2018

The Modulatory Effect of Expectations on Memory Retrieval During Sentence Comprehension

Luca Campanelli
The Graduate Center, CUNY

Julie A. Van Dyke
Haskins Laboratories

Klara Marton
The Graduate Center, CUNY

Recommended Citation
Campanelli, L., Van Dyke, J. A., & Marton, K. (2018). The modulatory effect of expectations on memory retrieval during sentence comprehension. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), Proceedings of the 40th Annual Conference of the Cognitive Science Society (pp. 1436–1441). Austin, Texas: Cognitive Science Society.

The Modulatory Effect of Expectations on Memory Retrieval During Sentence Comprehension

Luca Campanelli[1,2] (campanelli.l@gmail.com), Julie A. Van Dyke[2] (jvandyke@haskins.yale.edu), Klara Marton[1] (kmarton@gc.cuny.edu)

[1]Speech-Language-Hearing Sciences, The Graduate Center, CUNY, New York, NY 10016 USA

[2]Haskins Laboratories, New Haven, CT 06511 USA

Abstract

Memory retrieval and probabilistic expectations are recognized factors in sentence comprehension that capture two different critical aspects of processing difficulty: the cost of retrieving and integrating previously processed elements with the new input words and the cost of incorrect predictions about upcoming words or structures in a sentence. Although these two factors have independently received substantial support from the extant literature, how they interact remains poorly understood. The present study investigated memory retrieval and expectation in a single experiment, pitting these factors against each other. Results showed a significant interference effect in both response time to the comprehension questions and reading time at the last (spillover) sentence region. We also found that the interference effect on reading time (but not on comprehension question response time) was canceled when the word at the retrieval site was highly predictable. Overall, our findings are consistent with the hypothesis of a modulatory effect of expectations on memory retrieval and with the idea that expectation-based facilitation results from pre-activation of the target word ahead of time.

Keywords: working memory; interference; expectations; dual tasking

Introduction

Understanding the nature and the source of processing complexity in human sentence comprehension has been a central goal in linguistics and psycholinguistics. The present work focuses on two classes of explanations: Cue-based retrieval theory (Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006) and expectation theory (Hale, 2001; Levy, 2008). The cue-based retrieval theory defines processing difficulty in terms of memory interference from similar words. The expectation theory characterizes processing difficulty in terms of degree of experience with the input in the past: the less common a word or construction is, the more difficult it will be to process. We selected these two theories because of their substantial empirical support from the extant literature and, most importantly, because they make different theoretical assumptions that lead to aligned, complementary, or opposite predictions depending on the nature of the stimuli and the task at hand.

Consider the examples in (1): both theories correctly predict that (1b) is more difficult to process than (1a). Corpus studies (Roland, Dick, & Elman, 2007) show that in English, subject relative clauses occur more frequently than object relative clauses, and thus comprehenders' greater experience with grammatical structures as in (1a) would account for the difference in ease of processing between (1a) and (1b). From the point of view of cue-based retrieval theory, a key difference between subject and object relative clauses is at the verb in the embedded clause. In both (1a) and (1b) a retrieval operation must take place when the verb *attacked* is encountered; however, whereas in (1a) there is only one plausible subject (*reporter*) for *attacked*, in (1b) two nouns (*reporter* and *senator*), already encoded in memory, need to be correctly associated with subject and object roles, and the similarity between them can lead to memory interference and slower processing.

(1) a. The reporter who attacked the senator admitted the error.
b. The reporter who the senator attacked admitted the error.

If in constructions like (1) both theories correctly account for comprehenders' performance, they make opposite predictions in other cases. Consider the sentences in (2) (Grodner & Gibson, 2005). When the distance between the noun *administrator* and the verb *supervised* is increased, like in (2b) as compared to (2a), the cue-based retrieval theory predicts slower (if there is interference) or comparable (if there is no interference) reading times at the verb *supervised*. This effect is explained in terms of possible greater memory cost (interference) when two elements are more distant. The expectation theory predicts the opposite pattern: faster reading times at the verb *supervised* in (2b) than (2a). The explanation is that the expectation of a verb becomes stronger as more material is interpolated between the initial noun and its verb.

