysis revealed that the NP2 complexity effect at the verb remains intact even after including reading times from the word before the verb as a covariate ($\hat{\mu} = -0.022$; CrI Lower = -0.038; CrI Upper = -0.006: $P(\beta < 0) = .997$). Second, consideration of only correctly answered trials shows that the effects at the verb are magnified, while differences at the preceding region are minimized (see Appendix for model summaries). Some of the variation across conditions immediately prior to the verb thus comes from trials where encoding or retrieval processes may have been compromised. Further supporting this interpretation, it was found that several poorly-performing participants (who averaged 56% correct on the critical trials) were the primary source of reading times differences at the word region preceding the verb. In the case of these participants, it is indeed possible that encoding difficulties continued on into the retrieval region.[1] Taken together, these observations support the interpretation that the effects at the verb and subsequent word reflect cognitive processes that begin at the verb.

2 EXPERIMENT 2

If complexity effects arise during sentence processing because additional semantic or conceptual features distinguish representations from one another, this raises the question of whether all types of unique features distinguish comprehension-based representations. There may be nothing special, mnemonically speaking, about syntactic and semantic features in comprehension. Experiment 2 consequently looks at whether unique features in general stimulate faster processing at retrieval sites in comprehension. But this experiment also has a secondary purpose. In Experiment 1, longer encoding times match up with shorter reading times at or directly after the retrieval site. Thus, one take on the previous results is that additional semantic features stimulate more processing, which facilitates downstream retrieval. By manipulating the homogeneity of superficial features in Experiment 2, we address both issues due to the expectation that isolated word stimuli will not only generate contextually unique features (by definition), but will also lead

[1] This might be taken as justification to exclude these participants altogether; however, we see no reason to exclude participants because they encounter more encoding problems or read less accurately than their peers.

to extended processing times during the encoding phase. The question is how this will bear, if at all, on the processing of words that trigger the retrieval of these encodings.

2.1 PARTICIPANTS

Forty-four UC-San Diego students participated in this study, in exchange for course credit. All subjects identified themselves as monolingual American English speakers without any known history of color blindness. The results from two participants were removed due to comprehension question accuracies below 67%.

2.2 METHODOLOGY & MATERIALS

Thirty-two items were constructed with an object noun phrase in a transitive main clause modified by an object relative clause, as in (6) below. Textually, the conditions were identical to each other.

(6) The congressman interrogated the **general** who the lawyer for the Bush administration advised ___ to not comment on the detainees.

To manipulate processing during the encoding phase, the head noun of the object NP ("general" above) appeared either in the same color as the surrounding sentence text (white), or else in an incongruent color (bright green). Additionally, the color of the word that triggered retrieval ("advised") also varied between congruent and incongruent. This second manipulation provides a needed check to ensure that participants do not read later word regions faster because of anticipation for an incongruently colored word. Moreover, in the condition with the green head noun and green verb, we can assess whether reinstating features of the encoding phase aids in retrieval. Hence, each item had four conditions (WHITE-WHITE, WHITE-GREEN, GREEN-WHITE, GREEN-GREEN), but each subject saw only one condition of each item.

Participants received instructions that the color of the words in the sentences was immaterial to the task and that they did not need to respond to color changes. Yes/no comprehension questions followed each item, and participants received negative feedback if they answered a question incorrectly. Sixty fillers accompanied these critical items: 20 with 0 green words, 20 with 1 green word, and 20 with 2 green words. For filler items with 1 green word, the word was randomly selected from all words in the sentence. For fillers with 2 green words, one appeared randomly in the the first half of the sentence and the other in the second half. All fillers had a syntactic structure different from that used in the critical items.

The materials were presented in a self-paced, center presentation paradigm via a propriety software package. Only one version of each item appeared on each of four experimental lists, whose contents were pseudo-randomized such that at least one filler intervened between each critical item. A fixation cross in the center of the screen appeared before each trial, and a comprehension question followed every experimental trial, including fillers. Participants received feedback only on incorrectly answered trials.

The outlier removal process, computation of residual log reading times, and Bayesian analysis procedure all followed those used in Experiment 1. As in that experiment, there were no differences in comprehension accuracy (GREEN-GREEN=76%, GREEN-WHITE= 77%, WHITE-WHITE = 76%, WHITE-GREEN= 76%). Here, we analyze residual log reading times at the head noun of the matrix object phrase and the relative clause verb that triggers its retrieval.

2.3 RESULTS

At the object head noun, incongruent, green words slow reading times, compared to the congruent, white words (see Figure 3). Similarly, looking at reading times at the retrieval region (*advised* in ((6)), a perceptually incongruent, green verb slows reading speed compared to a congruent, white one.

Table 2. Model summary for Experiment 2. Summary includes the posterior 95% Credible Interval (CrI), i.e., the lower CrI refers to the 2.5% bound and the upper CrI refers to the 97.5% bound. P($\beta < 0$) indicates the probability that incongruence slows reading times effects, i.e., values closer to 0 indicate slowing down and values closer to 1 indicate speeding up due to incongruence.

Region	*Factor*	*Mean*	*CrI lower*	*CrI upper*	*P($\beta < 0$)*
Head noun	Noun color	0.039	0.007	0.071	.008
	Verb color	0.005	−0.025	0.034	.362
	Noun color × Verb color	−0.001	−0.031	0.028	.532
RC verb	Noun color	−0.007	−0.029	0.016	.731
	Verb color	0.032	0.008	0.056	.001
	Noun color × Verb color	−0.008	−0.034	0.017	.734

In contrast to the pattern observed in Experiment 1, the increased encoding time at the object head noun due to superficial incongruence leads to relatively weak facilitation effects at the retrieval site, as shown in Table 2. In fact, the mean parameter value resides less than one standard deviation (= .011) from zero, according to the model results. The mean value for the condition where both the noun and the verb are incongruently colored reflects slightly faster reading than for the condition where only the verb is incongruent (GREEN-GREEN: −0.015, SE = .021 ; WHITE-GREEN: 0.011, SE = .024). This difference of roughly one standard error is why the model acknowledges a relatively weak effect of noun color (and an interaction with verb color) on reading times at the verb. At regions after the verb, there is no evidence that processing an incongruently colored target noun facilitates processing.

2.4 DISCUSSION

Increased processing times triggered by incongruent stimuli at the encoding site had weak effects on processing at the retrieval site when compared to the complexity effects observed in Experiment 1. Only when the relevant perceptual features were reinstated at the retrieval site was there any numerical retrieval advantage for perceptually incongruous stimuli. Even in this case, the facilitating effects were quite mild and would be deemed insignificant on classical frequentist methods of analysis. These findings imply that contextually unique features do not necessarily lead to improved memory performance, nor does increased processing time.

These findings may initially seem to contrast with memory results for recognition/recall of items presented in lists. For instance, **von Restorff** (1933) observed better recognition for words that appeared in superficially incongruent states. Similar findings of improved memory performance for superficially incongruent linguistic items (within mixed lists, but not unmixed lists) appear in **Bruce et al.**, 1976, **Hunt and Elliot**, 1980, **Hunt**, 1995, **Dunlosky et al.**, 2000, *inter alia*.

However, the current evidence reinforces the idea that the memory retrieval context is of utmost importance—a point frequently reiterated by memory researchers such as Tulving, Nairne, and others. In the present case, color or other superficial orthographical features rarely matter in written, sentence comprehension. Particularly if subjects are requested to ignore such information, there is little reason for subjects to recruit such potentially distinctive features in memory retrieval, whether or not they elicit more processing. In contrast, standard list recall or recognition tasks are novel encoding and retrieval contexts for participants—we are not standardly shown a list of words and then asked to retrieve them later, so we have few if any entrained habits. Consequently, in such novel circumstances, participants reasonably utilize all manner of perceptual features in recovering representations from memory.

