


Integration and Prediction in Language Processing: A Synthesis of Old and New

Fernanda Ferreira and Suphasiree Chantavarin

Department of Psychology, University of California, Davis

Current Directions in Psychological Science
2018, Vol. 27(6) 443–448
© The Author(s) 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0963721418794491
www.psychologicalscience.org/CDPS


Abstract

Current theories of language processing emphasize prediction as a mechanism to facilitate comprehension, which contrasts with the state of the field a few decades ago, when prediction was rarely mentioned. We argue that the field of psycholinguistics would benefit from revisiting these earlier theories of comprehension that attempted to explain integration and the processes underlying the formation of rich representations of linguistic input and that emphasized informational newness over redundancy. We suggest further that integration and anticipation may be complementary mechanisms that operate on elaborated, coherent discourse representations, supporting enhanced comprehension and memory. In addition, the traditional emphasis on language as a tool for communication implies that much linguistic content will be nonredundant; moreover, the purpose of anticipation is probably not to permit the prediction of exact lexical or syntactic forms but instead to induce a state of preparedness that allows the comprehender to be receptive to new information, thus facilitating its processing.

Keywords

language processing, inferences, prediction, integration

Current psycholinguistic theories emphasize the importance of prediction for efficient language comprehension (Kuperberg & Jaeger, 2016). This trend dovetails with the broader treatment of cognitive agents as “prediction engines” that anticipate the future and update knowledge databases when their predictions are disconfirmed (Bansal, Ford, & Sperling, 2018; Clark, 2013). Today’s focus on prediction in language processing stands in stark contrast to models of comprehension and integration processes from a few decades ago, which essentially ignored prediction. Instead, the fundamental goal of those earlier models was to explain how people link incoming linguistic information to previous input and to background knowledge to generate rich, elaborated representations of texts and conversations (Long & Lea, 2005). The possibility that linguistic content might be predicted was occasionally considered, but discussions centered around ideas such as *forward inferencing*—a comprehender’s anticipation of characters, consequences, or other concepts likely to feature later in the text (Graesser, Singer, & Trabasso, 1994)—and evidence for forward inferences was minimal.

Today, psycholinguists are clearly taken with prediction, and lexical prediction in particular. Why is prediction so popular now? One reason is that prediction of

words fits well with top-down models of comprehension (Lupyan & Clark, 2015) because it suggests that meaning can get ahead of the input. Perhaps even more influential are Bayesian approaches to cognition (Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010), which assume that decision making is based on a rational weighting of the agent’s priors (previous knowledge) and current evidence (the input). Language processing is thought to reflect the way that knowledge and contextual constraint combine to make a particular linguistic form likely, leading to prediction.

Integration Versus Prediction, and Redundancy Versus Information

When we compare the current literature with work from a few decades ago, we see two radically different pictures. In the older tradition, the primary mechanism supporting language comprehension was thought to be integration—the linking of new ideas and concepts to

Corresponding Author:

Fernanda Ferreira, University of California, Davis, Department of Psychology, 1 Shields Ave., Davis, CA 95616
E-mail: fferreira@ucdavis.edu

what is already known or established (Gernsbacher, 1991; Kintsch & van Dijk, 1978). A key background assumption was that the purpose of language is to communicate; if we say something, it is because we possess information that we wish to convey to someone who lacks that knowledge. To make that task easier for the comprehender, researchers argued that speakers organize their sentences so that early content indexes what is under discussion, and the rest of the sentence adds to what is already known. As the input is processed, the comprehender's representation is continuously updated with new information, with the representation growing progressively richer and more elaborate. The more coherent and organized the representation, the deeper the comprehension, and the stronger the comprehender's memory for the linguistic episode.