(2) a. The administrator who the nurse supervised ...
b. The administrator who the nurse that was from the clinic supervised ...

Interestingly, neither of these predictions (comparable or slower vs. faster reading time with increased distance) holds up cross-linguistically: the predictions of the cue-based retrieval theory are correct in English and Russian but not, for example, in Hindi (Grodner & Gibson, 2005; Levy, Fedorenko, & Gibson, 2013) and the expectation theory is incorrect for English but correct in other languages such us Hindi and German (Levy & Keller, 2013; Vasishth & Lewis, 2006).

Overall, it appears that the cue-based retrieval theory and the expectation theory capture different aspects of processing difficulty: the former a "backward-looking" cost, that is, the cost of retrieving and integrating previously processed material with the incoming words; the latter a "forward-looking" cost, that is, the cost of updating or dropping predictions that are incompatible with the current word (Demberg & Keller, 2008). This theoretical difference is accompanied by data that cannot be fully explained by either approach. Based on these considerations, many researchers now agree that a model of sentence processing complexity needs to include both memory retrieval and expectation features in order to explain the full range of data available (e.g., Demberg & Keller, 2008; Levy et al., 2013; Staub, 2010; Vasishth & Drenhaus, 2011). Some attempts in this direction have been made in recent years but no integrative model has been formulated and a pressing question remains unanswered: how do memory retrieval and expectation work together? Here, we hypothesize that retrieving and integrating a previously processed word is easier for highly predictable sentences (sharp expectation) as compared to weakly predictable sentences. According to this view, expectation-based facilitation at the retrieval site results from pre-activation of the target word ahead of time. Our prediction for the current study is therefore that precise expectations may diminish the damaging effects of memory interference by boosting the availability of the target element relative to its competitors.

The Present Study

This study examined the modulatory effect of lexical-semantic expectation on memory retrieval. In order to maximize the validity and the comparability of our findings with previous research, this study combined two established research paradigms in a single experiment. 1) The effect of retrieval interference was isolated by using the sentence reading dual-task paradigm developed by Van Dyke and McElree (2006). This paradigm has proven to be robust at identifying retrieval interference effects, as the original finding has now been replicated several times (e.g., Sekerina, Campanelli, & Van Dyke, 2016; Van Dyke, Johns, & Kukona, 2014). Baseline and interference conditions were created by crossing memory load and main clause verb type. Memory load was manipulated such that participants either did or did not maintain in memory a list of three nouns (e.g., *website–handbag–password*) while reading the stimulus sentence. The interference condition was determined by whether the nouns in the memory list were (interference) or were not (no interference) plausible direct objects for the sentence main verb (e.g., *website*, *handbag*, and *password* are plausible direct objects for the verb *created* but not for the verb *performed*).

Table 1: Example of experimental items in the Memory Load condition. Slashes indicate regions of presentation.

Condition	Memory list	Sentence
a. NoInt, LowExp	website-handbag-password	It was the *dance*/that the person/who lived/in the city/performed/early last month.
b. Int, LowExp	website-handbag-password	It was the *dance*/that the person/who lived/in the city/created/early last month.
c. Int, HighExp.	website-handbag-password	It was the *dance*/that the choreographer/who lived/in the city/created/early last month.

2) The effect of expectation was isolated by manipulating the main clause verb's cloze probability, that is, the predictability of the verb in the context[1] of the stimulus sentence (Taylor, 1953). This manipulation has been successfully adopted in many studies examining the effect of expectation on sentence comprehension (e.g., Levy, 2008, 2013). The full design consisted of 6 conditions, with two levels of expectation examined within the Interference conditions, but only low expectation within the No Interference conditions (see Table 1 for examples of Load conditions). The No Load conditions included the same sentences, but without the memory list.