Table 3. Similarity values and predicted sampling probabilities for two retrieval contexts

Cue	*Traces*	*Similarity*	*Samp. prob.*
[C C 2 3 1]	[C C 1 2 3]	.55	.26
	[C C 2 3 1]	1.0	**.48**
	[C C 3 1 2]	.55	.26
[C C 2 3 1 Q R N]	[C C 1 2 3]	.47	.24
	[C C 2 3 1 Q R N]	1.0	**.52**
	[C C 3 1 2]	.47	.24

In short, this experiment establishes that the uniqueness effects in language comprehension depend heavily on the retrieval context. What counts as unique critically depends on the nature and demands imposed at the retrieval site. Ultimately, if some set of representational features are unimportant for memory retrieval, then their congruence with other local feature appears to also have little import for memory retrieval.

3 GENERAL DISCUSSION

Increased processing during the encoding phase leads to more efficient retrieval processing in sentence comprehension, but only under certain conditions. Experiment 1 illustrated that increased processing associated with the downstream target benefits retrieval-related processing, whereas processing related to non-targets had relatively weak, short-lived effects that only arose when the target itself was not elaborated. Experiment 2 expanded on this by showing that not just any sort of extra processing facilitates memory (even for targets) — indeed, the results suggest that it is not about processing *per se* so much as the role of the features themselves in the retrieval process. In many respects, these results parallel the findings of studies assessing the effects of elaboration on long-term memory performance for linguistic stimuli (**Jacoby and Craik**, 1979; **Eysenck**, 1979; **Stein et al.**, 1978; **Reder**, 1980; **Bradshaw and Anderson**, 1982; **Reder et al.**, 1986; **McDaniel et al.**, 1988). At the same time, they add to these studies by showing that memory performance improves as meaning-related processing increases for linguistic stimuli in the context of sentence comprehension. Secondly, they demonstrate that these effects occur even in covert retrieval settings, where the time constraints of real-time comprehension limit the options for retrieval strategies. Third, the results from the final experiment demonstrate that unique representational target features and increased processing do not always lead to improved memory retrieval.

Both sets of findings —- the advantage of additional processing for targets compared to non-targets, and the fact that increased processing time does not necessarily benefit memory retrieval — can be understood through the lens of the short-term, feature-based retrieval model of **Nairne** (1990, 2001, 2006), with some minor new assumptions (several other memory models make similar predictions, e.g., **Oberauer and Kliegl**, 2006 and **Shiffrin**, 2003, although the details differ). In Nairne's model, memory items are represented as a vector of features, e.g., [C X 1 2 3]. Retrieval cues consist of lingering, typically blurry, records of the immediate past, e.g., [C X ? 2 3], as well as cues from the local retrieval context. In turn, these two sets of cues form a memory probe that is compared against a set of candidate memory items. The ultimate objective is to "redintegrate" the retrieval cues with a memory item, as the cues by themselves cannot be directly interpreted (**Ericsson and Kintsch**, 1995). The probability of retrieving an event E_1, given a retrieval probe X_1 depends upon the similarity or feature-overlap of X_1 and E_1, as well as the similarity of X_1 to other memory candidates:

$$P_r(E_1|X_1) = \frac{s(X_1, E_1)}{\sum s(X_1, E_n)} \quad (1)$$

The similarity between a memory item and a retrieval probe is determined by the number of mismatching features divided by the total number of compared features (*d*):

$$s(X_1, E_1) = e^{-d(X_1, E_1)} \quad (2)$$

Because retrieval probes consist of remnants of the original encoding process that need to be interpreted by comparing them against candidate memory items, any contextually unique features in a target will improve the chances for successful retrieval. In short, a target's recoverability increases if it possesses a feature that no other competitor shares.

Nairne (2006) employs this model to explain isolation or distinctiveness effects, since odd/bizarre items possess features that mismatch with the features of some homogeneous background set. For instance, imagine a context where the original encoding is perfectly intact and acts as the sole source of retrieval cues, e.g., $X_1 = E_1$. Any contextually unique features will increase the dissimilarity or mismatch between the retrieval cues and competitors, even though contextual uniqueness does not directly affect the similarity value between the target and retrieval probe.

An implied consequence of such a theory is that simply adding features to a target is predicted to increase the odds of sampling from memory, so long as these features are unique. Table 4 shows how the probability of sampling a target increases as the number of mismatching features between the target and non-targets increases, even though the number of shared features remains constant (see **Hofmeister et al.**, 2013 for an application of this model to the processing and acceptability of multiple wh-questions in English). The added features Q, R, & N in the undegraded probe lack any correlates in the competitors, meaning that the mismatch between them and the probe increases, effectively upping the chances for sampling the target.

As Figure 4 illustrates (left panel), the effect of adding mismatching or contextually unique features faces some restrictions: increasing the number of mismatches yields diminishing returns, ultimately asymptoting at a level that depends upon the number of features involved and the number of feature matches. In less formal terms, adding a little unique, diagnostic information can be quite helpful for memory retrieval, but adding lots of unique information is not likely to contribute much more. This model also predicts that the number of competitors affects retrieval probability much more dramatically than the number of overlapping features. On the right side, Figure 4 shows that going from one competitor to three competitors which each share two features with the probe nearly halves the chances of retrieval. In contrast, the difference between two competitors with 2 vs. 10 matching features never exceeds 10% (see left side of Figure 4).

A key component of this type of model is that a fragile copy of the original encoding process stored in primary memory provides a source of retrieval cues. This makes explicit the idea that syntactic and semantic features not directly invoked by the local sentence context can influence retrieval processes, in contrast to assumptions that only the similarity of features "grammatically derived from the current word and context" enter into considerations of similarity-based interference (**Lewis et al.**, 2006, p. 448).[2] Sentence processing models built upon the latter kind of assumption face difficulty explaining some classic retrieval interference effects in the sentence processing literature (**Logačev and Vasishth**, 2012). For instance, **Gordon et al.** (2001) show that processing in object-cleft sentences like (7) is easier at

[2] Current sentence processing models are not without means to explain effects of complexity on memory retrieval. For instance, on the ACT-R-based theory of **Lewis and Vasishth** (2005), processing syntactic material that modifies some previously constructed representation requires the restoration of the stored memory item. This retrieval process, in turn, raises the overall activation level of the item, making it easier to retrieve subsequently. Thus, complexity-based effects on retrieval emerge most straightforwardly as the byproduct of encoding processes. Moreover, additional study time potentially allows for more accurate encoding, providing greater chances that target features will be cued at the retrieval site (see also **Shiffrin**, 2003). However, as retrieval cues are limited to those provided by local grammatical context, there is no guarantee that unique semantic or syntactic features will factor into estimates of similarity and thus retrieval difficulty.

the subcategorizing verb ("saw") when the two NPs are of different types (proper name vs. definite description), but that such effects are absent in subject relativization constructions:

(7) It was John/the barber that the lawyer/Bill saw in the parking lot.

These effects are commonly understood in terms of similarity-based interference: if the target noun phrase overlaps in form with another local noun phrase that appears before the verb, memory retrieval difficulty ensues, ostensibly because the retrieval cues match multiple memory representations. As the second NP occurs after the verb in subject relatives, no possibility for interference exists. Notably, the verb triggering retrieval ("saw") does not itself supply cues as to the nominal type of the clefted element; indeed, no language appears to explicitly code whether a verb requires a lexical, pronominal, or some other type of nominal argument. So, if the similarity effects arise because retrieval cues match multiple representations, then those cues must come from a source besides the verb. The original encoding of the target provides the most obvious source of such cues. Not only does this open up a way to explain similarity-based effects due to overlapping referential form, it can also accommodate phonological similarity effects such as the observed reading time contrast at the embedded verb in sentences like "The baker that the banker sought found the house" vs. "The runner that the banker sought found the house" (**Acheson and MacDonald**, 2011).[3]

The current findings add a further data point to our developing picture of similarity-based interference in sentence processing: non-target distinctiveness has a weaker role to play in retrieval interference than target distinctiveness. These effects can be straightforwardly accommodated with some specifications about how similarity is calculated. Following Nairne (2006), let's assume that similarity at retrieval sites is calculated by establishing mismatches with the lingering features of a target's encoding remnant and any other features in the retrieval probe. A memory probe such as [C X 1 2 3] will mismatch equally with a competitor representation like [C X 4 5 6] as [C X 4 5 6 L M], e.g., 3 out of 5 probe features will mismatch with competitor features. In other words, it is the number of features in the probe that determine how many mismatches there can be, and not the number of features in a memory retrieval candidate. Adding unique features to some non-target, therefore, will not directly affect the probability of sampling the target because it does not contribute to the set of retrieval cues.

