



Prediction in reading: A review of predictability effects, their theoretical implications, and beyond

Roslyn Wong¹ · Erik D. Reichle¹ · Aaron Veldre^{1,2}

Accepted: 13 September 2024 / Published online: 31 October 2024
© The Author(s) 2024

Abstract

Historically, *prediction* during reading has been considered an inefficient and cognitively expensive processing mechanism given the inherently generative nature of language, which allows upcoming text to unfold in an infinite number of possible ways. This article provides an accessible and comprehensive review of the psycholinguistic research that, over the past 40 or so years, has investigated whether readers are capable of generating predictions during reading, typically via experiments on the effects of *predictability* (i.e., how well a word can be predicted from its prior context). Five theoretically important issues are addressed: What is the best measure of predictability? What is the functional relationship between predictability and processing difficulty? What stage(s) of processing does predictability affect? Are predictability effects ubiquitous? What processes do predictability effects actually reflect? Insights from computational models of reading about how predictability manifests itself to facilitate the reading of text are also discussed. This review concludes by arguing that effects of predictability can, to a certain extent, be taken as demonstrating evidence that prediction is an important but flexible component of real-time language comprehension, in line with broader predictive accounts of cognitive functioning. However, converging evidence, especially from concurrent eye-tracking and brain-imaging methods, is necessary to refine theories of prediction.

Keywords Brain imaging · Eye movements · Predictability effects · Prediction · Reading · Reading models

Reading is seemingly one of the simplest tasks that individuals engage in on a daily basis, yet understanding how it unfolds poses an ongoing challenge for psycholinguistic research. Successful reading is a complex process—comprehenders must decode written text into abstract representations of individual letters and words, retrieve their meanings from long-term memory, and integrate this information into the unfolding discourse representation. Despite its complexity, this entire process unfolds within hundreds of milliseconds with little conscious effort. One explanation for how readers are able to accomplish this feat so quickly and effortlessly lies in the idea of *prediction*.

In its broadest sense, prediction refers to any type of process that uses information about the past or present to estimate the immediately relevant future. In the context of language comprehension,¹ readers encountering a phrase like “The day was breezy so the boy went outside to fly a . . .” can easily predict the word “kite” based on their prior knowledge and experiences (DeLong et al., 2005). These activated representations subsequently allow readers to process information faster when eventually encountered in the input stream. Psycholinguistic research over the past 40 or so years has provided evidence that readers rely on prediction—typically, by investigating the effects of *predictability*—that is, how well a word can be predicted from its prior context. This has led to several reviews in the literature, most of which focus on explaining the underlying mechanisms of prediction (Ferreira & Chantavarin, 2018; Huettig, 2015; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018; Ryskin &

✉ Roslyn Wong
roslyn.wong@mq.edu.au

¹ School of Psychological Sciences, Macquarie University, Sydney, Australia

² Graduate School of Health, University of Technology Sydney, Sydney, Australia

¹ Although this article is about the role of prediction in reading, what has been learned about this topic will likely have broader implications for the more general understanding of language comprehension.

Nieuwland, 2023), or on reviewing empirical evidence from eye tracking (Staub, 2015) or electrophysiology (Van Petten & Luka, 2012). However, whether evidence of predictability effects means that prediction is a central component of real-time language comprehension remains under debate (Huettig & Mani, 2016). Many more studies of predictability effects across different areas of research have been published over the past decade, providing further important insights into the nuanced role of prediction during online processing. As such, the aim of the present review is to update the current state of knowledge about prediction and predictability effects, explore their theoretical implications, and outline potential future research directions. The overarching theoretical issues are whether effects of predictability can be taken to index genuine linguistic prediction during reading and whether, by extension, prediction must occur on some level for successful language comprehension.

This article is broadly organized into four sections. The remainder of this first section defines prediction in the context of both cognition and language comprehension. The following section reviews eye-tracking and brain-imaging research on the effects of predictability in studies of reading before exploring the mechanistic accounts of predictability effects provided by computational models of word identification and reading. The next section considers a number of prominent theoretical issues that provide more refined insights into the nature of predictability effects. The final section discusses the overarching theoretical issues mentioned above and outlines future research directions.

Prediction in cognition

Prediction arguably plays a pervasive role in every aspect of human cognition. For example, in visual perception, individuals presented with a series of static images that imply motion, like a person diving off a cliff, anticipate ongoing mental representations of the unfolding event (Freyd, 1983; Senior et al., 2000). Similarly, in motor behavior, individuals predict upcoming actions based on internal forward models to facilitate motor control (Wolpert & Flanagan, 2001)—for instance, when using the fingertips to grip an object whose load is increased by a self-generated action such as a moving arm, individuals adjust their grip force without delay to prevent the object from slipping (Flanagan & Wing, 1997; Johansson & Cole, 1992). While early investigations of anticipatory mechanisms and representations began in the context of these lower-order sensory modalities of visual and motor cognition, prediction has since become implicated in more complex cognitive processes such as music processing (Land & Furneaux, 1997), emotion processing (Herwig et al., 2007), and theory of mind (Frith & Frith, 2006).

This experimental evidence is consistent with the view that the primary function of the brain is to anticipate future events and stimuli. While the classical view of the brain emphasized the role of lower-level sensory information in the construction of higher-level internal representations (Marr, 1982), the current twenty-first century perspective reverses this approach: “The brain is no longer viewed as a transformer of ambient sensations into cognition, but a generator of predictions and inferences that interprets experience according to subjective biases and statistical accounts of past encounters” (Mesulam, 2008, p. 368). Indeed, Clark (2013) posits that prediction offers a “deeply unified account of perception, cognition, and action” (p. 186) and goes as far as to surmise that the brain is fundamentally a “prediction machine.”

In line with this perspective, an increasing number of brain function models have become framed around predictive coding and minimizing prediction error (Clark, 2013; Friston, 2010; Hohwy, 2013, 2020). According to this framework, the brain is a hierarchically structured device that predicts sensory inputs via the interplay of backward (top-down) and forward (bottom-up) flows of information. The “backward” flow delivers predictions from higher to lower hierarchical levels based on what the system “knows” about the world and its current context such that incoming sensory signals that are consistent with these predictions are “explained away.” However, if there is a discrepancy between the sensory signals encountered and those predicted, the “forward” flow computes an error—that is, residual “unexpected” information that is propagated back up the hierarchy to refine top-down hypotheses about the current sensory data. Thus, the primary function of this multi-level bidirectional information exchange is to minimize overall prediction error or “free energy”—the probability of being in a state of surprise (Friston, 2010). This ensures that, in the long term, predictions about the world are being continuously updated and refined through knowledge and experiences (Bar, 2007; Lupyan & Clark, 2015), and an optimal model of the causes of incoming sensory signals are achieved across the different hierarchical levels (Friston, 2010).

Although empirical evidence that the brain implements predictive coding remains mixed (see Walsh et al., 2020, for a review), the general shift towards a predictive view of the brain is important for explaining cognitive processes for several reasons. Firstly, it has been posited that the brain would not be able to make sense of the world as rapidly as it does if it only relied on bottom-up information (Mesulam, 2008). Secondly, and perhaps more importantly, it has been argued that incoming sensory data is just too ambiguous and complex to deal with in a bottom-up fashion. As such, top-down processes, like the continual generation of predictions, help the brain perceive stability and coherence in its environment (Bar, 2007). Thus, prediction is now viewed as a, if not *the*, fundamental principle of human information processing.

Prediction in language comprehension

Given that language processing is realized by the brain, it is perhaps unsurprising that prediction has been used to explain how real-time language comprehension unfolds so rapidly and effortlessly (Ferreira & Chantavarin, 2018; Huetig, 2015; Kuperberg & Jaeger, 2016; Lupyan & Clark, 2015; Pickering & Gambi, 2018). If the language processor is able to predict linguistic input based on prior knowledge and experiences, this information should facilitate subsequent processing when the input is eventually encountered. From this viewpoint, the role of prediction in theories of language comprehension should not be controversial. However, for decades, there have been doubts as to whether comprehenders can anticipate specific linguistic content in a way that goes beyond the effects of low-level intralexical priming.

The first major reason for this skepticism towards a central role for prediction in reading is that theories of language comprehension have traditionally espoused a strong bottom-up bias. For example, the classic modular views of language processing (Fodor, 1983; Forster, 1979) posit that words are recognized solely based on sensory input, and that context only has a postlexical impact, affecting the ease with which words are integrated into the unfolding sentence or discourse representation. Thus, the idea that contextual information could have a prelexical impact by encouraging the anticipation of upcoming linguistic input seemed untenable. Even more lenient views of language processing, such as those offered by the *cohort model* (Marslen-Wilson, 1987) and the *shortlist model* (Norris, 1994), which claim to be more interactive, still have a fundamentally bottom-up emphasis. In these models, contextual information can influence word recognition and lexical access, but only after the sensory input has activated an initial set of potential candidates.

The second major reason for why researchers have historically rejected the idea of prediction in reading lies in the inherently generative nature of language, which allows infinite possible linguistic expressions for each upcoming word of a sentence (Jackendoff, 2002; Morris, 2006). For example, consider the *cloze task*, in which individuals are instructed to continue a phrase or complete a sentence with the first word that comes to mind (Taylor, 1953). If most individuals converge on the same word, the context is considered *predictive* or *constraining* and the completion is deemed *predictable* or *high cloze probability*. However, most naturally occurring text is only weakly or moderately constraining (see, e.g., Provo Corpus; Luke & Christianson, 2016, 2018), so participants will often provide multiple plausible responses for these contexts (Bloom & Fischler, 1980). As such, predicting what will come next in a sentence is a relatively ineffective and potentially costly strategy unless contextual constraint is unusually strong—comprehenders are unlikely

to make correct predictions very often and it may be more efficient to wait for the sentence to unfold naturally over the next few words (Forster, 1981; Jackendoff, 2002).

These antiprediction views, however, began to shift in the late 1990s. The modular views of language processing gave way to interactive views that emphasized the availability of contextual information even before the current sensory input had been processed (Ferreira & Lowder, 2016; McClelland, 1987; Morton, 1969; Sedivy et al., 1999). Meanwhile, the argument about generativity was challenged by the fact that, while few words are highly predictable in natural reading, many words or aspects of words are still moderately predictable. For example, consider again the first example presented in this article: “The day was breezy so the boy went outside to fly a kite.” Even though most comprehenders are unlikely to accurately predict in advance the word “fly” in this sentence, they can be fairly confident that a verb will be in this location. This syntactic information can then be used in conjunction with semantic and/or world knowledge to converge on a sufficiently plausible continuation. In line with these shifts in perspective, the past 40 or so years has seen growing experimental evidence of the predictive nature of the language processor.

The earliest evidence that comprehenders generate linguistic predictions during online processing comes from studies using lexical decision and speeded pronunciation (or “naming”) tasks. Many of these studies found that individuals were faster to make lexical decisions or name words that were contextually supported, such as “locker” in “John kept his gym clothes in his...,” compared with words that were not contextually supported, such as “closet” (Fischler & Bloom, 1985; Kleiman, 1980; Schwanenflugel & LaCount, 1988; Schwanenflugel & Shoben, 1985; Stanovich & West, 1983; Traxler & Foss, 2000). However, there are reasons to question the extent to which these findings from behavioral tasks can be generalized to normal reading given that the specific task demands of lexical decision and naming may recruit processing strategies that are not part of normal reading comprehension (Rayner & Liversedge, 2011).