Method

Participants

Participants included 36 young adults (15 females) between 22 and 37 years of age (M = 30.3; SD = 5.4). All participants were native speakers of American English and reported no history of cognitive or language delay. Participants' education level is reported in Table 2.

Materials

The experimental material consisted of object cleft sentences (Table 1). There were 10 trials for each condition for a total of 10 × 6 = 60 experimental trials. In addition, there were 120 filler sentences with different syntactic constructions as in Van Dyke and McElree (2006).

Table 2: Sample education.

Education	N
High School graduate	6
Some college, no degree	6
Associates degree	5
Bachelors degree	13
Graduate degree (Masters, Doctorate, etc.)	6

The experimental material was developed in two stages. First, we created the stimulus sentences for the 60 experimental trials: some of the sentences were adapted from Van Dyke and McElree (2006); the remaining sentences were created following the procedure described in that study. Second, to determine the cloze probability for the expectation manipulation we conducted two independent norming studies using Amazon Mechanical Turk. Cloze probabilities were obtained for the two Low Expectation conditions (a. and b. in Table 1: *It was the dance that the person who lived in the city* ______ *early last month*) and the Interference-High Expectation condition (c. in Table 1: *It was the dance that the choreographer who lived in the city* ______ *early last month*). Following the standard cloze procedure, for each norming study fifty native English speakers were asked to complete the sentences with plausible words. MTurk workers were allowed to participate in only one of the two norming studies. From these responses we selected verbs with probability of 5% or smaller for the low expectation conditions and verbs with probabilities greater than 30% for the high expectation conditions. In addition, the sentence elements that were

[1] The contextual element that was varied in the sentence was the main subject (e.g., *person* vs. *choreographer* in Table 1).

variable across conditions (main clause subject and verb) were matched on average word frequency and word length to the greatest extent possible.

Plausibility of the target object and the memory nouns was verified with an additional norming study using Amazon Mechanical Turk. Fifty MTurk workers were asked to judge on a scale from 1 (not at all) to 7 (very well) how well the target object of the sentence (e.g., *dance*) and the three memory nouns (e.g., *website*, *handbag*, and *password*) fit into the sentence. For this task, workers were presented with a simplified version of the experimental sentences in the following form: *The person created the _website_*. For plausible objects, we retained only words with average score greater than 5. For implausible objects, only words with average score smaller than 2 were kept.

Procedure

Six lists were created such that each item occurred only once in a list, following a Latin-Square design. The lists were counterbalanced by subject and presented using Inquisit Web 4 (Inquisit 4, 2015), which provides millisecond precision for stimulus presentation and response time. In the Memory Load condition, first, participants saw the words of the memory list for 3 seconds and were instructed to memorize them; then they read the sentence using the self-paced phrase-by-phrase methods (for details, see Van Dyke & McElree, 2006). At the end of the sentence, participants answered a comprehension question (true/false) by pressing "1" or "3" on the keyboard and then they typed the words of the memory list into the computer. In the No Memory Load condition, participants were only required to read the sentence and answer the comprehension question. Half of the comprehension questions were true and half were false. For the experimental sentences, correct answers required understanding the main clause subject-verb-object relation (e.g., for b. in Table 1, *Did the person create the dance? / Did the person create the password?*); for the filler sentences, the comprehension questions probed different aspects of the sentences.

Participants for the main experiment were recruited through MTurk. Only workers who consistently demonstrated a high degree of reliability in performing a wide range of tasks across a large number of requesters were allowed to participate. Requirements included Master qualification[2], approval rate greater than or equal to 99%, number of HITs approved greater than or equal to 10000, US location, and screen resolution greater than or equal to 800×600 pixels. Workers were invited to participate in the study only if they were in a quiet room and had the time to complete the task in one sitting. Forty-five workers accepted the HIT but only the 36 subjects described in the Participants section completed the task and were included in the present study.