The data hint nonetheless at some retrieval effects linked to the elaboration of non-targets, specifically when the retrieval target itself was syntactically and semantically simple. This would seem to initially contradict the above view that the uniqueness of non-targets does not directly bear on retrieval efficiency. There is no contradiction, however, if these non-targets effects are byproducts of encoding interference. That is, we presume that the uniqueness of features in non-target nominals affects how other local nominals, including downstream targets, are encoded, and indirectly influence retrieval operations *via such encoding effects*. Even more generally, encoding interference feeds into retrieval interference.

Already, evidence exists that similarity between linguistic representations in memory and those being encoded can lead to processing disruptions, during both encoding and retrieval stages (**Gordon et al.**, 2002; **Acheson and MacDonald**, 2011). For example, **Gordon et al.** (2002) provide evidence of reading slowdowns when words on a sentence-external memory list are similar to key words inside the sentence, e.g., proper names vs. definite descriptions, both at the encoding site for the sentence-internal words and later at retrieval sites for those same words. We would add to this by hypothesizing that encoding interference may contribute to the degradation of memory representations, following research that suggests that forgetting in short-term memory for linguistic representations can stem from feature overwriting (**Oberauer and Lange**, 2008; **Oberauer**, 2009). Because these features that are susceptible to overwriting also contribute to retrieval cues on the account sketched above, feature loss could compromise any cue-based retrieval process.

[3] Acheson and MacDonald (2011) illustrate similar effects in subject relatives, as well, suggesting that phonological similarity gives rise to encoding interference and not simply retrieval interference.

Applying these hypotheses to the results of Experiment 1, encoding interference emerges as an indirect (and accordingly, weaker) contributor to retrieval differences, beyond what is predicted by the model of memory retrieval inspired by Nairne. Specifically, similarity between the referring expressions determines encoding interference, which can affect the integrity of the trace for the target nominal. So, when NP1 is complex and NP3 is simple or vice versa, this translates to a reduced danger of feature overwriting, compared to when they are both simple. In turn, the potential for retrieval interference is mitigated when the two initial NPs mismatch in complexity, because the trace for NP2 is more likely to be intact. Things are somewhat more complicated when NP1 and NP2 are both complex: while overlapping in structural form, the NPs carry more unique semantic features than their simpler counterparts. In this case, we tentatively take the results to mean that encoding interference is relatively low, compared to the case where both NPs are simple, but not any lower than when just one such NP is complex. These ideas require further tests to be substantiated, as the current experiments were not designed to test them. Nonetheless, we maintain that the relatively weak effects of non-targets can best be explained by appealing to the effect of encoding interference on memory retrieval.

Notably, redintegration-based models of memory do not require that every perceivable feature matters for memory retrieval. Listeners or readers may preferentially not encode some features in typical language settings, such as modality-specific features or exclude such features from the retrieval probe based on prior experience of the efficacy of such features. The advantage of increased processing thus depends upon the discourse context and the extent to which processing engenders unique features that come into play during the retrieval stage. From this perspective, encoding manipulations cannot have a predictable effect on memory in the absence of information about the encoding and retrieval contexts – what other memory candidates are available and what the retrieval cues are.

The results of Experiment 2 align with this perspective, in light of the absence of isolation or superficial processing effects. Modality-dependent features, such as orthography, font style, text color, etc., often play a large role in various laboratory tests of memory and in effects such as the auditory recency effect, but they appear to have a lesser role in guiding retrieval in sentence processing contexts. Such contrasts, though, are explicable in terms of task demands and prior experience. Word recall and recognition tasks lie outside the typical range of personal pastimes, whereas sentence comprehension is an everyday occurrence. This arguably leads participants to utilize a wider range of possible retrieval cues in word recall tasks, whereas prior experience with sentence processing would bias against the use of modality-specific features to distinguish memory representations. Instead, modality-independent features — properties that largely remain constant across presentations or modalities such as syntactic category and meaning — provide the basis for restoring linguistic representations during sentence processing because of their diagnostic potential. Thus, it is due to the fact that discrimination between language representations in sentence comprehension depends on syntactic and semantic features that the uniqueness of these features bears on determinations of retrieval ease and success. Correspondingly, the primary source of retrieval difficulty in language comprehension – overlapping semantic and syntactic representations and the resulting interference – is what gives additional linguistic processing mnemonic value, and why other types of processing such as superficial processing have little mnemonic value.

4 CONCLUSION

These tests of implicit memory establish that elaboration effects occur in online sentence processing tasks, as they do in explicit tests of memory. In Experiment 1, we found that increased processing of syntactic and semantic features connected to the target benefits memory retrieval in sentence processing; however, additional processing directed towards non-targets had substantially weaker effects on processing at retrieval sites. In Experiment 2, it was established that the processing of superficial features or features connected to non-targets yielded insubstantial processing advantages at retrieval sites, despite leading to longer encoding times. As sentence processing demands differ from those of explicit memory tasks, it is unsurprising that the effects of encoding manipulations can differ drastically across tasks with inherently different retrieval contexts. This apparent dynamic interaction between encoding and retrieval led **Tulving** (1983, p. 239) to argue against any statements of the form that "encoding operations of class X are more effective than encoding operations of class Y" (see also **Neath and Surprenant**, 2005 for a recent review). In short, encoding manipulations are unpredictable without additional information about the nature of the retrieval task and the background of competing representations.

The comparison of memory findings in the broader psychology and psycholinguistics literature also led to a unified theoretical account of distinctiveness effects, applicable across tasks. Capturing the interplay between representational uniqueness and retrieval probability, Nairne's feature-based model provides a means for introducing retrieval cues that are unlikely to be cued by local grammatical memory triggers via the use of a fragile copy of the original encoding. This fills a critical gap in cue-based models of retrieval in sentence processing by pointing to alternative sources of retrieval cues beyond the local context, thus accounting for a variety of otherwise unexplained similarity-based effects in sentence processing.

ACKNOWLEDGEMENTS

Valuable feedback on this article was received from Marta Kutas, Rick Lewis, and Michael Shvartsman. This research was supported by NIH Training Grant T32-DC000041 via the Center for Research in Language at UC-San Diego.

REFERENCES

Acheson, D. J. and MacDonald, M. C. (2011), The rhymes that the reader perused confused the meaning: Phonological effects during on-line sentence comprehension, *Journal of Memory and Language*, 65, 2, 193–207

Bradshaw, G. and Anderson, J. R. (1982), Elaborative encoding as an explanation of levels of processing, *Journal of Verbal Learning and Verbal Behavior*, 21, 165–174

Bruce, D., Gaines, I., and Marion, T. (1976), Tests of an organizational hypothesis of isolation effects in free recall, *Journal of Verbal Learning and Verbal Behavior*, 15, 1, 59–72

Craik, F. and Lockhart, R. (1972), Levels of processing: A framework for memory research, *Journal of Verbal Learning and Verbal Behavior*, 11, 6, 671–684

Craik, F. and Tulving, E. (1975), Depth of processing and the retention of words in episodic memory, *Journal of Experimental Psychology: General*, 104, 3, 268–294

Dunlosky, J., Hunt, R., and Clark, E. (2000), Is perceptual salience needed in explanations of the isolation effect?, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 3, 649–657

Ericsson, K. and Kintsch, W. (1995), Long-term working memory, *Psychological Review*, 102, 2, 211–244

Eysenck, M. (1979), Depth, elaboration, and distinctiveness, in L. Cermak and F. Craik, eds., Levels of Processing in Human Memory (Lawrence Erlbaum Associates, Hillsdale, NJ), 89–118

Ferreira, F. and Clifton, C. (1986), The independence of syntactic processing, *Journal of Memory and Language*, 25, 3, 348–368

Fisher, R. and Craik, F. (1977), Interaction between encoding and retrieval operations in cued recall, *Journal of Experimental Psychology: Human Learning and Memory*, 3, 6, 701–11

Gallo, D., Meadow, N., Johnson, E., and Foster, K. (2008), Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task, *Journal of Memory and Language*, 58, 1095–1111