In contrast, in current approaches, it is assumed that the primary mechanism supporting language comprehension is prediction (Dell & Chang, 2013). Processing is viewed as fundamentally Bayesian, with comprehenders interpreting current input in light of their expectations, leading them to predict that certain words are imminent. A highly influential approach to expectation-based processing is surprisal theory (Hale, 2016; Levy, 2008; Smith & Levy, 2013), which assumes continuous word prediction across the entire range of probability values, from near zero to complete certainty. In support of this approach, eye-tracking studies have shown that predictable words are read faster than less predictable words (Rayner, Ashby, Pollatsek, & Reichle, 2004), and neural activity in left inferior gyrus and anterior superior temporal lobe correlates with surprisal values (Henderson, Choi, Lowder, & Ferreira, 2016). In addition, the electrophysiological response known as the N400, thought to reflect ease of semantic retrieval and integration, is reduced with greater contextual predictability (DeLong, Urbach, & Kutas, 2005). Related work shows that listeners examining a visual world while hearing sentences make anticipatory saccades to objects representing predictable upcoming words (Altmann & Kamide, 1999), and they anticipate corrections to speech errors (Lowder & Ferreira, 2016). Perhaps most striking are studies that report a larger N400 in response to determiners containing grammatical information incompatible with the features of a predicted upcoming noun, which implies genuine prediction because the determiner has little semantic content (Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). Although some investigators have recently expressed skepticism about prediction and have challenged some of this evidence (Huettig, 2015; Nieuwland et al., 2018; Staub, 2015), the notion that some aspects of meaning are preactivated, at least under some conditions, seems now well established in psycholinguistics.

The question we address here is whether the older tradition can be reconciled with the current emphasis on prediction. Our goal is actually more ambitious: We argue that these approaches should be unified to allow the field to move beyond repeated demonstrations of word prediction and toward the creation of deeper and more explanatory theories of language comprehension. To do so, we will defend two claims about language: The first claim is that integration underlies prediction, and the other claim is that language is not as redundant as some prediction-based approaches assume (e.g., Pickering & Garrod, 2004).

Comprehension and Newness in Earlier Models

As mentioned, the older tradition emphasized comprehension, with experiments designed to examine how people connect linguistic forms to each other and relate them to knowledge in long-term memory. For example, one highly influential theory (Kintsch, 1988) assumed that an incoming sentence was parsed into propositions that were stored in working memory and that, in turn, activated long-term memory representations, enabling the comprehender to establish important linguistic relations, such as coreference, as well as key semantic relations, such as causality. The critical properties of the representation that was built over successive processing cycles were coherence and semantic richness: All the linguistic bits had to hang together, and the explicit linguistic content needed to be combined with knowledge from long-term memory to form an elaborated discourse model of the text or conversation.

If the current sentence could not immediately be linked to the previous content, the comprehender was assumed to search back through the discourse representation to resolve the incoherence by identifying a relevant connection between the current sentence and the ones that came before. For example, to understand the sequence "Sam threw the report in the fire; ashes flew up the chimney," it is necessary to infer that the report burned (Singer & Ferreira, 1983). This link was referred to as a *backward* or *bridging inference*—backward because the current sentence triggers an attempt to retrieve the linguistic material that came before, to allow the two to be integrated. Backward inferences were contrasted with what were termed forward inferences, or what today many would call predictions, which were anticipations of an idea being mentioned in an upcoming sentence. For example, when encountering "Sam threw the report in the fire," the comprehender would infer or predict that the report burned as a consequence and, armed with that constructed proposition, would interpret later text.

In contrast to the current enthusiasm for prediction, in the older tradition there was widespread skepticism about the extent of forward inferencing during comprehension, motivating models such as *minimal inferencing* (McKoon & Ratcliff, 1992). Although it seemed plausible that forward inferences might be drawn during normal comprehension, evidence for that assumption was scant, leading to the conclusion that people make only backward connections, triggered by discourse incoherence. To explain this reluctance to predict, some researchers invoked the frame problem from artificial intelligence (Minsky, 1975), noting that it is in principle impossible to restrict the potentially relevant inferences one might draw from any statement. Evidence for forward inferencing was occasionally reported (Murray, Klin, & Myers, 1993), but comprehenders had to be highly motivated and had to make the formation of such inferences a deliberate goal of comprehension. Indeed, task demands were shown to affect even integration processes, highlighting the importance of motivation for all aspects of comprehension (Foertsch & Gernsbacher, 1994). This contrasts with the contemporary view, which treats prediction as normal and routine (although see Brothers, Swaab, & Traxler, 2017, for evidence that the task influences the extent to which readers engage in prediction).