Further evidence comes from studies using the *visual world paradigm* (Tanenhaus et al., 1995), in which participants’ eye movements are recorded as they look at a visual scene while listening to a sentence. In one of the most influential demonstrations, Altmann and Kamide (1999) presented participants with a visual context depicting a boy, a cake, and several other distractor objects while they listened to sentences such as “The boy will eat the cake” or “The boy will move the cake.” They found that, before the final word had even been presented, participants were faster to move their eyes towards the cake, the only edible object, in the “eat” compared with the “move” condition, suggesting that participants had predicted a compatible theme based on the

selectional information conveyed by the verb. This finding of anticipatory eye movements based on contextual information has since been replicated in a number of visual world studies (Kamide et al., 2003; Kukona et al., 2011; see Huettig et al., 2011, for a review). However, questions have also been raised about whether these findings reflect genuine linguistic prediction given that the visual context places constraints on the potential candidates that can be heard in each sentence (DeLong, Troyer et al., 2014b; Huettig et al., 2011; Kutas et al., 2011). In other words, it is unclear whether evidence of prediction in visual world studies reflects anticipation due to linguistic or visual information.

Given the interpretational issues with these paradigms, it is unsurprising that researchers have turned to investigating lexical prediction during online processing via studies of reading, which have the potential to present a broad range of linguistic structures. This has coincided with the development of online methodologies such as eye-tracking and brain-imaging techniques. These methods are well-suited to the investigation of how real-time language comprehension unfolds because of their ability to provide continuous streams of data with high temporal resolution. In studies of reading, the clearest demonstration of prediction comes from evidence that a word has been activated even before it has been encountered. For example, DeLong et al. (2005) capitalized on a phonological rule in English where the indefinite article is realized as *a* before a consonant and *an* before a vowel, and presented sentences like “The day was breezy so the boy went outside to fly . . .” which were completed by either the predictable noun-phrase “a kite” or the less predictable, but plausible, noun-phrase “an airplane.” They found that neural activity in the form of an N400 component correlated with the predictability of the noun—the more predictable the noun, the smaller the neural activity, reflecting the ease of semantic processing. Critically, this inverse correlation was also obtained before the noun was presented—that is, on the preceding article that carried no semantic information. This led DeLong et al. to conclude that comprehenders had used the preceding context to anticipate the predictable noun, or at least its first phoneme, and by extension its appropriate preceding article (see also DeLong et al., 2012; Martin et al., 2013). While these findings could be taken as evidence of prediction, they should be interpreted with caution because subsequent studies failed to replicate these effects on the article (Ito et al., 2017; Nieuwland et al., 2018; but see Urbach et al., 2020). More generally, these manipulations are difficult to implement in English because there are few systematic linguistic rules that allow the properties of lexical items to influence preceding words (but see Fleur et al., 2020; Otten & Van Berkum, 2009; Van Berkum et al., 2005, for evidence from Dutch; Foucart et al., 2014; Wicha, Bates, et al., 2003a; Wicha, Moreno, et al., 2003b; Wicha et al., 2004, for evidence from Spanish).

Thus, most studies of reading in English have focused on demonstrating lexical prediction via the effects of *predictability*—how well a word can be predicted from its prior context. The basic idea is that, if a word can be predicted in advance of its presentation, then the way in which this word is processed when eventually encountered may depend on its level of predictability (Kutas et al., 2011). The most common approach to operationalizing predictability is via a word’s *cloze probability*, which can be determined by aggregating the responses provided for a given sentence context in a cloze task (i.e., the number of productions divided by the total number of responses; Taylor, 1953). Words with cloze values close to 1 are almost perfectly predictable in their contexts while words with low cloze values are less predictable. The next section presents a review of behavioral and neural research of the effects of word-level predictability during reading before exploring the mechanistic accounts of predictability effects provided by computational models of word identification and reading.

Predictability effects in reading

Evidence from eye-movement studies

Eye-movement recording is a noninvasive behavioral method for measuring online cognitive processing during reading. In the context of psycholinguistic research, eye movements offer certain advantages compared with traditional behavioral tasks. Firstly, participants do not need to complete a secondary task during reading, such as making decisions about words or naming them aloud, which could impose additional processing demands. Secondly, and perhaps more importantly, eye movements are fundamental to the reading process itself, meaning that ongoing lexical processing is captured by where and when readers move their eyes. It is therefore unsurprising that the eye-tracking methodology has been used extensively to investigate a range of language representations and processes (see Rayner, 1998, 2009, for reviews), including those related to predictability effects.

Ehrlich and Rayner (1981) first demonstrated the effects of predictability on eye movements in a pair of experiments that manipulated predictability in two ways. The first experiment presented the same target in two different contexts, while the second experiment presented two different target words in the same contexts, so that, in both manipulations, one target had a high cloze probability while the other had a low cloze probability. Across both experiments, Ehrlich and Rayner found that words that could be predicted from the preceding context were more likely to be skipped and to receive shorter reading times if fixated than words that could not be predicted, consistent with the idea that the predictability of a word determines the time required to process it.

Similar predictability effects have since been reported across a number of eye-movement studies using either experimentally controlled approaches, which involve manipulating a specific target word in a sentence context (see Staub, 2015, for a review), or corpus-based approaches, which involve analysing every word in a corpus of sentences or texts (Andrews et al., 2022; Luke & Christianson, 2016; see Kliegl et al., 2004, for evidence from German). Across these studies, predictability effects are typically observed on skipping (Abbott et al., 2015; Cutter et al., 2020; Fitzsimmons & Drieghe, 2013; Frisson et al., 2017; Rayner et al., 2001, 2011; Rayner & Well, 1996; Rich & Harris, 2023; Wong et al., 2022, 2024; see Brysbaert et al., 2005, for a meta-analysis) and first-pass reading measures such as first fixation and gaze duration (Fitzsimmons & Drieghe, 2013; Frisson et al., 2017; Rayner et al., 2001, 2011; Rayner & Well, 1996; Rich & Harris, 2023; Wong et al., 2022, 2024), suggesting that word predictability exerts its influence during early stages of processing (e.g., word recognition and lexical access; Vasishth et al., 2013). However, predictability effects are also observed on “late” reading measures such as total fixation duration (to the extent that it reflects second-pass time rather than first-pass time) and rate of regressions (Frisson et al., 2017; Rayner et al., 2011; Rayner & Well, 1996; Wong et al., 2022, 2024), possibly reflecting the influence of contextual information on later stages of processing (e.g., postlexical integration; Clifton et al., 2007). Thus, across different eye-movement studies, there is robust evidence that the predictability of a word is an important determinant of where readers look and for how long.

Evidence from brain-imaging studies

Brain-imaging techniques allow inferences to be made about neural activity during online cognitive processing. Neural activity, which reflects communication between neurons, is an electrochemical process that generates electrical and magnetic activity, which can be detected via noninvasive electromagnetic techniques such as electroencephalography (EEG) and magnetoencephalography (MEG), respectively. Increases in neural activity are also accompanied by increases in blood flow to the relevant brain region, which can be detected via hemodynamic techniques such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). These four techniques have been used widely in psycholinguistic research, with a significant number of contributions coming from EEG, especially when this technique is used to obtain *event-related potentials* (ERPs). Evidence of predictability effects during reading using these techniques are reviewed in turn, with a focus on ERPs.

Evidence from EEG

EEG is a noninvasive technique for measuring neural activity via electrodes placed on the scalp. It is well-suited to the study of language representations and processes because it provides high-resolution temporal information, which makes it possible to infer the stages of processing influenced by an experimental manipulation, even when participants’ only task is to read for comprehension. This technique can be used to obtain ERPs, which are small voltage fluctuations that are time-locked to an event or stimulus of interest, such as the onset of the presentation of a word, reflecting the summation of synchronized postsynaptic activity generated by a large population of neurons during information processing.

In EEG studies, predictability effects during reading have been shown to impact an ERP component known as the *N400*—a negative-going waveform with a centro-parietal distribution that begins 250 ms after the presentation of a word and peaks 400 ms after stimulus onset. Kutas and Hillyard (1980) initially discovered this neural waveform when they presented sentences like “He spread the warm bread with...” and found larger N400 amplitudes for semantically anomalous words like “socks” than acceptable control words like “butter.” Subsequent research revealed that the sensitivity of this ERP component was not restricted to semantic anomalies, per se: Larger N400 amplitudes were also observed for unpredictable compared with predictable words even when both were equally plausible. Kutas and Hillyard (1984) first demonstrated this when they presented sentences completed by either a high, medium, or low cloze probability target word and found an inverse relationship between predictability and the N400 component. That is, the higher a word’s cloze probability, the smaller the corresponding N400 amplitude, perhaps reflecting the fewer neural resources required to process a word. This graded relationship between predictability and the N400 component has been replicated in a number of studies (Brothers et al., 2015, 2017, 2020; Davenport & Coulson, 2011; DeLong, Quante, et al., 2014a; Federmeier et al., 2007; Hubbard et al., 2019; Kuperberg et al., 2020; Kutas, 1993; Lai et al., 2021, 2023; Rommers & Federmeier, 2018b; Thornhill & Van Petten, 2012).

The N400 component has also been shown to be sensitive to manipulations that go beyond predictability at the lexical level (see Federmeier, 2022; Kutas & Federmeier, 2007, 2011, for reviews). This ERP waveform is also modulated by predictability manipulations at the orthographic and phonological level (DeLong et al., 2005; Ito et al., 2016; Laszlo & Federmeier, 2009; Martin et al., 2013), semantic level (Federmeier & Kutas, 1999; Kutas et al., 1984; Metusalem et al., 2012; Thornhill & Van Petten, 2012; Wlotko & Federmeier, 2015), and even for emojis (Weissman et al.,

2024). The N400 is also sensitive to other lexical properties of words such as their frequency (Rugg, 1990), concreteness (Barber et al., 2013), and position in the sentence (Van Petten & Kutas, 1990, 1991). Indeed, the N400 component does not even appear to be language-specific because it is also elicited in response to any type of input that produces activity in long-term semantic memory, including faces and objects, numeric symbols, and environmental sounds (see Federmeier et al., 2016, for a review). Taken together, these findings suggest that the N400 component may be a default neural response to any potentially meaningful perceptual stimuli. Nonetheless, the fact remains that its amplitude is strongly correlated with a word's cloze probability.

The timing of ERP predictability effects, however, suggests that they may occur relatively late in the time course of normal reading—that is, 400 ms after presentation of a word—or, at least, that they take some time to appear in the ERP record. Given that the typical eye fixation only lasts 200–250 ms, by 400 ms after a word has been fixated, readers have likely completed lexical access for that word and moved their eyes onto the subsequent word (Dimigen et al., 2011; Rayner & Clifton, 2009). As such, there remains an ongoing debate in the literature as to whether reduced N400 effects during reading reflect processing difficulty during the lexical or postlexical stage of processing (Kutas & Federmeier, 2011; Nieuwland et al., 2020). This, in turn, has implications for whether these effects can be taken as reflecting genuine prediction processes, which would be expected to influence early stages of processing. While there has been some evidence to suggest predictability effects on earlier ERP components, the reliability of these findings remains unclear (see Nieuwland, 2019, for a review).