Gordon, P., Hendrick, R., and Johnson, M. (2001), Memory interference during language processing., *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 6, 1411–23

Gordon, P., Hendrick, R., Johnson, M., and Lee, Y. (2006), Similarity-based interference during language comprehension: Evidence from eye tracking during reading, *Journal of Experimantal Psychology: Learning, Memory, Cognition*, 32, 6, 1304–1321

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

Hirshman, E., Whelley, M., and Palij, M. (1989), An investigation of paradoxical memory effects, *Journal of Memory and Language*, 28, 594–609

Hofmeister, P. (2011), Representational complexity and memory retrieval in language comprehension, *Language and Cognitive Processes*, 26, 3, 376–405

Hofmeister, P., Jaeger, T., Sag, I., Arnon, I., and Snider, N. (2013), The source ambiguity problem: distinguishing effects of grammar and processing on acceptability judgments, *Language & Cognitive Processes*, 28, 48–87

Hunt, R. (1995), The subtlety of distinctiveness: What von Restorff really did, *Psychonomic Bulletin & Review*, 2, 1, 105–112

Hunt, R. and Elliot, J. (1980), The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention, *Journal of Experimental Psychology: General*, 109, 1, 49–74

Hunt, R. and Worthen, J. (2006), Distinctiveness and Memory (Oxford University Press, Oxford)

Husain, S., Vasishth, S., and Srinivasan, N. (2014), Strong expectations cancel locality effects: Evidence from Hindi, *PloS one*, 9, 7, e100986

Hyde, T. and Jenkins, J. (1973), Recall for words as a function of semantic, graphic, and syntactic orienting tasks, *Journal of Verbal Learning and Verbal Behavior*, 12, 5, 471–480

Jacoby, L. and Craik, F. (1979), Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context, in L. Cermak and F. Craik, eds., Levels of Processing in Human Memory (Lawrence Erlbaum Associates, Hillsdale, NJ), 1–21

Kluender, R. and Kutas, M. (1993), Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies, *Journal of Cognitive Neuroscience*, 5, 2, 196–214

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

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

Logačev, P. and Vasishth, S. (2012), Case matching and conflicting bindings interference, in M. Lamers and P. de Swart, eds., Case, Word Order and Prominence (Springer), 187–216

McDaniel, M., Dunay, P., Lyman, B., and Kerwin, M. L. (1988), Effects of elaboration and relational distinctiveness on sentence memory, *The American Journal of Psychology*, 101, 357–369

McElree, B. (2000), Sentence comprehension is mediated by content-addressable memory structures, *Journal of Psycholinguistic Research*, 29, 2, 111–123

Moscovitch, M. and Craik, F. (1976), Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall, *Journal of Verbal Learning and Verbal Behavior*, 15, 4, 447–458

Nairne, J. (1990), A feature model of immediate memory, *Memory & Cognition*, 18, 3, 251–269

Nairne, J. (2001), A functional analysis of primary memory, in H. Roediger, J. S. Nairne, and A. Suprenant, eds., The Nature of Remembering: Essays in Honor of Robert G. Crowder (American Psychological Association, Washington, D.C.), 282–296

Nairne, J. (2006), Modeling distinctiveness: Implications for general memory theory, in R. Hunt and J. Worthen, eds., Distinctiveness and Memory (Oxford University Press, New York), 27–46

Neath, I. and Surprenant, A. (2005), Mechanisms of memory, *Handbook of Cognition*, 221–238

Nicol, J. and Swinney, D. (1989), The role of structure in coreference assignment during sentence comprehension, *Journal of Psycholinguistic Research*, 18, 1, 5–19

Oberauer, K. (2009), Interference between storage and processing in working memory: Feature overwriting, not similarity-based competition, *Memory & Cognition*, 37, 3, 346–357

Oberauer, K. and Kliegl, R. (2006), A formal model of capacity limits in working memory, *Journal of Memory and Language*, 55, 4, 601–626

Oberauer, K. and Lange, E. (2008), Interference in verbal working memory: Distinguishing similarity-based confusion, feature overwriting, and feature migration, *Journal of Memory and Language*, 58, 3, 730–745

Osterhout, L. and Swinney, D. (1993), On the temporal course of gap-filling during comprehension of verbal passives, *Journal of Psycholinguistic Research*, 22, 2, 273–286

Table 4. Model summary for correctly answered trials only from Experiment 1 for regions from one word before relative clause verb to one word after. Summary includes the posterior 95% Credible Interval (CrI), i.e., the lower CrI refers to the 2.5% bound and the upper CrI refers to the 97.5% bound. P($\beta <$ 0) indicates the probability that complexity slows reading times effects, i.e., values closer to 0 indicate slowing down and values closer to 1 indicate speeding up due to complexity.

Region	*Factor*	*Mean*	*CrI lower*	*CrI upper*	*P($\beta < 0$)*
Preceding region	NP1 complexity	−0.002	−0.021	0.016	.596
	NP2 complexity	−0.009	−0.026	0.007	.871
	NP1 × NP2 complexity	−0.003	−0.018	0.012	.653
RC verb	NP1 complexity	0.002	−0.013	0.017	.369
	NP2 complexity	−0.019	−0.035	−0.003	.989
	NP1 × NP2 complexity	0.013	−0.004	0.030	.070

Reder, L. (1980), The role of elaboration in the comprehension and retention of prose: A critical review, *Review of Educational Research*, 50, 1, 5–53

Reder, L., Charney, D., and Morgan, K. (1986), The role of elaborations in learning a skill from an instructional text, *Memory and Cognition*, 14, 1, 64–78

Shiffrin, R. (2003), Modeling memory and perception, *Cognitive Science*, 27, 3, 341–378

Stein, B., Morris, D., and Bransford, J. (1978), Constraints on effective elaboration, *Journal of Verbal Learning and Verbal Behavior*, 17, 6, 707–714

Tanenhaus, M., Carlson, G., and Seidenberg, M. (1985), Do listeners compute linguistic representations?, in A. Zwicky, L. Kartunnen, and D. Dowty, eds., Natural Language Parsing: Psycholinguistic, Theoretical, and Computational Perspectives (Cambridge University Press, London and New York), 359–408

Tulving, E. (1983), Elements of Episodic Memory (Clarendon Press, Oxford)

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

Van Dyke, J. A. (2007), Interference effects from grammatically unavailable constituents during sentence processing, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 2, 407

Van Dyke, J. A. and McElree, B. (2011), Cue-dependent interference in comprehension, *Journal of Memory and Language*, 65, 3, 247–263

von Restorff, H. (1933), Über die Wirkung von Bereichsbildungen im Spurenfeld, *Psychological Research*, 18, 1, 299–342

Watkins, M., LeCompte, D., and Kim, K. (2000), Role of study strategy in recall of mixed lists of common and rare words, *Journal of Experimental Psychology: Learning Memory and Cognition*, 26, 239–245

5 APPENDIX

5.1 EXPERIMENT 1 MATERIALS

1. The federal prison warden punished the condemned political prisoner who the guard at the old facility supplied with illegal goods like cigarettes.

2. The important, prospective client harassed the attractive personal secretary who the co-worker at the giant firm accused of drinking on the job.

3. The regular, valued customer trusted the sleazy car salesman who the mechanic at the auto shop helped to hide the engine problems.

4. The experienced stage choreographer embraced the graceful ballet dancer who an admirer in the front row begged for an autograph for his collection.

5. The Marine Corps sergeant dismissed the wounded American soldier who the commander at the army base ordered to search every car that approached the checkpoint.

6. The noisy, obnoxious drunk heckled the hilarious stand-up comedian who the bartender with a strong accent defended from the back of the room.

7. The accident insurance investigator questioned the emotional crash survivor who the pilot with a broken arm pulled from the wreckage of the plane.

8. The brutal military policeman arrested the peaceful Buddhist monk who a councillor in the capital city saved from being imprisoned and tortured.

9. The conservative US congressman interrogated the victorious four-star general who a lawyer for the White House advised to not comment on the prisoners.

10. The young, rebellious teenager hated the struggling rock musician who a scout for a record company discovered several years ago in Paris.