Researchers in this era were also rather stringent about the standards for evidence of prediction. Alternative explanations had to be ruled out, beginning with the most obvious: simple integration. Merely showing that a highly contextually constrained word or phrase was read faster than one that was less constrained was considered insufficient to support a claim of prediction, because the form might simply be more easily processed once encountered. A result that could be attributed to priming (i.e., as the output of a “dumb” or automatic processing route; Huettig, 2015) was also not viewed as evidence for prediction, because spreading activation within a lexical or conceptual structure is a passive mechanism that operates within a single stage or level of the comprehension system (Duffy, Henderson, & Morris, 1989), and genuine prediction was understood to be an active, constructive, top-down process. Conscious prediction was also dismissed as uninteresting, because presumably any intelligent person can complete the sequence “Bill put on his socks and . . .”; indeed, this is what off-line cloze tasks measure, and cloze data set the stage for experimental investigations of prediction but are not themselves evidence for it. Thus, a prediction effect had to be one that (a) could not be attributed to integration or associative priming (i.e., prediction must be a dynamic, generative process that occurs before and not after receiving the input) and (b) could not be explained as arising from

an atypical situation that might unnaturally encourage conscious prediction (e.g., an experiment in which words are presented at unusually slow rates).

Earlier models of comprehension also described the importance of the *given-new* strategy, according to which speakers order the content of their sentences to facilitate the integration of new ideas (Haviland & Clark, 1974). Speakers generally place ideas already on common ground in early syntactic positions (typically, the subject of the sentence), allowing comprehenders to retrieve the topic under discussion, and speakers then reserve the remainder of the sentence for new information. The placement of given information before new information was believed to benefit comprehenders because it allows them to locate the relevant entity before attributing previously unknown properties or features to it. But if new information tends to occur later in sentences, it cannot also be true that redundancy and predictability increase across sentence positions and are greatest at or near the end.

Reconciling the Two Traditions: Preparedness, Not Prediction

The two approaches we have outlined here are superficially rather different and may even appear incompatible, but we believe not only that they are reconcilable but also that a new, hybrid approach has the potential to inspire exciting new research on language comprehension. The structure of our argument is shown in Figure 1.

Starting from the bottom, we assume the division of a sentence into given and new information. Information that is given makes contact with linguistic material that came before, as well as with background knowledge, and integration of the input with preceding context and knowledge leads to the creation of a rich semantic representation. At the same time, the enriched representation is used to anticipate upcoming content, which will typically be a set of semantic features, and sometimes (probably rarely; Luke & Christianson, 2016) those features might be sufficient to identify a single lexical candidate. The critical innovation here is to replace the term *prediction* with the concept of *preparedness*: A comprehender who has formed a rich representation of preceding discourse will be in a state of preparedness that leads to greater receptivity to certain semantic (and potentially also syntactic and phonological) features over others, which may take the form of preactivation.¹ This type of anticipation is less like knowing that the word “library” will probably occur next in a sequence such as “He returned the books to the school. . . ” and more like the state that a barista might be in after asking a customer for his or her order