Evidence from other brain-imaging techniques

Aside from EEG, other brain-imaging techniques used to measure neural activity during online cognitive processing include MEG, PET, and fMRI. However, evidence of predictability effects in studies of reading using these methods is comparatively scarce.

MEG is a noninvasive technique for measuring neural activity via magnetic sensors located a few centimeters from the scalp. The technique can be used to obtain *event-related fields* (ERFs), which, like their electrical counterparts (i.e., ERPs), provide high-resolution temporal information even during the simplest reading comprehension tasks. Additionally, because the anatomical origins of magnetic signals are easier to localize (see Van Petten & Luka, 2006, for a review), this technique also has the advantage of providing high-resolution spatial information. Predictability effects during reading have been shown to impact an ERF component known as the magnetic *N400* (N400m), which occurs around 300 to 500 ms poststimulus onset, corresponding to

the N400 waveform observed in the ERP record. For example, Halgren et al. (2002) observed smaller N400m amplitudes for predictable compared with unpredictable words in a sentence reading task (see also Helenius et al., 1998, for evidence from Finnish), although it remains unclear whether this study separated lexical predictability from contextual plausibility. Similar to its electrical counterpart then, the N400m likely reflect a general response to meaningful perceptual stimuli (Ihara et al., 2007; Maess et al., 2006). More recently, predictability effects in the MEG record have also been linked to differences in oscillatory activity in left temporal regions of the brain (see Wang, Hagoort et al., 2018a, for evidence from Dutch). This region has also been shown to produce more similar patterns of neural activity for predictions of the same word in different contexts compared with predictions of different words (see Wang, Kuperberg et al., 2018b, for evidence from Chinese). However, these emerging areas of research and their links to predictive processing remain poorly understood.

The other techniques, PET and fMRI, measure changes in regional cerebral blood flow associated with brain activity. PET involves the injection of a radioactive, positron-emitting contrast agent and, perhaps due to this invasive procedure, is not commonly used in psycholinguistic research with no studies to date having investigated predictability effects during reading. fMRI, on the other hand, yields similar patterns of activation as the PET method (see Price, 2012, for a review), but is less invasive and does not involve radiation exposure. fMRI, also, has the benefit of providing high-resolution spatial information, although its low-resolution temporal information has limited its use in psycholinguistic research. In fMRI studies, predictability effects during reading have been linked to decreased activation in left frontal and temporal regions of the brain (Dien et al., 2008; see also Baumgaertner et al., 2002; Hartwigsen et al., 2017, for evidence from German).

Evidence from brain-imaging techniques provides important insights into the neural bases of predictability effects during reading. However, the extent to which these findings generalize to normal reading remains unclear. Brain-imaging studies often present stimuli using a *rapid serial visual presentation* (RSVP) *paradigm* in which sentences or texts are displayed one word at a time in the center of the screen at a fixed pace of 400–1,000 ms per word, which is assumed to provide sufficient time to separate neural responses to individual words. This presentation method clearly differs from normal reading in many aspects. Firstly, the word-by-word presentation format prevents readers from engaging in natural eye-movement behavior including skipping words, rereading text, and extracting upcoming information from the parafovea. Secondly, the fixed-pace presentation rate does not allow readers volitional control over the rate of input. Instead, words are typically presented for longer than

the standard 200–250 ms duration of fixations, which could contribute to an increased likelihood of conscious prediction strategies (Dambacher et al., 2012; Wlotko & Federmeier, 2015). Finally, readers are required to maintain central fixation and suppress eye movements and blinks during the task. As such, reading under RSVP is quite unnatural and could impose additional processing demands that are not required in normal reading.

To partly address the limited ecological validity of RSVP, some brain-imaging studies have presented stimuli using a *self-paced reading* (SPR) paradigm in which sentences or texts are presented one word at a time but at the reader's own pace (Ditman et al., 2007). Behavioral studies using this paradigm to investigate predictability effects have reported the classic processing benefits on reading times for expected compared with unexpected words, as measured by the latency of button presses (Brothers et al., 2017; Jongman et al., 2022). ERP studies using this paradigm have also revealed the expected smaller N400 amplitudes for predictable compared with unpredictable words (Ng et al., 2017; Payne & Federmeier, 2017b, 2019). Although these SPR findings appear to yield the same patterns of effects as the RSVP paradigm, they should also be interpreted with caution given the limitations of word-by-word presentation formats described above. Thus, it is possible that brain imaging studies using RSVP or SPR paradigms do not accurately capture the processes underlying normal reading, which is an important implication when considering discrepancies in the conclusions of eye-tracking and brain-imaging research.

Evidence from concurrent eye-tracking and brain-imaging techniques

Researchers have begun to address the gap between brain-imaging studies and natural reading via co-registration methods, which involve the simultaneous recording of eye-movement and brain activity data from the same participants (see Himmelstoss et al., 2020, for a review). These methods of investigating continuous brain activity under natural reading have the advantage of enabling a more comprehensive assessment of the neural correlates underpinning online cognitive processing. They also allow for the analysis of brain activity that is time-locked to the onset of a fixation on a critical word, rather than to the onset of an externally triggered stimulus. However, only a small number of studies of reading have used co-registration methods to investigate predictability effects during reading. One approach has been the combination of eye movements and EEG to extract *fixation-related potentials* (FRPs), which has revealed the classic effects of cloze probability on both eye-movement measures and the amplitude of the N400 component (Burnsky et al., 2022; Kretzschmar et al., 2015; see also Dimigen et al., 2011; Kretzschmar et al., 2009, for evidence from

German). Another approach has been the combination of eye movements and fMRIs to yield *fixation-related fMRIs* (Marsman et al., 2012), which has demonstrated that lexical predictability is both a significant predictor of reading times and involves a broad network of brain regions (Carter et al., 2019; see also Schuster et al., 2016, 2020, 2021; Weiss et al., 2023, for evidence from German). Although co-registration studies appear to replicate the findings of conventional brain-imaging studies, the interpretation of these findings should be qualified by the fact that events that occur at the same time in both records do not always necessarily reflect the same underlying processes. As with ERP data, the time-locked effects observed in FRP data, such as the N400, occur relatively late in the time course of normal reading. This means that not only have readers' eyes likely moved onto the subsequent word, but the FRP waveforms risk being contaminated by neural processing from unrelated events. Nonetheless, with further ongoing development and refinement, the recording of neural activity during natural reading is likely to be a useful methodology for consolidating existing evidence of predictability effects in eye-tracking and brain-imaging research.

Predictability effects within computational models of word identification and reading

On the basis of the above findings, a number of computer models have been put forward to formalize how predictability manifests itself to facilitate the reading of text. Computer models of reading can be grouped by the types of phenomena they attempt to explain—the identification of isolated words, the syntactic and semantic operations necessary to understand phrases and sentences, the processing required to represent more extended discourse, or how these processes are coordinated with vision and attention to produce the patterns of eye movements observed during reading (see Reichle, 2021, for a review). Because of this diversity and the fact that the models are formally implemented to varying degrees (e.g., flowcharts vs. computer programs), they differ considerably in terms of their assumptions about the processes involved in predictability effects. This section reviews computer models to provide a brief overview of their different assumptions and to highlight some of their theoretically interesting points of contrast. Firstly, models in each of the four previously mentioned groups developed to explain behavior during reading are considered, followed by models developed to simulate brain activity during reading.

Word-identification models

Models of word identification have largely been developed to explain the identification of single isolated words as measured using either lexical decision or naming tasks. As

such, although the principles instantiated by these models are intended to “scale up” to explain how words are identified in natural reading, the details of how this occurs are unclear. Many central questions—such as how a word’s predictability might influence its identification—are either ignored or discussed at an abstract level. For example, the earliest word identification models were either silent about the role of predictability (Carr & Pollatsek, 1985; Coltheart, 1978; Glushko, 1979; Paap et al., 1982) or assumed that higher-level semantic and/or syntactic information could (somehow) influence lexical representations in a top-down manner (Forster, 1976; Morton, 1969). Although subsequent models were fully implemented and thus made more precise assumptions about how words are identified, the role of predictability in these models continued to be largely ignored (Gomez et al., 2008; Van Rijn & Anderson, 2003; Whitney, 2001). For example, the *interactive-activation model* (McClelland & Rumelhart, 1981) and most of its progeny (Coltheart et al., 2001; Davis, 2010; Perry et al., 2007; Zorzi et al., 1998) assume that words are represented by a hierarchical network of interconnected nodes, with information about the visual features of a printed word being propagated from feature nodes upwards through letter and word nodes in a highly interactive manner, with the word node receiving the most bottom-up activation eventually suppressing all other word nodes as it is identified. Complementing this description is the assumption that the activation of word nodes can also be supported through the top-down propagation of activation from higher-level nodes representing the context in which the words might be embedded. The latter assumption has not been implemented in models of word identification; however, as will be discussed below, it has been implemented in models of eye-movement control that incorporate variants of the interactive-activation model within their frameworks. One other interesting example of how word identification models might accommodate predictability effects is provided by the *Bayesian Reader* model (Norris, 2006). In this model, the frequency with which a given word has been previously encountered in text influences the prior odds of perceiving a given sequence of letters as that word. The linguistic context in which a word is embedded might similarly influence those prior odds, although precisely how this occurs also remains unspecified.

Sentence-processing models

The second category of reading models are those related to the processing of sentences. It is important to first note the starting assumption of these models which is that words have been accurately identified. The main objective of sentence processing models then is to explain how information about the meanings and syntactic categories of those words are used to construct the larger units of meaning corresponding

to phrases and whole sentences. The earliest sentence processing models, for example, mainly used syntactic information from phrase structure rules to “decide” the likely syntactic categories of upcoming words in order to parse sentences as efficiently as possible (Frazier, 1978; Frazier & Fodor, 1978; Woods, 1970). By this conceptualization, syntactic information was privileged because it was assumed to influence sentence processing more rapidly than semantic information. Later models, however, challenged this assumption by positing that semantic and pragmatic information could be used rapidly to guide the parsing of sentences (Elman, 1990; Jurafsky, 1996; Just & Carpenter, 1992; Lewis & Vasishth, 2005; MacDonald et al., 1994; McClelland et al., 1989; Tabor et al., 1997; Van Dyke & Lewis, 2003). Although these models differ markedly in their implementation, they share the same underlying assumption that lexical, syntactic, semantic, and pragmatic information can be used jointly to constrain the identities of upcoming words.