11. The alleged bombing accomplice misled the undercover federal agent who a witness in the terrorism trial notified of suspicious activity outside the Parliament building.

12. The celebrity hair stylist offended the renowned fashion designer who a journalist at the spring exhibition asked about the likely trends for next year.

13. The senior electrical engineer disliked the nerdy computer programmer who a hacker on the company network ridiculed for having pathetic security safeguards.

14. The critical newspaper reviewer applauded the famous young actor who the director of the art film ignored during the opening night festivities.

15. The senior foreign diplomat contacted the ruthless military dictator who the activist from the United Kingdom encouraged to preserve natural habitats and resources.

16. The hard-nosed newspaper reporter interviewed the wealthy teen celebrity who the photographer for the tabloid magazine embarrassed last week at a charity dinner.

17. The capital murder defendant feared the corrupt homicide cop who the judge in the court case silenced after a disturbing courtroom outburst.

18. The hard-working factory employee obeyed the professional shift supervisor who an inspector from the county government cautioned about the poor work conditions.

19. The cable news analyst ridiculed the Labour mayoral candidate who the leader of the political party supported despite the anxiety of other members.

20. The chief CIA interrogator questioned the heartless former mercenary who the commander of the armed rebellion shot in the foot without any explanation.

21. The wanted, dangerous fugitive robbed the rich American tourist who the guide for the London tour warned about straying too far from the group.

22. The wealthy German industrialist threatened the poor Russian peasant who the investor from a foreign land protected without the least bit of hesitation.

23. The powerful business executive infuriated the liberal socialist politician who a lobbyist the oil company bribed to vote for the upcoming bill.

24. The influential legal advisor lectured the handsome blonde prince who the official learning to speak English thanked at the end of the ceremony.

25. The elderly passing pedestrian dodged the angry taxi driver who a bystander near the bus stop identified later on at the police station.

26. The happy, suburban housewife offended the helpful new assistant who the manager on a lunch break called into his office after the incident.

27. The trained hospice nurse consoled the dying elderly patient who the doctor on a double shift forgot due to a lack of sleep.

28. The private office secretary aggravated the young female intern who a partner at the law firm hired less than three weeks ago.

5.2 EXPERIMENT 2 MATERIALS

1. The warden punished the prisoner who the guard at the federal prison supplied with illegal contraband like cigarettes.

2. The client harrassed the secretary who the coworker at the giant firm noticed drinking on the job occasionally.

3. The customer doubted the salesman who the mechanic at the auto shop believed to be a deplorable crook.

4. The housewife visited the priest who the bishop with an intimidating presence commended for a life of dedication to the church.

5. The choreographer embraced the dancer who the admirer in the front row begged for an autograph for his daughter.

6. The sergeant relieved the soldier who the commander of the army troops ordered to search every car that approached the checkpoint.

7. The drunkard heckled the comedian who the bartender on his lunch break defended from the back of the room.

8. The investigator questioned the survivor who the pilot with a broken arm pulled from the burning fuselage.

9. The policeman arrested the monk who the councilor in the capital city saved from being imprisoned and tortured.

10. The congressman interrogated the general who the lawyer for the Bush administration advised to not comment on the detainees.

11. The teenager hated the musician who the girl listening to the song adored since she saw the late night talk show performance.

12. The accomplice misled the agent who the witness in the impeachment trial notified about the presence of two suspicious characters.

13. The stylist offended the designer who the interviewer at the spring exhibition asked about the likely trends for next year.

14. The engineer debated the programmer who the hacker on the company network ridiculed for having pathetic security safeguards.

15. The reviewer criticized the actor who the director of the art film ignored during the opening night festivities.

16. The diplomat contacted the dictator who the activist from the United Kingdom encouraged to preserve natural habitats and resources.

17. The reporter interviewed the celebrity who the photographer for the tabloid magazine embarrassed last week at a charity dinner.

18. The defendant accused the cop who the judge in the murder case silenced after a disturbing courtroom outburst.

19. The guard aided the criminal who the agent from the U.S. marshals apprehended following a long and tiring chase.

20. The employee obeyed the supervisor who the inspector of the safety measures cautioned about the poor work conditions.

21. The captain evaluated the fireman who the veteran with 30 years' experience trained over the course of six months.

22. The investigator summoned the athlete who the coach of the football team invited to try out for a spot on the team.

23. The pundit ridiculed the candidate who the leader of the political party supported despite the misgivings of other members.

24. The interrogator questioned the mercenary who the commander of the armed rebellion abandoned without any explanation or warning.

25. The fugitive robbed the tourist who the guide for the huge group warned about straying too far from the group.

26. The industrialist threatened the peasant who the investor from a foreign land protected without the least bit of hesitation.

27. The executive infuriated the politician who the lobbyist for some oil companies bribed to vote for the upcoming bill.

28. The advisor lectured the prince who the dignitary learning to speak English thanked at the end of the ceremony.

29. The pedestrian dodged the driver who the bystander near the bus stop identified later on at the police station.

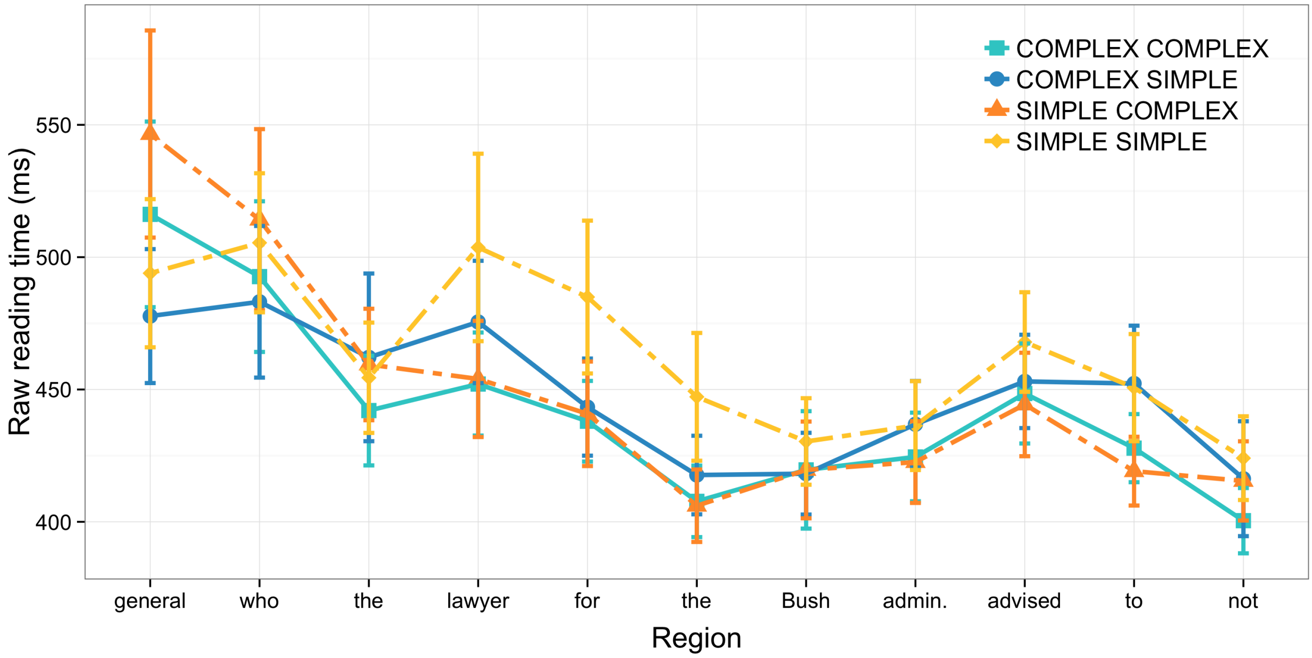


Figure 1. Raw reading times in Experiment 1 by region; error bars show 95% confidence intervals.

30. The customer offended the assistant who the manager of the department store called into his office after the incident.