The study required an average of 45 minutes to complete, and participants received $8 plus a bonus depending on their performance accuracy: $0.02 for each comprehension question answered correctly (total for 180 sentences = $3.6 maximum) and $0.02 for each single memory word recalled correctly (82 lists of 3 words = 246 = $4.92 maximum).

All procedures and materials were approved by the Institutional Review Board of the City University of New York.

Data Analysis

Mixed-effects logistic regression analysis was used to examine comprehension question accuracy, and linear mixed-effects regression was employed to analyze comprehension question response time and reading time. Both response times and reading times were log transformed before the analyses. All models included random intercepts for subjects and items and by-subject[3] random slopes (Baayen, Davidson, & Bates, 2008). As the three experimental manipulations were not completely crossed, a single factor with six levels was used as the predictor variable; then pre-planned contrasts were used to test our hypotheses: the effect of Load (Load conditions vs. NoLoad conditions), the effects of Interference and Expectation in the NoLoad conditions, and the effects of Interference and Expectation in the Load conditions.

The presence of outliers was examined by looking at average performance by subject and condition and models residuals. Absolute standardized values greater than 2.5 were trimmed, with exclusion of less than 4% of the data.

Data were analyzed with R version 3.4.3 (R Core Team, 2017) using the `glmer` and `lmer` functions from the **lme4** package, version 1.1-15 (Bates, Mächler, Bolker, & Walker, 2015). Contrasts for the effects of Interference and Expectation were carried out using the **emmeans** package (Lenth, Love, & Herve, 2018). The effect of Load was tested using the `linearHypothesis` function from the **car** package (Fox & Weisberg, 2011).

Results

Recall of Memory List (Load conditions only)

Recall of the memory words was scored according to a lenient criterion, in which accuracy was determined by the number of words correctly recalled, regardless of their order of presentation. Average recall was 93.2% and comparable in the three conditions, therefore no statistical analysis was performed. Descriptive statistics are reported in Table 3.

Table 3: Descriptive statistics for recall accuracy (percent).

Condition	Mean	SD	Min	Max
NoInt-LowExp	92.8	11.1	60	100
Int-LowExp	93.6	11.2	53	100
Int-HighExp	93.2	11.3	57	100

Comprehension Questions Accuracy

Response accuracy to the comprehension questions was above 97% in the NoLoad conditions (NoInt-LowExp = 98%; Int-LowExp = 97%; Int-HighExp = 97%) and higher than 90% in the three Load conditions (NoInt-LowExp =

[2] For more information about Master qualifications in Amazon Mechanical Turk, see https://www.mturk.com.

[3] By-item random slopes were tested but never retained because of convergence failures.

95%; Int-LowExp = 91%; Int-HighExp = 92%). Contrasts following mixed-effects logistic regression analysis confirmed a statistically significant effect of Load, ($\chi^2(1) = 23.9$, $p < .001$). None of the contrasts for the effects of Interference and Expectation in the NoLoad and Load conditions reached statistical significance ($p > .073$).

Comprehension Questions Response Time

The analysis of response time to the comprehension questions revealed a significant effect of Load ($\chi^2(1) = 15.1$, $p < .001$), such that participants' reaction times were overall slower in the Load than in the NoLoad conditions (1794 ms and 1661 ms, respectively). It also emerged a significant Interference effects in the Load conditions, indicating that the two interference conditions (Int-LowExp and Int-HighExp) were slower than the no interference condition. No other significant effects emerged (Table 4).

Table 4: Comprehension question response time: Contrasts for the effects of Interference and Expectation.