One interesting variant of these sentence processing models that warrants discussion is *surprisal* (Hale, 2001; Levy, 2008), because it provides both an explanation for and a metric of the variation that is observed in word processing difficulty as a function of predictability. Unlike the accounts of sentence processing described above, which focus on the coordination of distinct cognitive operations that determine how words are recognized and integrated into unfolding sentence representations, surprisal theory is based on an “inferential” view of sentence processing according to which the primary role of the language processor is to assign a probability distribution over all possible continuations of a sentence (Shain, 2024). Processing effort is thus argued to be proportional to the magnitude of change in this probability distribution given by the amount of new information conveyed by each additional word. This process is also quantified by the metric of surprisal (a word’s negative log probability; see section Surprisal and entropy), which functions as a “*causal bottleneck* between representations and behavioral observables” (Levy, 2008, p. 1132). As with the other models of sentence processing, however, the assumptions of surprisal theory are general in two senses. The first is that a variety of factors (e.g., lexical, syntactic, semantic) that are distinct from the process of probabilistic inference likely contribute to language comprehension, with the precise mechanism(s) being unspecified (see Staub, 2024, for a review). The second is that predictions about the relative benefit or cost that are due to a word’s surprisal or predictability are often qualitative (e.g., a word will be identified more rapidly in Condition A than Condition B) or imprecise (e.g., words in Region A will be read more rapidly than words in Region B). It should be noted though that these two limitations reflect the fact that these models tend to emphasize the products of sentence processing rather than how the time course of sentence processing influences measures of online reading (e.g., eye movements).

Discourse-processing models

The limitations of sentence processing models are even more pronounced with models of discourse representation which often start with abstract, propositional representations of phrases or sentences and are even further removed from online processing. As such, these models generally have very little to say about the identification of specific words, but instead describe how the literal content of a text is used in conjunction with information in semantic memory to make a variety of inferences that are necessary for understanding and elaborating upon the text. For example, the earliest discourse representation models assumed that readers either use their knowledge of common story “scripts” to actively predict upcoming elements (Rumelhart, 1975) or rely upon the passive spreading of activation among related concepts in semantic memory to link the propositions of a text with whatever prior information might be relevant (Kintsch & van Dijk, 1978). Although recent models tend to emphasize the latter, specifying in more detail how spreading activation affords the making of inferences (Goldman & Varma, 1995; Kintsch, 1998; Myers & O’Brien, 1998; van den Broek et al., 1996), there are examples of models that maintain the core assumption that readers’ knowledge of causality plays an important role in understanding discourse (Fletcher & Bloom, 1988; Frank et al., 2003; Golden & Rumelhart, 1993; Langston & Trabasso, 1999). Like sentence processing models, however, all models of discourse representation are limited to making fairly coarse predictions about both the nature of the mechanism(s) that are responsible for predictability effects and how these effects manifest themselves during reading.

Models of the reading architecture

The final group of reading models are those that specify how word identification and, to a lesser degree, sentence

processing interacts with the visual system and attention to determine when and where the eyes move during reading. Early models, which are not specific about these processes, make no attempt to explain predictability effects (Gough, 1972; Morrison, 1984; Reilly, 1993; see also McDonald et al., 2005). However, later models of the reading architecture, which provide more global descriptions of these processes albeit in an abstract, higher-level way, often make explicit provisions for explaining how predictability influences reading times. For example, the *E-Z Reader* model (Reichle et al., 1998, 2009, 2012) assumes that word n is identified in two successive stages, L_1 and L_2 , with the mean time to complete the first stage, $t(L_1)$, being described by Eq. 1 and the mean time to complete the second stage, $t(L_2)$ being a fixed proportion of $t(L_1)$. As the top “branch” of Eq. 1 shows, word n can be “guessed” from its preceding sentence context, causing the value of $t(L_1)$ to equal 0 ms, with a probability equal to word n ’s cloze probability (i.e., $predictability_n$). Because most words have low predictability, however, the duration of $t(L_1)$ is most often a linear function of three free parameters ($\alpha_1 = 124$, $\alpha_2 = 11.1$, and $\alpha_3 = 76$), the logarithm of word n ’s frequency of occurrence (i.e., $frequency_n$), and its predictability.² The model also assumes that, for word n to be predictable, all the preceding words must be identified and integrated; if the meaning of word $n - 1$ had not been integrated, the value of $predictability_n$ is equal to 0. Thus, although highly predictable words (e.g., function words) can sometimes be guessed outright, the time required to identify most words is roughly an additive function of its frequency and predictability. Collectively, these assumptions allow E-Z Reader to accurately simulate the results of eye-tracking experiments (e.g., Rayner et al., 2004) which show that word frequency and predictability rapidly affect both the propensity to skip and the fixation durations on those words.

$$t(L_1) = \begin{cases} 0, \\ \alpha_1 - \alpha_2 \log frequency_n - \alpha_3 predictability_n, \end{cases}$$

$$\begin{aligned} &\text{with } p = predictability_n \\ &\text{with } p = 1 - predictability_n \end{aligned} \quad (1)$$

Other models of eye-movement control make similar assumptions about predictability effects. For example, the *SWIFT* model (Engbert et al., 2002, 2005) additionally assumes that how a word’s predictability affects its rate of processing depends upon whether lexical processing is in its first or second stage, and whether the word is being fixated or processed in the parafovea. This is accomplished using Eq. 2, where F_n is a variable that modulates the activation level of word n as a function of its current stage of processing and location. In Eq. 2, t is an index of time and t_p is the point in time when word n reaches its maximum activation, $f (= 70.2)$ is a free parameter that increases the rate at which

word n becomes active prior to this point, $\theta (= 0.11)$ is a free parameter that modulates the overall effect of predictability, and k is the point of fixation. Thus, in the top “branch” of

² The values of these parameters (and those of the other models that will be discussed below) are typically selected to minimize the discrepancy between observed measures of reading behavior (e.g., mean fixation durations and probability measures) and the model’s predicted values of those same measures. It is also important to note that the explanation provided here ignores the additional complexity of retinal eccentricity; in E-Z Reader, $t(L_1)$ also increases as a function of the distance between the location of word n and the centre of vision due to decreases in visual acuity and increasing effects of crowding (see Veldre et al., 2023).

Eq. 2, if word n is in the initial stage of processing and being processed from the parafovea, its rate of activation will be attenuated for more predictable words because such words require less bottom-up support to be identified. However, after word n reaches its maximum activation, its activation declines but more slowly for more predictable words. Together, these assumptions allow SWIFT to accurately simulate the kinds of predictability effects reported in eye-movement experiments (e.g., Kliegl et al., 2006).

$$F_n = \begin{cases} +f(1 - \theta \text{ predictability}_n), & \text{if } t < t_{p(n)} \text{ and } k < n \\ +f, & \text{if } t < t_{p(n)} \text{ and } k \geq n \\ -(1 + \theta \text{ predictability}_n), & \text{if } t \geq t_{p(n)} \end{cases} \quad (2)$$

One additional model of eye-movement control that warrants discussion is Li and Pollatsek's (2020) *Chinese Reading Model* (CRM), which is designed to explain the patterns of eye movements observed during the reading of Chinese. Due to the unique writing system of the Chinese language, this model faces additional challenges that are not evident in the reading of languages that use alphabetic writing systems (see Reichle & Yu, 2024, for a discussion of these issues). Ignoring these script-related differences, however, the model is like both *Glenmore* (Reilly & Radach, 2006) and *OBI-Reader* (Snell et al., 2018) in that they combine variants of the interactive-activation model of word identification (McClelland & Rumelhart, 1981) described earlier and additional assumptions about the mechanisms that “decide” when and where to move the eyes. Importantly, the CRM also makes provisions for explaining predictability effects by adopting the assumption that, across time, t , the activation of a given word node n , $act_n(t)$, will increase at a rate that partially reflects its cloze probability. This assumption is described by Eq. 3, where *gain* ($= 0.18$) is a free parameter that determines the minimum effect of predictability. By adopting the preceding assumption, the CRM, like both E-Z Reader (Reichle et al., 1998, 2009, 2012) and SWIFT (Engbert et al., 2002, 2005), can provide quantitative accounts of eye-movement experiments (e.g., Rayner et al., 2005) that demonstrate how variation in the predictability of Chinese words influence various eye-movement measures on those words. Perhaps just as importantly, by doing this, the CRM makes good on the “promissory note” that, within the framework of the interactive-activation model (McClelland & Rumelhart, 1981), higher-level context can influence the time required to identify individual words.

$$act_n(t + \Delta t) = act_n(t) + \text{predictability}_n + \text{gain} \quad (3)$$

In considering the models of the reading architecture reviewed so far, it is necessary to acknowledge that Eqs. 1–3 describe three different ways in which a word's predictability might influence its rate of processing and identification.

Except for the CRM, which describes how words are identified in some detail, the other models provide only abstract descriptions of how predictability affects the rate of lexical processing. Perhaps for this reason, there have been two recent attempts to develop more comprehensive models of reading by embedding assumptions related to word identification, sentence processing, and/or discourse representation within the frameworks of models of eye-movement control. The first of these, *Über-Reader* (Reichle, 2021), incorporates an instance-based model of word identification (based on Ans et al., 1998) and a handful of assumptions about sentence processing and discourse representation within the framework of E-Z Reader (Reichle et al., 2012). Although this model does not currently explain predictability effects (see Reichle, 2021, pp. 505–506), two suggestions for addressing this limitation are offered. The first is that, in identifying a word, syntactic features consistent with the model's predictions about the word's syntactic category are used, in combination with the word's orthographic features, to facilitate its identification. The second is that general semantic features related to the sentence or discourse topic might likewise be used in combination with the word's orthographic features to facilitate its identification. These two approaches, if used concurrently, might allow for gradations in the specificity of the model's predictions about the identities of upcoming words. The second of these more comprehensive models, *SEAM* (Rabe et al., 2023), embeds the sentence processing model of Lewis and Vasishth (2005) within the framework of the most recent version of SWIFT (Seelig et al., 2020). Although this model arguably provides a more sophisticated account of sentence processing than does *Über-Reader*, like the latter it also fails to provide an account of predictability effects, nor is there any suggestion for how *SEAM* might be modified to explain these effects.

Models of brain activity during reading

The computer models described so far have been developed to simulate behavior during reading and, as such, cannot be directly applied to the explanation of brain activity during reading. In recent years, progress has been made towards the development of models that are capable of simulating the patterns of neural data observed during reading in brain-imaging studies and the role of predictability in this process. Perhaps unsurprisingly though, these models have predominantly been informed by neural activity measured using ERPs given the extensive research conducted using the EEG method compared with other brain-imaging techniques. Computational models of language electrophysiology can be ascribed to two general categories—large-scale models that are based on natural language corpora and small-scale models that are based on specific theoretical considerations.

Large-scale computational models link ERP activity to complex neural networks that have been trained to predict the next word in large corpora of spoken and/or written texts (Devlin et al., 2019; Radford et al., 2019). For example, a number of studies that have derived surprisal estimates of predictability from language models have shown that these probabilistic measures predict N400 amplitudes during natural reading (Aurnhammer & Frank, 2019; Frank et al., 2015; Michaelov et al., 2021). However, only one study by Lindborg and Rabovsky (2021) has suggested that this is because changes in the internal activation states of these models are predictive of N400 amplitudes. Overall, because surprisal, an output measure of these models, is typically used to predict neural activity only, it remains unclear mechanistically how a word's predictability is related to brain activity during reading.