31. The nurse consoled the patient who the doctor on a double shift forgot due to a lack of sleep.

32. The secretary aggravated the intern who the partner at the law firm hired less than three weeks ago.

6 FIGURES

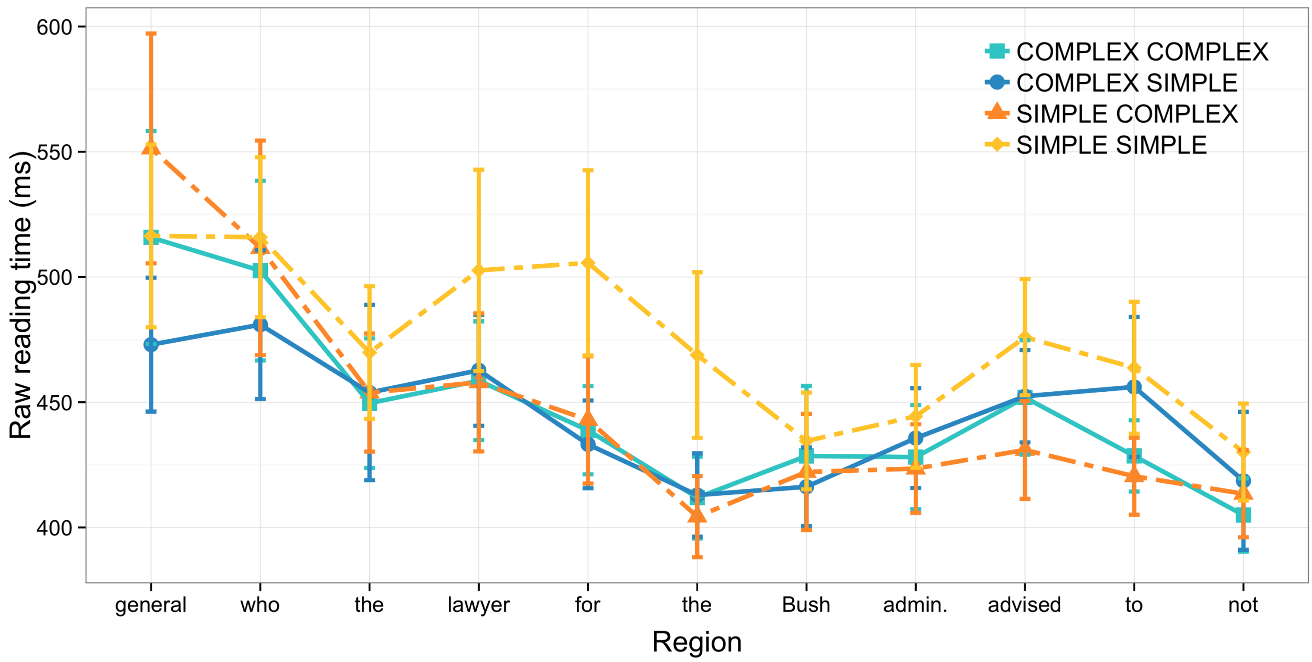


Figure 2. Raw reading times in Experiment 1 for correctly answered trials; error bars show 95% confidence intervals.

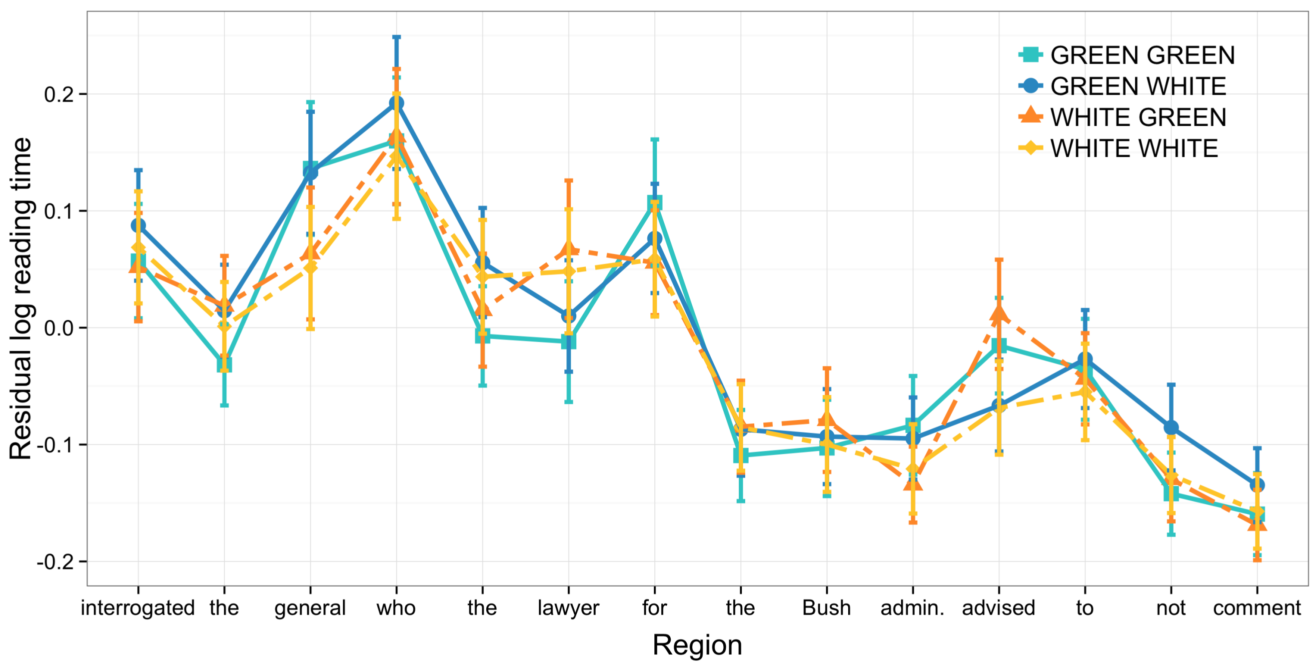


Figure 3. Residual log reading times at verb in Experiment 2; error bars show 95% confidence intervals.

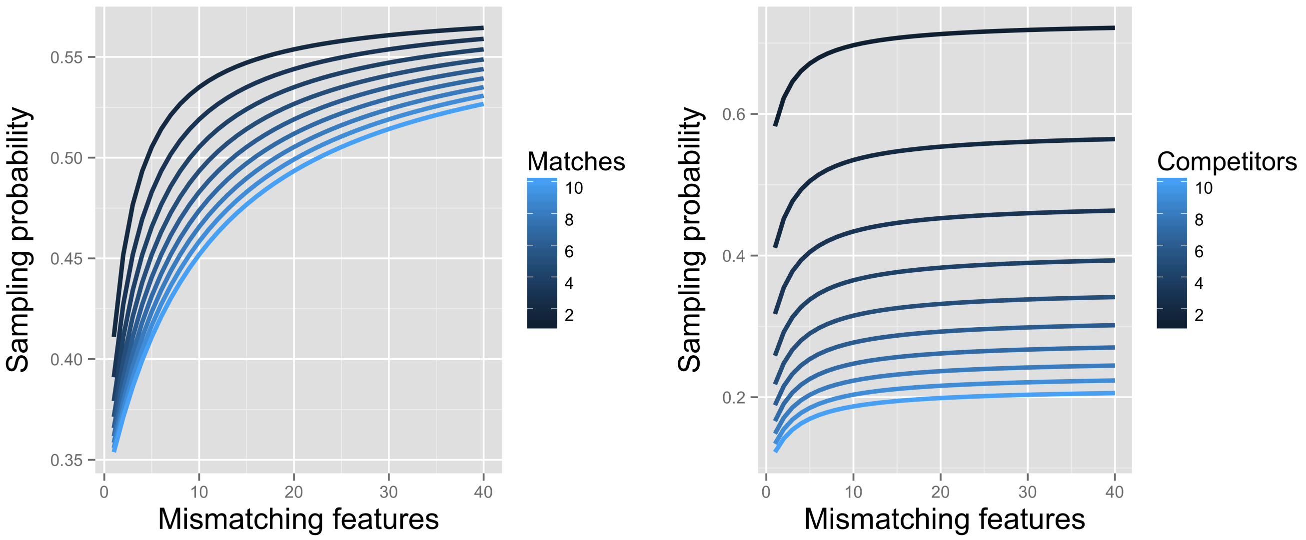


Figure 3. LEFT: Relationship between number of unique target features (mismatching with non-targets) and average sampling probability of target with two competitors. In descending order, the lines show the varying sampling probability curves for 2 to 10 probe features matching with each competitor. RIGHT: Relationship between number of unique target features and average sampling probability of target as a function of the number of competitors (from 1 to 10 in descending order), assuming two matching features between the probe and each competitor. The retrieval sampling curves illustrate the diminishing effects of mismatching features and the relatively greater effect of the number of competitors compared to the number of matching features.