Contrast[a]	Estimate	95% CI	*t* ratio	*p*-value
1–2	-0.04	[-0.07; 0.00]	-1.78	.076
1–3	-0.03	[-0.07; 0.01]	-1.46	.144
2–3	0.01	[-0.03; 0.05]	0.32	.752
4–5	-0.04	[-0.08; 0.00]	-2.14	.032
4–6	-0.07	[-0.11; -0.03]	-3.52	<.001
5–6	-0.03	[-0.07; 0.01]	-1.34	.18

Note. [a]1 = NoLoad-NoInt-LowExp; 2 = NoLoad-Int-LowExp; 3 = NoLoad-Int-HighExp; 4 = Load-NoInt-LowExp; 5 = Load-Int-LowExp; 6 = Load-Int-HighExp.

Reading Time

To adjust for between subject variability in reading time, region length, and region position, residual log transformed reading time was used as the dependent measure (Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994).

For the critical sentence region (retrieval verb, e.g., *performed*), no significant effects of Load, Interference, or Expectation were found ($p > .18$). Therefore, these results will not be discussed further.

Table 5: Reading time at the spillover region: Contrasts for the effects of Interference and Expectation.

Contrast[a]	Estimate	95% CI	*t* ratio	*p*-value
1-2	-0.01	[-0.03; 0.02]	-0.53	.599
1-3	0.05	[0.02; 0.07]	4.01	<.001
2-3	0.06	[0.03; 0.08]	4.49	<.001
4-5	-0.03	[-0.05; 0]	-1.99	.054
4-6	0.02	[0; 0.05]	1.71	.096
5-6	0.05	[0.02; 0.07]	3.47	.001

Note. [a]1 = NoLoad-NoInt-LowExp; 2 = NoLoad-Int-LowExp; 3 = NoLoad-Int-HighExp; 4 = Load-NoInt-LowExp; 5 = Load-Int-LowExp; 6 = Load-Int-HighExp.

A different pattern of results was found at the spillover region (e.g., *early last month*). Similarly to Van Dyke and McElree (2006), reading time in the Load conditions was faster than that in the NoLoad conditions ($\chi^2(1) = 5.6$, $p = .018$). For the NoLoad conditions, it emerged an Expectation, but not Interference, effect, indicating that reading time was faster in the HighExp condition than in the two LowExp conditions. Most importantly, we found significant, or approaching significance, effects of both Expectation and Interference in the Load conditions, such that reading was slowest in the Int-LowExp condition and fastest in the Int-HighExp condition (Table 5; Figure 1).

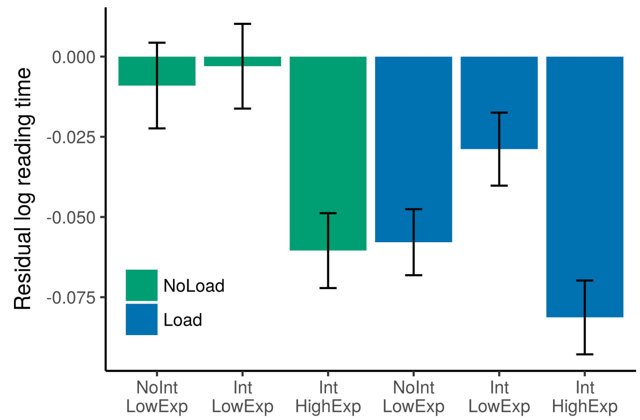

Figure 1: Reading time at the spillover region (±SE). More negative values indicate faster reading times.

Discussion

In this paper, we examined the effect of lexical-semantic expectation on memory retrieval during self-paced reading of object cleft sentences. The aim of the present study was twofold. First, although there is substantial empirical support for the independent contribution of memory and expectation in sentence comprehension, much less is known about how they interact during online processing. The present experiment is one of the first to examine their interaction by combining two established research paradigms in a single experiment.