Small-scale computational models, on the other hand, link ERP activity to changes within internal hidden layers of neural networks (see Nour Eddine et al., 2022, for a review). Early models, which were based on the processing of single words and word pairs, focused on simulating the N400 observed in response to lexical factors and simple priming manipulations (Cheyette & Plaut, 2017; Laszlo & Armstrong, 2014; Laszlo & Plaut, 2012; Rabovsky & McRae, 2014). However, because the process of word identification in these models is assumed to involve the mapping of orthographic features onto distributed semantic representations, the details of how sentence-level variables such as predictability influence this ERP component are largely ignored. For example, one class of models posits that N400 amplitudes reflect the amount of semantic activation produced by bottom-up input (Cheyette & Plaut, 2017; Laszlo & Armstrong, 2014; Laszlo & Plaut, 2012), while another class posits that they reflect the discrepancy between the internal semantic state of the model and the semantic features of the input (Rabovsky & McRae, 2014). Later models, which were based on the processing of sentences, focused on simulating the N400 observed in response to broader contextual factors including predictability (Brouwer et al., 2017, 2021; Fitz & Chang, 2019; Rabovsky, 2020; Rabovsky et al., 2018). According to these models, N400 amplitudes reflect either the amount of change in an internal neural network representation layer (Brouwer et al., 2017, 2021; Rabovsky, 2020; Rabovsky et al., 2018) or the difference between the next-word prediction of the model and the input actually presented (Fitz & Chang, 2019). But even though some of these models are able to simulate the effects of predictability on the N400, the magnitude of change associated with this component is generally computed outside the model's architecture, so the direct link between predictability and brain activity during reading in these models remains unclear.

In summary, computer models of reading have the capacity to provide insights into the patterns of behavioral and

neural data during reading in eye-movement and brain-imaging studies. However, most models to date lack sufficient quantitative detail to explain how predictability manifests itself to facilitate the reading of text. More specifically, existing models of brain activity have largely been informed by ERP data, so it remains unclear how their assumptions would apply to data from other brain-imaging techniques. Thus, the ongoing development and refinement of computer models of reading are important for evaluating the feasibility of different accounts of predictability effects and for motivating new empirical research. It will also be a critical challenge for future models to simulate the effects of lexical and contextual factors on both behavioral and neural data with minimal additional assumptions.

Theoretical issues relating to predictability effects

Eye-tracking and brain-imaging research, together with computer models of reading, provide robust evidence that the predictability of a word has a strong influence during online processing. This section explores several prominent theoretical issues that provide more refined insights into the nature of predictability effects during reading.

What is the best measure of predictability?

One important theoretical question that has implications for understanding the nature of predictability effects is how to best measure subjective predictability as experienced by human comprehenders. This review so far has described studies of reading that operationalize a word's predictability through the measure of cloze probability, which is calculated by aggregating responses provided for a given sentence context in a cloze task (Taylor, 1953). This section evaluates the cloze probability metric in more detail and reviews several alternative computational approaches including transitional probability and the information-theoretic metrics of surprisal and entropy.

Cloze probability

Cloze probability is the favored metric for estimating a word's predictability by virtue of the fact that it is derived from human comprehenders themselves, thereby capturing their expectations of what word will come next in a given sentence context. This means that, unlike more objective approaches, which are influenced by the co-occurrence of words in text, cloze probability provides a subjective but purer measure of what researchers are interested in: how predictable a word is in a given context. There are, however, a few limitations that have been identified about this

metric. Firstly, cloze probability can be noisy and subject to response biases—for instance, failure to follow task instructions or motivational issues may lead some participants to provide overly simple responses and others to produce overly elaborate or unnatural responses. Secondly, it is an open question whether cloze probability reflects a genuine measure of comprehenders' online predictions. Because the cloze task is an offline, untimed procedure, participants are able to engage in conscious reflection and, consequently, strategic task completion compared with online reading during which typical fixations last 200–250 ms. Thirdly, and relatedly, it remains unclear what exactly the cloze task captures. Staub et al. (2015) found that participants generated cloze responses faster when contextual constraint was higher, suggesting that the task captured a race between lexical units to accrue contextual activation rather than the effects of predictability per se. Luke and Christianson (2016) also found that the lexical, positional, and semantic properties of a word influenced participants' cloze responses (see also Smith & Levy, 2013), reflecting the general cognitive constraints of a production-based task, although they noted that the majority of the variance was explained by cloze values over and above these individual predictors. Thus, although cloze probability is arguably the most widely used measure of word predictability, perhaps because of the limitations outlined, some researchers have turned to more objective computational approaches.

Transitional probability

One alternative measure that has been suggested as a potential determinant of a word's predictability is forward *transitional probability* (TP), the conditional probability that word n will occur given word $n - 1$. Unlike cloze probability, this low-level statistical information is derived from text corpora and is independent of high-level contextual information. For example, based on a typical corpus of texts, the verb *accept* is followed by the noun *defeat* more often than the noun *losses*. TP effects were first demonstrated by McDonald and Shillcock (2003a), who observed that high TP nouns yielded shorter first fixation durations than low TP nouns, although this benefit did not extend to gaze duration or skipping (see also McDonald & Shillcock, 2003b, for similar evidence using corpus-based analyses). However, a subsequent study by Frisson et al. (2005) suggested that these apparent effects of TP may actually be effects of cloze probability because, although the nouns presented in McDonald and Shillcock's (2003a) study were very low in cloze probability, high TP nouns were rated as 10 times more predictable than low TP nouns (.08 vs .008). When Frisson et al. tested this possibility by presenting McDonald and Shillcock's verb–noun phrases in neutrally or highly constraining contexts, they found cloze probability effects on the early measures of first

and single fixation duration, in addition to a TP effect on gaze duration. Notably, however, cloze probability varied substantially between the high and low TP nouns in both context conditions, implying that this TP effect was present only when cloze probability had not been controlled. A subsequent experiment in which cloze probability was controlled revealed significant effects of cloze probability but not TP. On the basis of these findings then, word predictability appears to be more accurately captured by cloze probability rather than TP (but see Andrews & Reynolds, 2013; Li et al., 2021).

Surprisal and entropy

Another set of measures that has been used to estimate a word's predictability is *surprisal* and *entropy*, which are derived from statistical (e.g., n -gram, phrase structure grammars, neural networks) or large language models (e.g., masked, autoregressive) trained on large corpora with the objective of predicting the upcoming word of a sequence given prior context. This approach is gaining prominence in the literature because it allows probability estimates to be generated for items across the entire distribution, including those at the lower end that might otherwise be associated with a cloze probability of zero. The first measure, *surprisal*, captures the extent to which a word is unexpected in its context (Hale, 2001; Levy, 2008) and is measured by taking the negative log probability of a word given its prior context: $surprisal(w_i) = -\log p(w_i | w_1 \dots w_{i-1})$. Words with higher surprisal, a lower probability of occurring in a sentence, are harder to process, while words with lower surprisal, a higher probability of occurring in a sentence, are easier to process. Consistent with this account, increasing word surprisal has been linked to longer fixation durations (Amenta et al., 2023; Cevoli et al., 2022; Lowder et al., 2018; Onnis et al., 2022; Smith & Levy, 2013) and larger N400 amplitudes (Aurnhammer & Frank, 2019; Frank et al., 2015; Michaelov et al., 2021). The second measure, *entropy*, captures the degree of uncertainty about how a sentence will unfold (Shannon, 1948) and is measured by taking the negative sum of the probabilities of all outcomes in a sentence, $p(x)$, multiplied by the logarithm of the probabilities of the outcomes: $entropy H(X) = -\sum_{x \in X} p(x) \log_2 p(x)$. Entropy is higher when all possible continuations are probable and lower when there is certainty about the upcoming continuation. Because entropy captures the degree of certainty before a target word is reached, it has also been used to calculate a related metric, *entropy reduction*, $(H_i - H_{i-1})$, which provides an index of the amount of information gained with the identification of each additional word (Hale, 2003, 2006, 2016). Currently, however, behavioral and neural evidence of a relationship between processing effort and either entropy (Aurnhammer

& Frank, 2019; Cevoli et al., 2022; Roark et al., 2009; van Schijndel & Linzen, 2018) or entropy reduction (Frank, 2013; Frank et al., 2015; Linzen & Jaeger, 2016; Lowder et al., 2018; Wu et al., 2010) remains somewhat mixed.

Future directions

Researchers have access to a number of methodological options for measuring a word's predictability in context; however, determining which measure best captures subjective predictability as experienced by human comprehenders remains an ongoing debate. The emergence of language models over the past decade, in particular, has unlocked a new set of tools for explaining processing effort during reading. In light of the increasing use of computational approaches to examine predictability effects, it is important to highlight some specific issues that may require further empirical investigation.

Firstly, in comparison to surprisal, which is the most commonly used computational metric for estimating word predictability, substantially less research has focused on the potential role of entropy and, by extension, entropy reduction. The disproportionate focus is somewhat surprising given that entropy captures the state of the model's predictions *before* a target word is reached, arguably providing a better index of predictability than surprisal, which captures the state of the model's prediction *when* a target word is reached (Aurnhammer & Frank, 2019; Cevoli et al., 2022; Karimi et al., 2024; Pickering & Gambi, 2018). Given that overall evidence of the effect of entropy on processing effort is mixed, however, it would be useful for future research to clarify the link between this metric and different processing effort indices, and whether it makes a contribution to word processing that is distinct from that of surprisal.

Secondly, it remains unclear whether computational metrics based on language models explain variance in processing effort to the same extent as cloze probability derived from human comprehenders. Studies that have explicitly considered this issue have yielded mixed findings: Computational estimates appear to perform worse (Frisson et al., 2005; Smith & Levy, 2011), equivalently (Chandra et al., 2023), or even better (de Varda et al., 2023; Hofmann et al., 2022; Michaelov et al., 2021) than human estimates of predictability. These discrepant outcomes, however, could reflect methodological differences between the studies, including the language model used, the corpus on which the model was trained, the type of analyses conducted, and the processing indices being investigated. Thus, it will be important for future research to consider these technical choices more systematically to assess the predictive power of computational metrics of predictability.

Finally, to the extent that language models can yield reliable estimates of subjective predictability, it also remains

an open debate which models, if any, are analogous to the human comprehension system. Recent findings suggest that statistical language models perform more poorly than large language models in accounting for behavioral (de Varda et al., 2023; Shain et al., 2024) and neural indices of processing effort (Merkx & Frank, 2021; Michaelov et al., 2021), even though, intuitively, the former seem to be more cognitively plausible because of their incremental processing of words and constraints on working memory which mirror that of human comprehenders (Keller, 2010; Merx & Frank, 2021). Even among large language models, smaller models with fewer parameters appear to explain reading behavior better (de Varda et al., 2023; Shain et al., 2024), suggesting that overly powerful models with unlimited access to large sequences of words may overestimate contextual influences on predictability effects compared with human comprehenders (Oh & Schuler, 2023). With the ongoing development and refinement of large language models, it will be important for future research to consider which computational approaches reliably approximate subjective predictability while also being neurocognitively plausible models of language comprehension.

What is the functional relationship between predictability and processing difficulty?

Given the link between predictability and how efficiently a given word is processed, another theoretical question that has been the focus of substantial research is how to quantitatively define this relationship. The simplest account is that the relationship between predictability and processing effort is linear. That is, if predictions about upcoming words are generated according to a probability matching scale, increases in predictability should correspond to decreases in processing effort. An alternative account is that the functional relationship is logarithmic, meaning that the ratio of predictability, rather than its raw difference, determines the processing effort between items—a difference in predictability of .05 compared with .1 should yield the same effect on processing as a predictability difference of .5 compared with 1. The key difference between these two accounts then is how they expect contextual information to affect the processing of very unpredictable words—while logarithmic accounts expect even small differences in low predictability to lead to processing effects, linear accounts expect these effects to be negligible because processing effects are driven primarily by highly predictable words. Determining which function underlies predictability effects is important for understanding whether readers use context to predict upcoming words in proportion to their probability or to preactivate information across the entire lexicon.