Figure 1.TIF

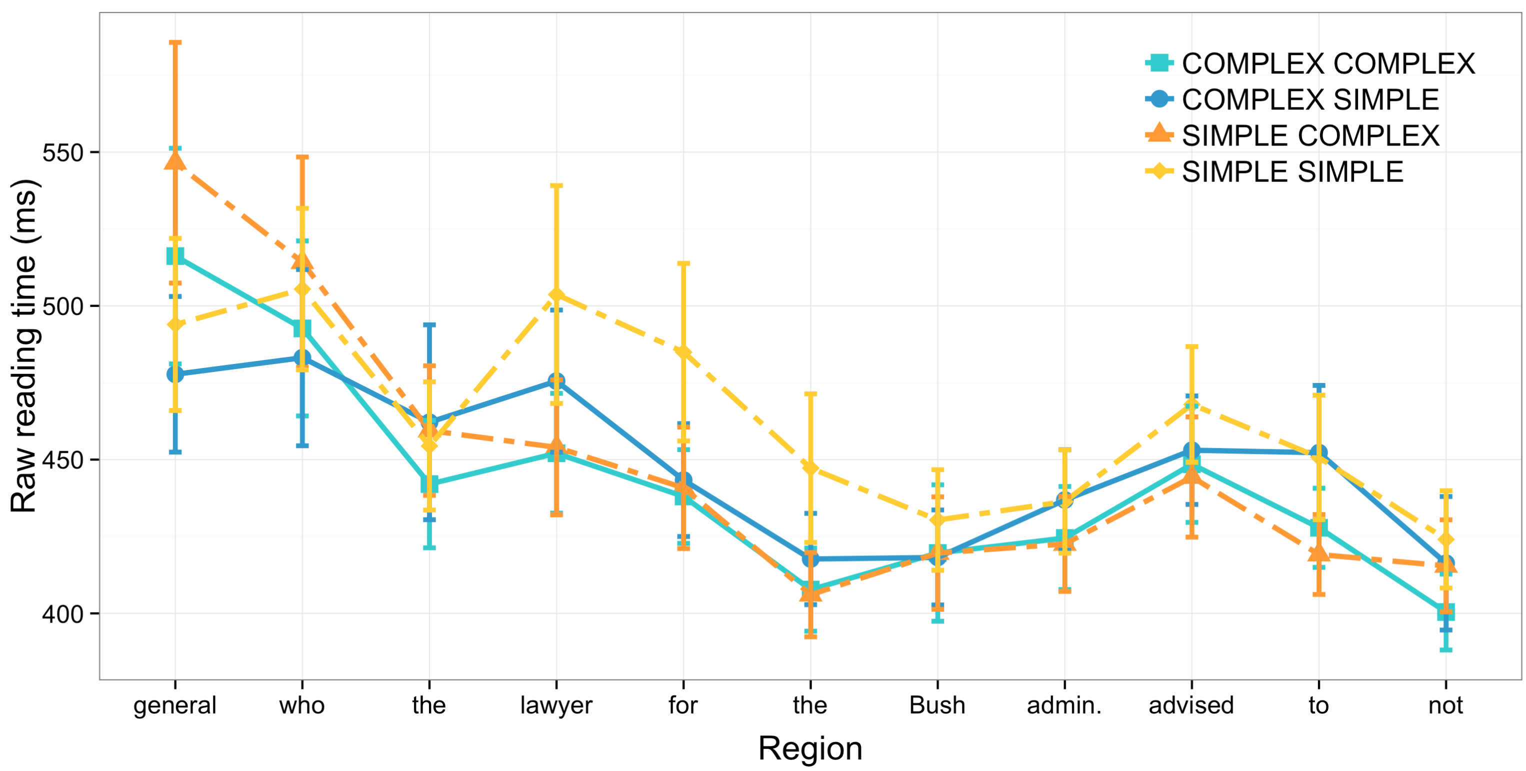

COMPLEX COMPLEX
COMPLEX SIMPLE
SIMPLE COMPLEX
SIMPLE SIMPLE
Raw reading time (ms)
550
500
450
400
general
who
the
lawyer
for
the
Bush
admin.
advised
to
not
Region