Second, the majority of the studies that examined the independent effects of memory and expectation operationalized memory cost using outdated models of memory, such as variations of the *capacity* theory, that lead to memory cost measures based on some sort of distance between dependent constituents (e.g., number of words or intervening discourse referents; Gibson, 2000; Just & Carpenter, 1992; Miller & Chomsky, 1963). Current memory models in experimental psychology and psycholinguistics have shifted their focus from global *capacity* to the specific content of the memory representations and whether the target representations can be reliably retrieved from memory (Van Dyke & Johns, 2012; Van Dyke, Johns, & Kukona, 2014; Van Dyke & Lewis, 2003). All previously processed items that are stored in working memory compete during retrieval, and processing difficulty is related to the reliability of the retrieval cues for discriminating targets from distracting information. On this view, it is retrieval interference from similar items, not the amount of information, that is the primary source of processing difficulty (Fedorenko, Gibson, & Rohde, 2006; Fedorenko, Woodbury, & Gibson, 2013;

Gordon, Hendrick, & Johnson, 2001; Gordon, Hendrick, & Levine, 2002; Van Dyke, 2007; Van Dyke & McElree, 2006, 2011). The present experiment advances our understanding of the relation between memory and expectation by operationalizing memory cost in terms of retrieval interference.

Two key findings emerged. First, reading time at the spillover region replicated established memory interference and expectation effects, thus supporting the validity of the method adopted. Second, we found that the interference effect on reading time (but not on response accuracy or response time to the comprehension questions) was completely canceled when the word at the retrieval site was highly predictable.

These results are in line with other, indirect, evidence that points to a modulatory effect of expectation on memory retrieval. In the experimental psychology and brain research literature, for example, recent work has shown that expectations of upcoming events improve speed and accuracy of stimulus detection, discrimination, and retrieval from memory (e.g., Bollinger et al., 2010; Esterman & Yantis, 2010; Gazzaley & Nobre, 2012; Summerfield & Egner, 2009). In the psycholinguistics literature, it is worth mentioning, among others, the study by Husain, Vasishth, and Srinivasan (2014, see also Nicenboim, Vasishth, Gattei, Sigman, & Kliegl, 2015). These authors found that locality effects (increased difficulty with increasing distance between elements) were detected when expectation was weak but not when expectation was strong.

Overall, this pattern of results further stresses the need for a unified framework for memory, expectation, and their interaction, and seems not to be consistent with additive or two-factor type of models of language processing. Levy (2008), for example, proposed a model in which expectation has an effect on early stages (e.g., lexical processing) and memory interference on a later, syntactic integration phase. Although there is evidence supporting such two-factor models, they do not make any explicit predictions about possible interactions between memory and expectation components.

A model that explicitly predicts an interaction between expectation and memory interference is proposed by Lau (2009). Lau hypothesizes that a strong enough expectation may trigger the attachment of the expected element in advance of its encounter. This is followed by a quick check of the bottom-up input against the prediction, without the need for a retrieval operation. However, the mechanisms by which the predicted elements are maintained active in memory are not clearly specified.

An interaction between expectation and memory interference could also be accommodated by left corner parsers, such as the ACT-R implementation in Lewis & Vasishth (2005). A transient shift of attention to the cued representation guided by contextual information can increase the baseline activation of the target word and facilitate subsequent memory retrieval.

Overall, our findings are consistent with the hypothesis that as the relevance of a specific element increases in a given sentential context, the cued information is (probabilistically) pre-activated. If we think of memory retrieval as a gradual accumulation of information in the focus of attention, expectation would exercise its effect via an advance accumulation of evidence before the retrieval is initiated. Such a head start for selection of information would reduce retrieval interference by boosting the availability of the target word relative to its competitors. The pre-activation view of expectation-based facilitation is not new in psycholinguistics, and received substantial support in the last decades (e.g., Federmeier, 2007; Kuperberg & Jaeger, 2016; Pickering & Gambi, in press, for discussion). This view is also compatible with recent neural models that describe the brain as a predictive machine, in which the neural system is assumed to constantly be predicting upcoming input and monitoring the degree of match between anticipated information and perceptual input (e.g., Clark, 2013; Schacter, Addis, & Buckner, 2007). Future research will need to further investigate the precise mechanisms by which expectations exercise their effects and timing of their interactions with memory retrieval.