Even though linear and logarithmic views of predictability make different predictions about how word processing unfolds, the empirical record on this question is currently mixed. Evidence of a linear predictability relationship comes from Brothers and Kuperberg's (2021) meta-analysis of eye-movement data ($N = 218$), which found that linear models of predictability as estimated by cloze probability provided a better fit to the data than logarithmic models. This was confirmed in two further tasks using SPR and cross-modal picture naming, although these tasks clearly differ from normal reading in many aspects. Other studies using computational estimates report evidence of a logarithmic predictability relationship (Shain et al., 2024; Smith & Levy, 2013; Wilcox et al., 2023; see also Rayner & Well, 1996, for evidence using cloze probability). For example, Smith and Levy (2013) found a logarithmic effect of predictability across six orders of magnitude (i.e., 1 to .000001) when using a tri-gram co-occurrence measure to analyze eye-tracking and self-paced reading measures from naturalistic corpora. Finally, it has been suggested from EEG data that predictability effects could even follow an additive linear and logarithmic function. This superimposed function was first reported by Szewczyk and Federmeier (2022), who observed that the logarithmic function captured variability in the N400 waveform for unexpected words, while the linear function captured additional attenuation in the component for expected words.

Thus, there is no clear consensus on the nature of the functional relationship between predictability and processing effort, although most findings appear to favor a logarithmic account. These mixed outcomes could reflect the impact of two key factors that have differed across investigations of these issues. Firstly, studies have differed in their operationalization of word predictability, with most studies supporting a logarithmic account using computational metrics. As previously mentioned, this approach provides estimates for items across the entire probability distribution, including for words that are unlikely to be produced by human comprehenders, but which are critical to the differentiation of linear and logarithmic effects. Secondly, studies have varied in their use of either constructed or naturalistic language materials. Most studies favoring a logarithmic view have utilized the latter, which, unlike constructed language materials, allow for word-by-word modelling and are therefore able to cover the critical low end of the probability distribution that distinguishes these competing accounts. It should be noted though that naturalistic texts are constrained by their lack of experimental control, allowing words to vary on multiple uncontrolled dimensions that may influence the predictor of interest (Angele et al., 2015; Rayner et al., 2007). As such, the discrepant findings observed across the literature could reflect the methodological choices made by

researchers, which appear to be more conducive to producing logarithmic predictability effects.

Future directions

It remains to be determined whether predictability effects are driven primarily by words at the upper or lower end of the probability distribution. Future investigations of this question should focus on clarifying the extent to which the methodological choices made by researchers contribute to the linking function that relates predictability and processing effort. For example, the reanalysis of existing data using different predictability estimates has the potential to provide valuable insights. Indeed, when Shain (2024) reanalyzed Brothers and Kuperberg's (2021) SPR data using surprisal estimates derived from the large language model GPT-2, they found that the data were consistent with logarithmic over linear predictability effects. More generally, although most studies have focused on measuring processing effort behaviorally as a function of predictability, few studies have examined this issue via neural indices of processing effort. While it is possible that different processing indices capture distinct underlying functions (e.g., Burnsky et al., 2022; Federmeier, 2022; Kretzschmar et al., 2015), additional insights from neural methods could potentially inform this debate of whether predictability effects are linear or logarithmic in nature.

What stage(s) of processing does predictability affect?

Another theoretical question that has interested researchers relates to the stage(s) of processing that is/are affected by effects of predictability. One way in which researchers have explored this issue has been by investigating the effects of predictability on word recognition in sentence contexts. In general, word identification in context can be conceptualized as comprising three stages: (1) the *prelexical processing stage*, which involves any processing that occurs before lexical access, such as the encoding of a word's visual and orthographic features; (2) the *lexical processing* or *lexical access stage*, which involves matching this abstract information to a representation in the mental lexicon; and (3) the *postlexical processing stage*, which involves any processing that occurs after lexical access. Several factors that have been identified as having an impact on distinct stages of word identification are word frequency, parafoveal preview, and stimulus quality. Investigating how these factors interact with contextual information has the potential to provide insights into the stage(s) of processing affected by word predictability. Based on the additive factors logic proposed by Sternberg (1969), interactive effects would suggest that

these variables affect the same stage of processing, while additive effects would suggest that these variables affect different stages of processing (but see McClelland, 1979). This section reviews evidence from studies that have investigated how these factors interact with predictability, with a focus on data from eye-movement and ERP methods.

Predictability and word frequency effects

The frequency of a word, as indexed by how often it occurs in corpora of spoken and/or written texts, is one of the most robust predictors of the speed and accuracy of word identification. High frequency words yield more skips (Drieghe et al., 2005; Rayner & Raney, 1996) and receive shorter first fixation and gaze durations (Inhoff & Rayner, 1986; Rayner & Duffy, 1986) than low frequency words, reflecting the early influence of word frequency on lexical processing. High frequency words also yield smaller N400 components than low frequency words, especially when words are presented in lists (Barber et al., 2004; Grainger et al., 2012; Rugg, 1990) or at the beginning of a sentence (Dambacher et al., 2006; Van Petten & Kutas, 1990). While it is debated whether this ERP component necessarily reflects lexical processing (Kutas & Federmeier, 2011; Nieuwland et al., 2020), frequency effects have also been reported in the earliest N1 time range, a negative-going waveform that occurs approximately 130–200 ms poststimulus onset (Hauk & Pulvermüller, 2004; Sereno et al., 1998). Thus, across different methodologies, evidence of frequency effects has been taken as evidence of lexical access (Sereno & Rayner, 2000, 2003).

Given the robust early influence of word frequency, researchers have gained insights into the temporal locus of predictability effects by investigating whether predictability interacts with frequency effects. Under modular views of language processing (Fodor, 1983; Forster, 1979), which imply that predictability affects a different stage of processing (i.e., postlexical) compared with frequency, these two variables should show additive effects on word recognition based on Sternberg's (1969) additive factors logic. Under interactive views of language processing (McClelland, 1987; Morton, 1969), however, which posit that predictability and frequency affect the same stage of processing, there should be evidence of an early interaction between these two variables. That is, as formalized by the inferential view of probabilistic models (Hale, 2001; Levy, 2008; Norris, 2006), which propose that frequency effects function as predictability effects when context is absent, the predictability effect should be more pronounced for low compared with high frequency words. Computational models of reading also vary in the assumptions they make about the relationship between frequency and predictability. For example, according to the E-Z Reader model discussed earlier, a word's predictability and frequency modulate the rate of lexical processing in

an additive manner (see Eq. 1). However, the relationship between word predictability and frequency is less clear in both SWIFT (Eq. 2) and the CRM (Eq. 3) because, in both models, the rate of lexical processing itself is affected by other variables (e.g., concurrent word processing in SWIFT and inhibition from orthographically similar words in the CRM).

The question of whether word frequency interacts with contextual information during reading has been examined empirically in three ways. The first method involves reaction time measures, which were mainly used in early behavioral studies. For example, Stanovich and West (1981, 1983) used a naming task to test the effects of predictability on pronunciation latencies of high and low frequency words. They found significant main effects of predictability and frequency, as well as an interaction in the direction described above—that is, larger predictability effects for low compared with high frequency words—suggesting that contextual information did interact with word frequency to determine how easily a word was identified. Similar patterns of interaction effects were also reported in studies using lexical decision tasks (Becker, 1979; West & Stanovich, 1982). However, several factors limit the generalizability of these behavioral studies including the presentation method and task-specific demands, which may recruit different strategic processes compared with normal reading.

The next two methods involve studies of eye movements and ERPs grouped by whether they use categorical or continuous manipulations of these variables. Studies that use categorical manipulations rely on constructed language materials that factorially manipulate cloze probability and corpus frequency. Eye-movement studies using this controlled design have reported main effects of each variable on fixation durations but no significant interaction (Ashby et al., 2005; Hand et al., 2010, 2012; Rayner et al., 2001, 2004; Staub, 2020; Staub & Goddard, 2019; but see Sereno et al., 2018). In contrast, ERP studies, have reported interactive effects of both variables on the N400 component, with larger effects of predictability observed for low compared with high frequency words (Dambacher et al., 2006; Sereno et al., 2020; Van Petten & Kutas, 1990; see also Sereno et al., 2020, for evidence on the P1). Thus, there appear to be two distinct patterns of combined predictability-frequency effects on online processing measures, which have implications for determining the temporal locus of predictability. The additive effects observed in the eye-movement record indicate that word predictability may not operate at the same “lexical access” stage of processing as word frequency; however, the opposite is suggested by the interactive effects observed in the ERP record. The findings of Kretzschmar et al.'s (2015) FRP study may provide some insight into this discrepancy: While there were additive effects of both variables on first fixation and gaze duration, there was a predictability effect

on the N400 component, but no frequency effect or interaction. The distinct overall processing patterns may therefore reflect different underlying functions of eye movements and ERPs (see also Burnsky et al., 2022; Federmeier, 2022).

Studies that use continuous manipulations of predictability and frequency, on the other hand, rely on naturalistic language materials, which allow researchers to analyze each word of a sentence or text rather than a single critical region. Early studies which calculated cloze probability for each word found only additive effects of predictability and frequency (Kennedy et al., 2013; Whitford & Titone, 2014). More recent studies using computational metrics based on language models have reported similar findings (Goodkind & Bicknell, 2021; Shain, 2024; but see Shain, 2019). For example, Shain (2024) found that predictability as indexed by surprisal derived from GPT-2 was a significant predictor of reading times from three different modalities (eye movements, SPR, and MAZE³) from six naturalistic corpora, independently of unigram (log-frequency) probability.

In summary, the nature of the relationship between predictability and frequency, as well as the specific stage(s) of processing affected by predictability effects, remain far from clear. The discrepant findings observed across the literature could again reflect the methodological differences including the predictability estimates, language materials, and processing indices being investigated. The following sections discuss other factors that might provide clearer insights into the stage(s) of processing affected by word predictability.

Predictability and parafoveal preview effects

This section has focused on the effects of predictability on word identification that are derived from text that includes, or is prior to, the fixated word. However, readers also extract information from upcoming words in the parafovea.

Parafoveal processing is typically investigated in eye-movement studies using the gaze-contingent *boundary paradigm* in which a target word is replaced by a preview word in its location until the reader's eyes cross an invisible, predefined boundary at the end of the pretarget word (Rayner, 1975). Because readers are not usually aware of this display change due to saccadic suppression (Matin, 1974), the paradigm can be used to infer the types of information extracted parafoveally by comparing reading measures on target words between conditions displaying different types of information in parafoveal vision. Boundary paradigm studies have

revealed a characteristic *parafoveal preview effect*, whereby targets following a valid preview (i.e., identical) compared with invalid preview (e.g., an unrelated word, pseudoword, or nonword) receive facilitated processing—although part of this effect also reflects a preview cost from encountering misleading information parafoveally (Kliegl et al., 2013).⁴ Numerous experiments have converged on the finding that readers are able to preprocess orthographic, phonological, morphological, and some semantic information from upcoming words (see Andrews & Veldre, 2019; Schotter et al., 2012; Vasilev & Angele, 2017, for reviews). Given that parafoveal preview effects are generally interpreted as emerging during prelexical and lexical processing, investigating whether contextual information interacts with parafoveal information can provide insights into the stage(s) of processing affected by word predictability.