Figure 2.TIF

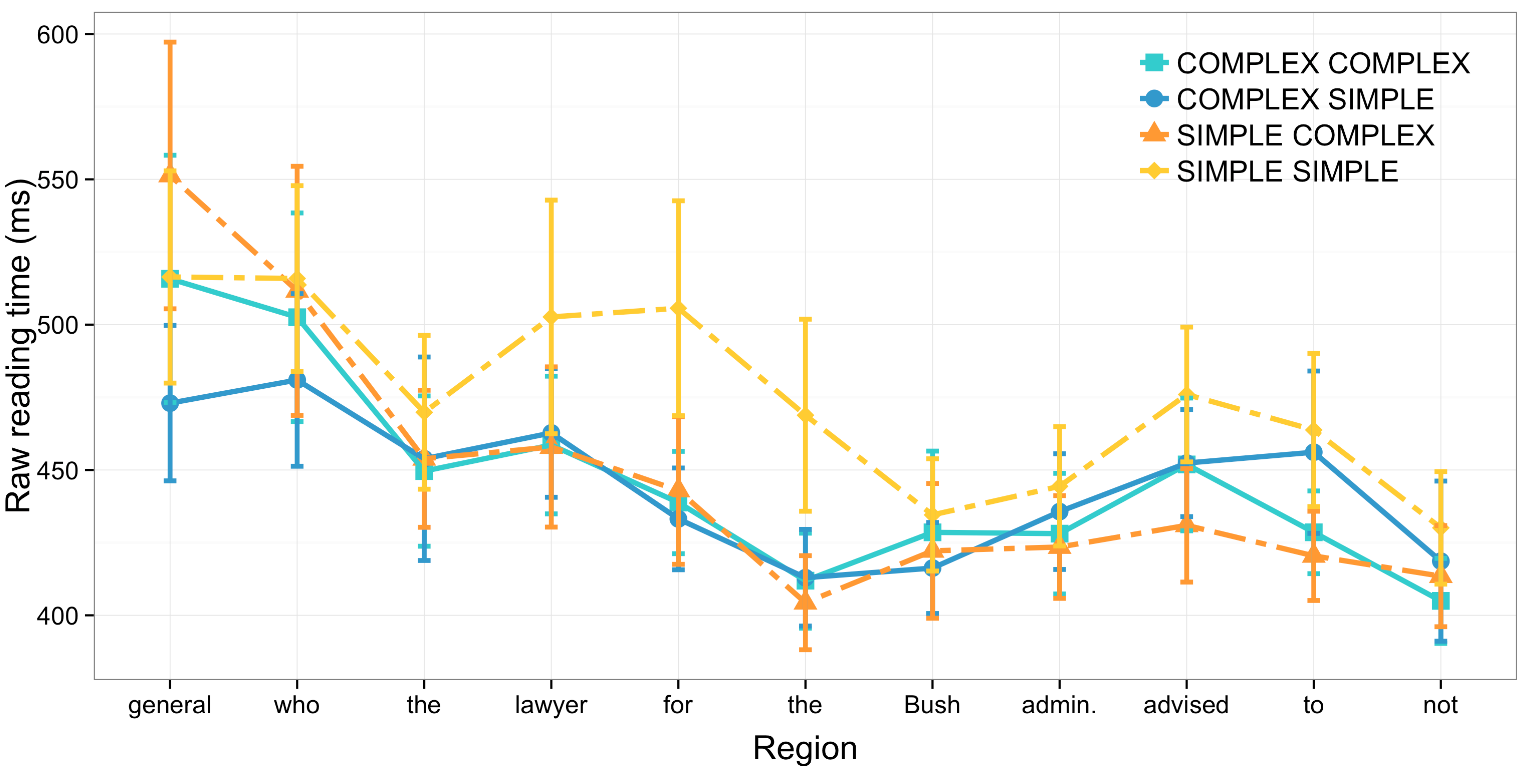
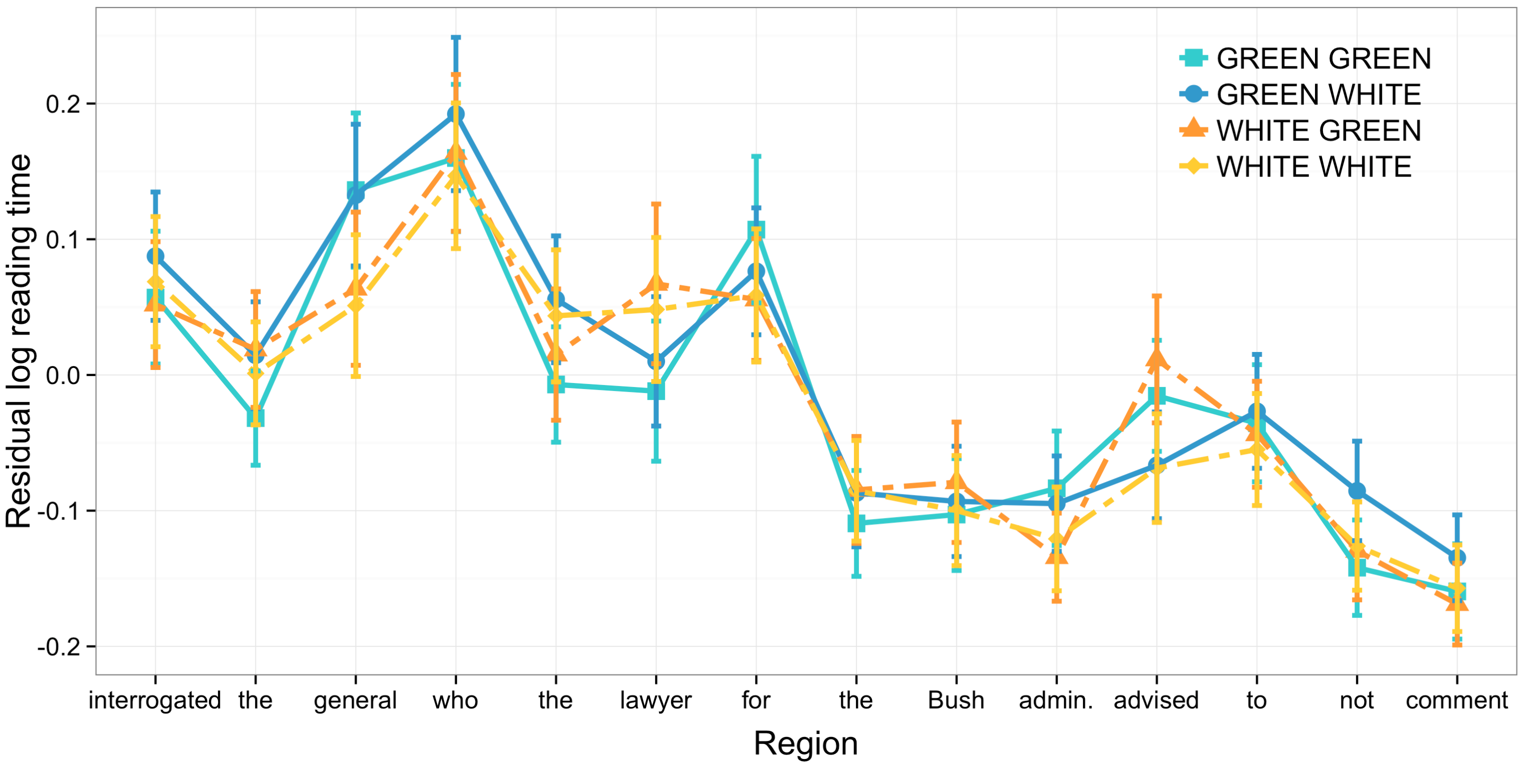
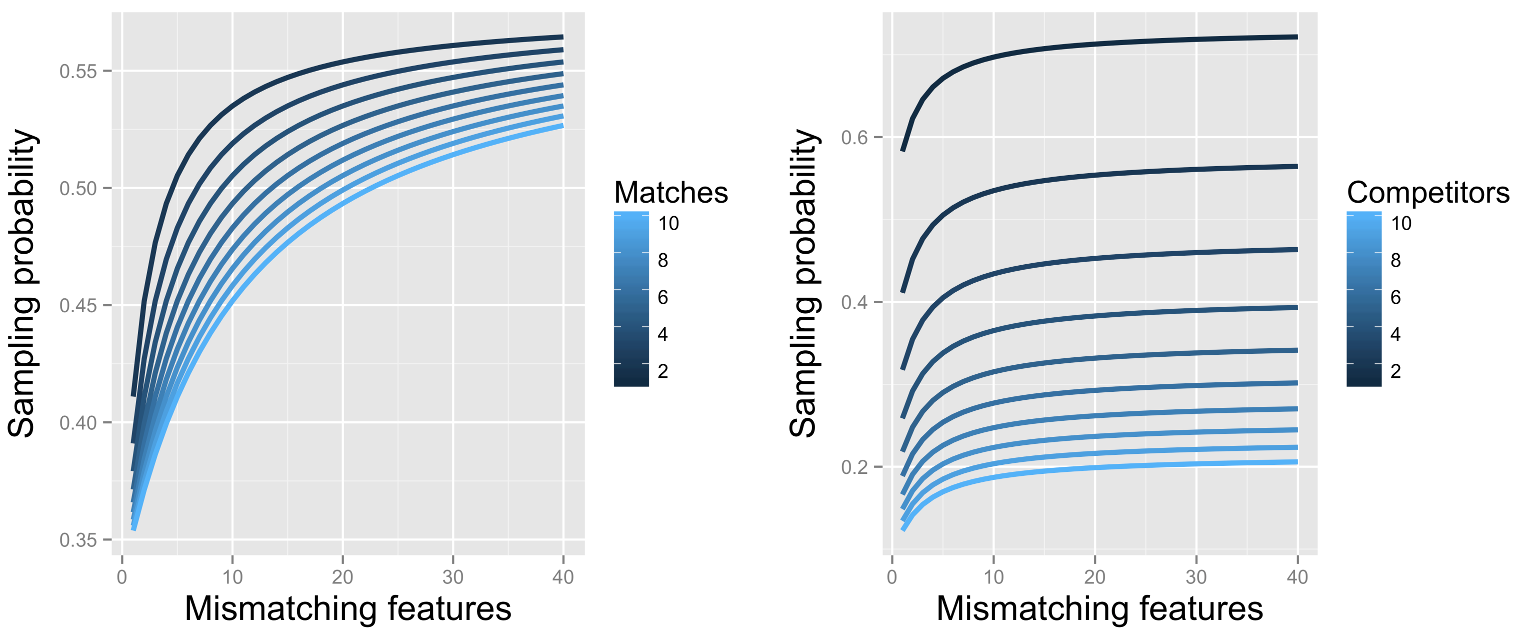

COMPLEX COMPLEX
COMPLEX SIMPLE
SIMPLE COMPLEX
SIMPLE SIMPLE
Raw reading time (ms)
600
550
500
450
400
general
who
the
lawyer
for
the
Bush
admin.
advised
to
not
Region

Figure 3.TIF



GREEN GREEN
GREEN WHITE
WHITE GREEN
WHITE WHITE
Residual log reading time
0.2
0.1
0.0
-0.1
-0.2
interrogated
the
general
who
the
lawyer
for
the
Bush
admin.
advised
to
not
comment
Region

Figure 4.TIF


Sampling probability
0.55
0.50
0.45
0.40
0.35
0
10
20
30
40
Mismatching features
Matches
10
8
6
4
2
Sampling probability
0.6
0.4
0.2
0
10
20
30
40
Mismatching features
Competitors
10
8
6
4
2