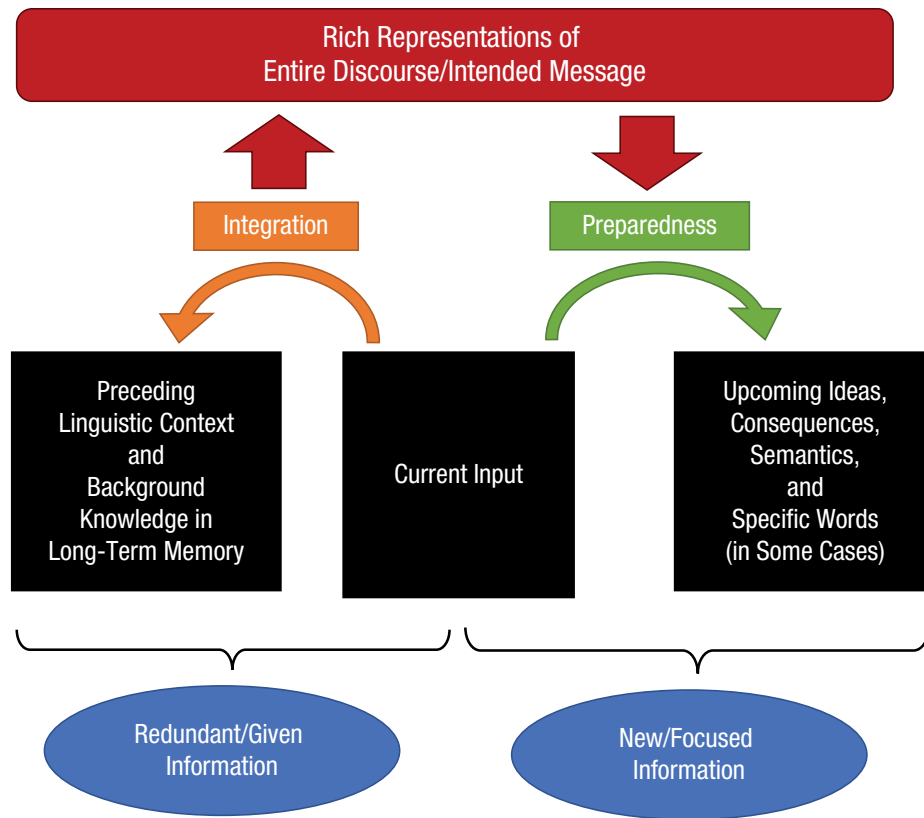


Fig. 1. Our approach linking integration and prediction, or *preparedness*, and reconciling the older and more recent theoretical traditions. We assume the division of a sentence into given and new information. Information that is given makes contact with linguistic material that came before, as well as with background knowledge, and integration of the input with preceding context and knowledge leads to the creation of a rich semantic representation. At the same time, the enriched representation is used to anticipate upcoming content, which will typically be a set of semantic features, and sometimes those features might be sufficient to identify a single lexical candidate.

or the state of a tennis player waiting to return an opponent's serve. Preparedness is similar to the concept of prediction in that it is a forward-looking process that occurs before the linguistic input is received. However, unlike prediction, preparedness emphasizes the prediction of upcoming new information rather than given information, and preparedness is contingent on the formation of a rich discourse representation.

One reason that we believe it is preferable to think in terms of preparedness rather than prediction is that although demonstrations of word-level prediction may be useful for illuminating the architecture of the language-processing system, word prediction is probably rare in real life (Luke & Christianson, 2016). Demonstrations of high predictability emerge from study designs in which sentences with highly predictable final words are manufactured and normed specifically to test the experimental hypotheses. This is not to imply that anticipation is useless, however, because we believe anticipation prepares the comprehender to assimilate

new information (Ferreira & Lowder, 2016). For example, consider a sequence such as "A: What did you eat for lunch? B: I had a" The rest of B's response is presumably not known or else A would not have asked the question, but A can use world knowledge and perhaps also the information in the discourse up to that point to anticipate a semantic category or some features of the likely answer, which will smooth the incorporation of the new information when it arrives. In this view, precise lexical prediction is not common, but preparedness for relevant semantic categories and features is part of normal language processing. Moreover, the idea of preparedness meshes nicely with the motivation and task effects discussed earlier (Brothers et al., 2017; Foertsch & Gernsbacher, 1994), because preparedness likely reflects arousal, which is closely related to motivation. This approach, then, allows us to retain the basic insight behind the given-new strategy—that if specific words in sentences are predictable at all, they are those occurring in initial sentence positions—while allowing

anticipation to be useful, because preactivation of relevant semantic features prepares the comprehender to integrate new information.

Overall, then, this approach reconciles the two approaches we have identified here: As emphasized in the older tradition, accessing of preceding linguistic context and background knowledge facilitates integration of new input, but as newer approaches recognize, a rich representation enables the comprehender to prepare to receive new information, which will lead to precise predictions in (rare) cases of strong constraint. Our proposal treats integration and anticipation not as alternative, competing mechanisms but as essentially two sides of the same coin (Kuperberg & Jaeger, 2016). The older and current approaches both acknowledge that comprehension involves the continuous updating of discourse representations, whether through bottom-up integration or the use of top-down expectations. But our approach emphasizes integration and its role in building rich, elaborated representations of already processed input, which support mechanisms that enable the comprehender to prepare to receive new information.

Questions for Future Research

This synthesis of the two traditions and research literatures inspires many new research questions. Here, we identified just three. The first relates to our concept of preparedness and how it differs from prediction involving the retrieval of prestored information given biasing context (e.g., predicting “pepper” following “Please pass the salt and . . .”). Are lexical prediction and preparedness fundamentally different, or is the retrieval of an associate such as “pepper” given “salt” different only in degree from what is involved in anticipating what a server will list as this evening’s dinner specials? Second, we have assumed that anticipation and integration are complementary mechanisms, but do they take place in parallel during comprehension, or does comprehension proceed in phases that alternate episodes of integration and anticipation? And finally, we can ask what amount of predictability in texts is appealing to listeners and readers. Predictability has been shown to facilitate processing, but if a conversation or story is highly predictable, it will be boring. At the same time, language that is wildly unpredictable is also problematic because it is hard to understand, which comprehenders also find disagreeable. Is there a “sweet spot” of prediction, then? What is the right amount of predictability for language to be interesting and not overly challenging, and can this be related to individual differences in knowledge and cognitive ability? We believe these and related questions are exciting avenues of research for moving the field of psycholinguistics forward, taking the field

beyond demonstrations that predictable words are easy to process.