Acknowledgments

Thanks to Irina Sekerina and Ashley Lewis for comments and helpful discussions. This material is based upon work supported by the National Science Foundation under Grant No. BCS-1625083 and the National Institutes of Health, Grant No. R01 HD-073288.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*.

Bollinger, J., Rubens, M. T., Zanto, T. P., & Gazzaley, A. (2010). Expectation-driven changes in cortical functional connectivity influence working memory and long-term memory performance. *Journal of Neuroscience, 30*, 14399–14410.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*, 181–204.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition, 109*, 193–210.

Esterman, M., & Yantis, S. (2010). Perceptual expectation evokes category-selective cortical activity. *Cerebral Cortex, 20*, 1245–1253.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*, 491–505.

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language, 54*, 541–553.

Fedorenko, E., Woodbury, R., & Gibson, E. (2013). Direct evidence of memory retrieval as a source of difficulty in non-local dependencies in language. *Cognitive Science, 37*, 378–394.

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

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (2nd ed.). Thousand Oaks, CA: Sage.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences, 16*, 129–135.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT.

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

Grodner, D. J., & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentenial complexity. *Cognitive Science, 29*, 261–291.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the 2nd meeting of the North American Chapter of the Association for Computational Linguistics (NAACL)* (Vol. 2, pp. 1–8). Stroudsburg, PA: ACL.

Husain, S., Vasishth, S., & Srinivasan, N. (2014). Strong expectations cancel locality effects: Evidence from Hindi. *PLoS ONE, 9*, e100986.

Inquisit 4. (2015). Retrieved from www.millisecond.com

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

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31*, 32–59.

Lau, E. F. (2009). *The predictive nature of language comprehension* (Doctoral dissertation). University of Maryland, College Park, Maryland.

Lenth, R., Love, J., & Herve, M. (2018). emmeans: Estimated marginal means, aka least-squares means (Version 1.1).

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*, 1126–1177.

Levy, R. (2013). Memory and surprisal in human sentence comprehension. In R. P. G. van Gompel (Ed.), *Sentence processing* (pp. 78–114). New York: Psychology Press.

Levy, R., Fedorenko, E., & Gibson, E. (2013). The syntactic complexity of Russian relative clauses. *Journal of Memory and Language, 69*, 461–495.

Levy, R., & Keller, F. (2013). Expectation and locality effects in German verb-final structures. *Journal of Memory and Language, 68*, 199–222.

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

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

Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 419–491). New York: Wiley.

Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing information during working memory: Beyond sustained internal attention. *Trends in Cognitive Sciences, 21*, 449–461.

Nicenboim, B., Vasishth, S., Gattei, C., Sigman, M., & Kliegl, R. (2015). Working memory differences in long-distance dependency resolution. *Frontiers in Psychology, 6*.

Pickering, M. J., & Gambi, C. (in press). Predicting while comprehending language: A theory and review. *Psychological Bulletin*.

R Core Team. (2017). R: A language and environment for statistical computing (Version 3.4.3). Vienna, Austria: R Foundation for Statistical Computing.

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language, 57*, 348–379.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience, 8*, 657–661.

Sekerina, I. A., Campanelli, L., & Van Dyke, J. A. (2016). Using the visual world paradigm to study retrieval interference in spoken language comprehension. *Frontiers in Psychology, 7*, 873.

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception, & Psychophysics, 78*, 1839–1860.

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition, 116*, 71–86.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences, 13*, 403–409.

Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. Journalism Quarterly, 30, 415–433.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*, 285–318.

Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 407–430.

Van Dyke, J. A., & Johns, C. L. (2012). Memory interference as a determinant of language comprehension. *Language and Linguistics Compass, 6*, 193–211.

Van Dyke, J. A., Johns, C. L., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition, 131*, 373–403.

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

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

Vasishth, S., & Drenhaus, H. (2011). Locality in German. *Dialogue & Discourse, 2*, 59–82.

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