Balota et al. (1985) conducted the first eye-movement study to investigate this question. Participants read sentences such as “Since the wedding was today, the baker rushed the wedding cake/pie”, which were completed by either a predictable or unpredictable but plausible word that was preceded by one of five parafoveal previews: valid (i.e., the target word itself, e.g., “cake”); a nonword that was visually similar to this target (e.g., “cahc”); the other possible target word (e.g., “pies”); a nonword that was visually similar to this other target (e.g., “picz”); or an unrelated and visually dissimilar word (e.g., “bomb”). Balota et al. found higher skipping rates for predictable compared with unpredictable words only when following a valid identical preview and, to a lesser extent, an invalid but visually similar preview, suggesting that predictability-based skipping required a preview that at least visually matched the predictable word (but see Drieghe et al., 2005). They also found that the predictability benefit on gaze duration was present only following a preview that was identical or visually similar. This interactive pattern, whereby the predictability effect is eliminated for invalid but not valid previews, has been observed on early reading measures across a number of studies (e.g., Juhasz et al., 2008; Luke, 2018; Staub & Goddard, 2019; Veldre & Andrews, 2018b; White et al., 2005b) and has been taken to suggest that predictability effects depend on the ambiguity of perceptual input during early orthographic processing (Staub & Goddard, 2019; but see Parker et al., 2017; Parker & Slatery, 2019). When readers are denied valid previews, orthographic processing of the target word must unfold entirely in foveal vision, meaning that perceptual information is too

³ The MAZE task (Forster et al., 2009; Freedman & Forster, 1985) is another measure of online sentence processing in which two words are presented at the same time for each upcoming position of a sentence and participants must choose the word that grammatically continues the sentence.

⁴ Although participants are usually unaware of display changes, there is some evidence that display change awareness modulates preview effects, albeit inconsistently (cf. Veldre & Andrews, 2018a; White et al., 2005a). To date, however, there is no evidence that display change awareness affects the relationship between predictability effects and parafoveal processing.

clear for predictability information to have an effect on early processing stages. In contrast, when readers have access to valid previews, orthographic processing of the target word begins before they are directly fixated, based on degraded visual input with the support of contextual information. On the basis of these eye-movement studies, word predictability appears to exert its influence during the earliest parafoveal stages of processing.

The relationship between predictability and parafoveal preview has been investigated to a lesser extent in ERP studies because the RSVP paradigm does not allow for the processing of upcoming words in the parafovea. The *visual hemi-field flanker RSVP paradigm* partly addresses the limitations of RSVP by presenting each word of a sentence centrally, flanked to the left and right by the preceding and following word, respectively (Barber et al., 2011). Studies using this methodology have reported graded N400 predictability effects for words presented in parafoveal vision which are subsequently attenuated when the same words appear foveally (Payne et al., 2019; Stites et al., 2017). Another more ecologically valid approach has been the co-registration of EEG and eye movements under natural reading conditions. Using this methodology, Burnsky et al. (2022) found the expected interaction on early reading measures, that is, predictability effects for valid but not invalid previews, while the same interaction on the N400 was observed for fixations time-locked to the onset of the pretarget, but not target, word. This neural evidence provides further support for the idea that word predictability may exert its effects when words are viewed in parafoveal vision (Staub & Goddard, 2019).

In summary, there appears to be reliable evidence across methodologies that predictability affects the processing of parafoveal information, suggesting that the temporal locus of word predictability is either during prelexical processing or very early lexical processing. If this is the case, predictability would also be expected to interact with factors that affect the earliest stages of letter or word processing, such as stimulus quality. This possibility is considered in the next section.

Predictability and stimulus quality effects

The quality of a stimulus has been shown to influence how long readers spend processing a word (e.g., contrast reduction, Becker & Killion, 1977; dot-pattern degradation, Meyer et al., 1975). One naturalistic way to manipulate stimulus quality is *contrast reduction*, which involves adjusting the contrast between a word and its background. A number of studies have found that words presented in faint text tend to receive longer first fixation and gaze duration than words presented normally (Drieghe, 2008; Reingold & Rayner, 2006; White & Staub, 2012), suggesting that

stimulus quality affects early stages of word identification, during which visual features are encoded and abstract letters are computed (Besner & Roberts, 2003). If it is the case that word predictability modulates an early stage of processing, as implied by its interaction with parafoveal information, predictability should also interact with the quality of the visual stimulus.

The question of whether stimulus quality interacts with higher-order information including predictability has received relatively limited investigation. The earliest relevant findings come from studies using lexical decision which reported that priming effects were larger for degraded compared with intact targets (Balota et al., 2008; Borowsky & Besner, 1993). On the assumption that a semantic prime in a lexical decision task functions like contextual information during normal reading in that both involve the preactivation of general semantic information and/or a specific lexical unit (Staub, 2015), these findings suggest that word predictability may operate at the same stage of processing as stimulus quality. Recent studies investigating this issue using a more typical predictability manipulation have revealed a different set of findings. Staub (2020) recorded participants' eye movements as they read sentences containing target words that were factorially manipulated for predictability and stimulus quality. While they found main effects of both variables, the interaction effect was very small and statistically unreliable, albeit in the expected direction—that is, the predictability effect was larger for faint compared with normal text. A subsequent FRP study conducted by Burnsky et al. (2022) using a similar design replicated these eye-movement effects. In the FRP record, they found an effect of predictability, but no evidence of a stimulus quality effect or an interaction between the two variables. Thus, in the two studies to directly investigate this issue, there appears to be little evidence of the interactive effects that would be expected if predictability affects the same stage of processing as stimulus quality.

Future directions

These findings across three separate literatures demonstrate that it remains unclear which precise stage, or stages, of processing is/are affected by a word's predictability. While it appears that word predictability does not influence lexical processing given that it does not reliably interact with frequency effects, evidence that it influences an earlier stage of processing is mixed given that it interacts with preview validity but not stimulus quality. A tentative conclusion is that predictability effects operate at a very specific early stage of processing when words are viewed in parafoveal vision, which would be consistent with the idea that readers engage in the prediction of upcoming words in advance of their presentation. Further research into the relationship

between predictability and other variables that have been linked to early stages of word or letter processing may provide additional insights (e.g., font difficulty, Rayner et al., 2006; Staub, 2020; letter rotation, Blythe et al., 2019; text mirroring, Chandra et al., 2020).

One final approach to addressing this theoretical question could be through the investigation of the distributional effects of predictability and other factors on parameters of the ex-Gaussian distribution. The ex-Gaussian distribution is a convolution of a normal distribution and an exponential distribution specified by three parameters: μ (the mean of the distribution), σ (the variability of the distribution), and τ (the degree of the rightward skew of the distribution). When fit to individual data, this distribution indicates whether an experimental manipulation impacts processing effort by shifting the distribution and/or changing the skew. Word predictability has been found to impact the μ parameter only in fixation duration distributions (Sheridan & Reingold, 2012; Staub, 2011; Staub & Benatar, 2013) with high predictability words shifting the distribution to the left. Word frequency (Reingold et al., 2012; Staub et al., 2010) and parafoveal preview (White & Staub, 2012) have been found to impact both the μ and τ parameters, while stimulus quality manipulated via contrast reduction has been found to impact the μ parameter (White & Staub, 2012). Although these distributional effects appear to suggest that predictability is more similar to stimulus quality than to frequency or parafoveal preview, these findings should be interpreted with caution, as these effects have not been assessed within a single study (see, e.g., Shain, 2024). Further research into this approach thus has the potential to inform how early effects of predictability arise during reading.

Are predictability effects ubiquitous?

Another ongoing question in the psycholinguistic literature is whether predictability effects are an automatic consequence of online processing. In order to address this issue, it is necessary to investigate whether and how predictability effects interact with factors such as interindividual differences among readers, age, and task demands.

Predictability effects and individual differences

One factor that has been shown to mediate predictability effects during reading is individual differences among readers (Huetting, 2015). In particular, these processes have been linked to language expertise, with a number of studies demonstrating larger predictability effects and/or more efficient use of predictability information in highly- compared with less-skilled readers, as indexed by measures of literacy (Ng et al., 2017; Steen-Baker et al., 2017) and verbal ability (Cheimariou

et al., 2021). Relatedly, studies of non-L1 English adults have revealed smaller predictability effects in L2 compared with L1 speakers (see Ito & Pickering, 2021, for a review). Yet there is concurrent evidence to suggest that, even among L1 speakers, poorer readers, as indexed by their reading speed, show more pronounced predictability effects than better readers (Ashby et al., 2005; Slattery & Yates, 2018; but see Payne & Federmeier, 2019). Poorer readers' impoverished lexical representations (Perfetti, 2007; Perfetti & Hart, 2002) may require them to rely more on contextual information for word identification (Perfetti & Lesgold, 1979; Stanovich, 1984). More generally, predictability effects have been linked to cognitive resources such as cognitive control (Zirnshtein et al., 2018) and attention (Brothers et al., 2017; Hubbard & Federmeier, 2021) among other factors that have typically been investigated using the visual world paradigm (e.g., working memory, Huetting & Janse, 2016; Ito et al., 2018). Thus, the magnitude of predictability effects during online processing appears to vary among readers, and is systematically influenced by individual differences.

Predictability effects and age

Predictability effects during reading have also been shown to interact with age, a factor that lies at the intersection between language skills and cognitive abilities. On the one hand, normal aging is associated with the accumulation of language skills (e.g., vocabulary, Alwin & McCammon, 2001; Verhaeghen, 2003; word-related knowledge, Salthouse, 1993), which implies that older adults should be more sensitive to contextual information compared with younger adults (Payne et al., 2012; Ryskin et al., 2020). On the other hand, however, aging is accompanied by an array of changes in visual (e.g., reduced central and peripheral acuity, general visual pathology; see Owsley, 2011, for a review), and cognitive processes (e.g., lower processing, reduced attention and executive control, and smaller working memory capacity; see Verhaeghen, 2013, for a review), which suggest that older readers should use predictability information less efficiently compared with their younger counterparts. Perhaps because of these competing trajectories, investigations of how aging impacts predictability effects to date have yielded mixed findings.