Recommended Reading

- Ferreira, F., & Lowder, M. W. (2016). (See References). A discussion of how information structure guides language processing and anticipation of new information.
- Huetig, F. (2015). (See References). An overview of research on prediction in language processing expressing some skepticism regarding the importance of prediction for successful comprehension.
- Kuperberg, G. R., & Jaeger, T. F. (2016). (See References). A contemporary, nuanced review of the potential importance of prediction for language comprehension.
- Otero, J., & Kintsch, W. (1992). Failures to detect contradictions in a text: What readers believe versus what they read. *Psychological Science*, *3*, 229–235. An accessible, representative study from the literature investigating integration and discourse comprehension.
- Otten, M., & Van Berkum, J. J. A. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes*, *45*, 464–496. A study using cognitive neuroscience techniques to distinguish discourse-based prediction from word priming.

Action Editor

Randall W. Engle served as action editor for this article.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This research was supported by National Science Foundation Grant BCS-1650888 (to F. Ferreira) and National Institutes of Health Grant 5R56AG053346 (to F. Ferreira, John M. Henderson, and Tamara Swaab).

Note

1. We thank John M. Henderson for suggesting this idea and terminology.

References

- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*, 247–264.
- Bansal, S., Ford, J. M., & Spering, M. (2018). The function and failure of sensory predictions. *Annals of the New York Academy of Sciences*, *1426*, 199–220.
- Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language*, *93*, 203–216.
- Clark, A. (2013). Are we predictive engines? Perils, prospects, and the puzzle of the porous perceiver. *Behavioral & Brain Sciences*, *36*, 233–253.

- Dell, G. S., & Chang, F. (2013). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), Article 20120394. doi:10.1098/rstb.2012.0394
- DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*, 1117–1121.
- Duffy, S. A., Henderson, J. M., & Morris, R. K. (1989). Semantic facilitation of lexical access during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 791–801.
- Ferreira, F., & Lowder, M. W. (2016). Prediction, information structure, and good-enough language processing. *Psychology of Learning and Motivation*, *65*, 217–247.
- Foertsch, J., & Gernsbacher, M. A. (1994). In search of complete comprehension: Getting “minimalists” to work. *Discourse Processes*, *18*, 271–296.
- Gernsbacher, M. A. (1991). Cognitive processes and mechanisms in language comprehension: The structure building framework. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 217–264). New York, NY: Academic Press.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371–395.
- Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*, 357–364.
- Hale, J. (2016). Information-theoretical complexity metrics. *Language and Linguistics Compass*, *10*, 397–412.
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*, 512–521.
- Henderson, J. M., Choi, W., Lowder, M. W., & Ferreira, F. (2016). Language structure in the brain: A fixation-related fMRI study of syntactic surprisal in reading. *NeuroImage*, *132*, 293–300.
- Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163–182.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363–394.
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*, 32–59.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126–1177.
- Long, D. L., & Lea, R. B. (2005). Have we been searching for meaning in all the wrong places? Defining the “search after meaning” principle in comprehension. *Discourse Processes*, *39*, 279–298.
- Lowder, M. W., & Ferreira, F. (2016). Prediction in the processing of repair disfluencies: Evidence from the visual-world paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 1400–1416.
- Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology*, *88*, 22–60.
- Lupyan, G., & Clark, A. (2015). Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychological Science*, *24*, 279–284.
- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99*, 440–466.
- Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 211–277). New York, NY: McGraw-Hill.
- Murray, J. D., Klin, C. M., & Myers, J. L. (1993). Forward inferences in narrative text. *Journal of Memory and Language*, *32*, 464–473.
- Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., . . . Kohu, Z. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *eLife*, *7*, Article e33468. doi:10.7554/eLife.33468
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic theory of dialogue. *Behavioral & Brain Sciences*, *27*, 190–226.
- Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 720–732.
- Singer, M., & Ferreira, F. (1983). Inferring consequences in story comprehension. *Journal of Verbal Learning and Verbal Behavior*, *22*, 437–448.
- Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*, 302–319.
- Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language & Linguistics Compass*, *9*, 311–327.
- Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 443–467.