Eye-movement studies suggest that older readers show equivalent, if not stronger, predictability effects during reading compared with younger readers (Andrews et al., 2022; Cheimariou et al., 2021; Choi et al., 2017; Rayner et al., 2006; Veldre et al., 2022; see Zhang et al., 2022, for a review). This has been argued to reflect an overall *risky reading strategy* according to which older adults rely on contextual information during reading to compensate for declines in visual and cognitive abilities (McGowan & Reichle, 2018; Paterson et al., 2020; Rayner et al., 2006; but see Choi et al., 2017; Veldre et al., 2021, 2022; Zhang et al., 2022). ERP

studies, on the other hand, report the opposite pattern of effects. Compared with younger readers, older adults consistently show either smaller N400 effects, delayed N400 effects, or both (Dave, Brothers, Swaab et al., 2018a, Dave, Brothers, Traxler et al., 2018b; Federmeier & Kutas, 2005; Payne & Federmeier, 2017a; Wlotko et al., 2012; Wlotko & Federmeier, 2012, see Payne & Silcox, 2019; Wlotko et al., 2010, for reviews), suggesting that they are in fact less sensitive to contextual information. With age, there is also limited evidence of the late frontal positivity, which has been taken to index the processing cost of making an incorrect prediction (Dave, Brothers, & Swaab, 2018a; Wlotko et al., 2012), although a subset of older adults with higher verbal ability appear to show evidence of this neural waveform (Dave, Brothers, Traxler et al., 2018b). While these mixed findings may reflect differences in the stimuli presentation rate and format across methodologies, it appears that, at least in samples of older readers, predictability effects may not always be observable during online processing.

Predictability effects and task demands

Another important factor that has been shown to modulate predictability effects during reading is task demands, which have been investigated using several different approaches. One approach has been to use instructions to explicitly manipulate readers' engagement with prediction during reading, for example, by asking participants to read sentences in a passive comprehension block and then in an active prediction block where they are instructed to predict passage-final words and report the accuracy of their predictions. Using this paradigm, a number of studies have found enhanced N400 predictability effects, reflecting greater facilitation for expected input when predictive strategies were encouraged in active compared with passive blocks (Brothers et al., 2017; Dave, Brothers, Traxler, et al., 2018b; Lai et al., 2023). It is unclear, however, whether these graded effects extend to the late frontal positivity, which indexes the cost of an incorrect prediction (see Lai et al., 2023, for a discussion). A second approach used by researchers has been to vary information in readers' broader linguistic environment to implicitly modulate their propensity to use prediction during reading. Using this type of manipulation, Brothers et al. (2017) observed diminished effects of predictability on self-paced reading times when the broader linguistic environment contained a low proportion of confirmed predictions (<20%), suggesting that readers do not always rely on predictive strategies when this strategy is likely to be futile (Kuperberg & Jaeger, 2016). While these studies investigating predictability effects under different task instructions have typically used less natural reading paradigms, they do provide further support for the notion that predictability

effects during online processing may be more complex and multifaceted than previously assumed.

Future directions

Evidence of predictability effects is near ubiquitous during online processing; however, it is likely that these effects also vary across groups and individuals, and under different circumstances. It will be important for future research to investigate whether there are other factors that could potentially mediate predictability effects during reading. For example, language production mechanisms have been implicated in predictive processes during language comprehension (Dell & Chang, 2013; Federmeier, 2007; Pickering & Garrod, 2007), but most studies investigating this link have been in the context of the visual world paradigm and/or correlational in nature (see Huettig, 2015; Pickering & Gambi, 2018, for reviews). It will also be informative for future research to assess whether a language's writing system has an impact on predictability effects. Most research to date has been conducted using alphabetic languages. For instance, consider Chinese, a logographic script written as strings of characters with no spacing cues to demarcate word boundaries, which raises questions about whether and how word predictability information might be used during online processing (e.g., Rayner et al., 2005). Investigations of these issues will provide further insights into whether predictability effects unfold automatically during reading given certain factors and task demands.

What processes do predictability effects actually reflect?

In this article so far, evidence of predictability effects has been taken to index genuine prediction processes. This section reviews this claim in further detail before discussing some alternative processes that may account for predictability effects during reading: graded preactivation, integration, and probabilistic inference. Determining what processes underlie predictability effects is tied to the broader debate about the cognitive architecture of the human language comprehension system.

Lexical prediction processes

The most intuitive account of predictability effects is that they reflect prediction processes—if a word can be predicted in advance of its presentation, then the way in which this word is processed when eventually encountered may reflect its level of predictability (Kutas et al., 2011). Most psycholinguistic research to date has explicitly, if not implicitly, defined prediction as the “all-or-none process

of activating a linguistic term (a word) in advance of perceptual input” (DeLong, Troyer et al., 2014b, p. 632). This predictive process, termed *lexical prediction*, is assumed to involve the prediction of a specific lexical candidate, which is accompanied by processing benefits if readers’ predictions turn out to be correct but processing costs if they turn out to be incorrect. Most evidence of predictability benefits outlined so far in this article could be taken as support for the first half of this assertion, even if most naturally occurring text is neither predictive nor constraining (Luke & Christianson, 2016). However, support for the second half of this assertion—the processing costs when unpredictable words violate a more expected completion in their context—remains more elusive. Eye-movement studies have found minimal evidence of prediction error costs across various studies using constructed (Frisson et al., 2017; Wong et al., 2022; but see Wong et al., 2024) and naturalistic language materials (Andrews et al., 2022; Luke & Christianson, 2016; but see Cevoli et al., 2022). In contrast, ERP studies have reported evidence of enhanced neural activity for prediction violations in the form of a positivity, approximately 600–900 ms poststimulus, with an anterior scalp distribution (Brothers et al., 2017, 2020; DeLong et al., 2011; DeLong, Quante, et al., 2014a; Federmeier et al., 2007; Hubbard & Federmeier, 2023; Hubbard et al., 2019; Kuperberg et al., 2020; Kutas, 1993; Lai et al., 2021, 2023; Rommers & Federmeier, 2018a; Thornhill & Van Petten, 2012; see Van Petten & Luka, 2012, for a review). This neural component has been distinguished from a positivity with a parietal scalp distribution that occurs in the same time frame for unpredictable words that are anomalous in strongly constraining contexts (Brothers et al., 2020; DeLong, Quante, et al., 2014a; Kuperberg et al., 2020; Thornhill & Van Petten, 2012; see Van Petten & Luka, 2012, for a review). Although the late frontal positivity appears relatively late in the time course of normal reading, one hypothesis is that it reflects higher-order processes related to the suppression or inhibition of the incorrectly predicted word (Federmeier et al., 2007; Kutas, 1993; Ness & Meltzer-Asscher, 2018). However, because there is also evidence of this ERP waveform for unexpected input in weakly to moderately constraining contexts where the most expected completion is unlikely to have been preactivated (Brothers et al., 2015; Davenport & Coulson, 2011; Hubbard et al., 2019; Thornhill & Van Petten, 2012), another hypothesis is that it also reflects integration of the unexpected input and/or updating of the broader discourse representation (Brothers et al., 2015; DeLong, Quante, et al., 2014a; Kuperberg et al., 2020). While this discrepancy in observing prediction error costs could in part be attributed to the differing methodologies and processing indices, the absence of conclusive evidence of the processing costs that would be expected to

accompany incorrect predictions remains a challenge to the argument that readers make use of lexical prediction during reading.

Graded preactivation processes

Another account that has emerged in more recent years is that predictability effects do not involve specific lexical predictions but rather the partial preactivation of upcoming words (Brothers & Kuperberg, 2021; Federmeier, 2022; Luke & Christianson, 2016; Staub, 2015; Staub et al., 2015). This predictive process, termed *graded prediction*, is assumed to involve the passive activation of information related to the current discourse as part of the natural organization of long-term memory in response to incoming linguistic input (Federmeier, 2022). As such, multiple lexical candidates can be generated for each upcoming word of a sentence without incurring any processing costs if they turn out to be incorrect. This type of prediction is accounted for by the evidence of predictability benefits considered so far in this article, as well as several other findings. Firstly, a number of studies have shown that readers preactivate relevant information about upcoming words, including at the syntactic level (Cutter et al., 2022; Dikker et al., 2010; Lau et al., 2006; Staub & Clifton, 2006; see Ferreira & Qiu, 2021, for a review) and orthographic and phonological levels (DeLong et al., 2005; Ito et al., 2016; Laszlo & Federmeier, 2009; Luke & Christianson, 2012; Martin et al., 2013). Secondly, as described above, few studies have found processing costs for incorrect predictions and have instead observed facilitated processing for unpredictable words when semantically related to the best completion (Andrews et al., 2022; Federmeier & Kutas, 1999; Frisson et al., 2017; Luke & Christianson, 2016; Thornhill & Van Petten, 2012; Wlotko & Federmeier, 2015; Wong et al., 2022, 2024). Thus, there is growing evidence to support the notion that readers engage in the graded prediction of upcoming words, even if their full lexical identity cannot be predicted from prior context.

Although graded prediction differs in its assumptions about how language processing should unfold compared with lexical prediction, it is unlikely these two accounts reflect two completely distinct processes. An emerging proposal is that prediction is a dynamic process involving the activation of multiple sources of information over different time courses (Burnsky et al., 2022; Federmeier, 2022; Huettig, 2015; Pickering & Gambi, 2018; Szewczyk & Federmeier, 2022). For example, Federmeier (2022) posits two processing modes for language comprehension: *connecting*, which involves linking incoming linguistic input with long-term semantic memory to activate information in a graded and parallel manner, and *considering*, which involves using comprehension strategies like prediction to transform initial graded semantic representations into

more stable multidimensional representations. Pickering and Gambi (2018) similarly propose a two-systems account of prediction: *prediction-by-association*, which is characterized by spreading activation between related concepts (Collins & Loftus, 1975; Neely, 1977), and *prediction-by-production*, which involves using the production system to covertly anticipate upcoming linguistic content (see also Dell & Chang, 2013; Huettig, 2015; Pickering & Garrod, 2004, 2013). Thus, under these dual-systems accounts of prediction, the initial stage of (graded) preactivation is passive but obligatory, which prepares the language processor for the subsequent active but optional stage of (lexical) prediction. This second stage is nonobligatory because it can depend on readers having sufficient time (e.g., slower presentation rates, Ito et al., 2016; Wlotko & Federmeier, 2015) and resources (e.g., language skills, cognitive abilities, and age). In other words, predictive processes occur along a continuum: At any given point in a sentence, readers may either preactivate general semantic information or, with the availability of time and resources, activate a more specific word form.

Postlexical integration processes

Another account of predictability effects is that they reflect *postlexical integration processes*, that is, when a comprehender combines linguistic information that is activated as a result of processing the current input with a representation of the preceding input (Kutas et al., 2011; Pickering & Gambi, 2018; Van Petten & Luka, 2012). Consider again the first example presented in this article: “The day was breezy so the boy went outside to fly a . . .” for which the high cloze completion “kite” is processed most efficiently. According to the prediction view, “kite” receives facilitated processing because comprehenders are more likely to have activated it in advance of its presentation, making it easier to process than a low cloze completion like “airplane.” According to the integration view, however, “kite” is still easier to process even if it has not been predicted because comprehenders are more likely to have activated linguistic information pertaining to “kite” based on the prior context, but importantly, not the lexical item itself. As such, when “kite” is encountered, it is easier to integrate by combining the word’s meaning with the representation of the preceding input.

While these accounts of predictability effects can be distinguished theoretically, it is difficult to find empirical evidence that is compatible with prediction but not integration processes. As discussed in the section Prediction in language comprehension, there are a handful of EEG studies that demonstrate clear evidence of prediction by revealing that a word has been activated even before it has been encountered by the comprehender (DeLong et al., 2005, 2012; Martin et al., 2013). However, these studies are in the minority due to

the difficulty of designing such manipulations in English, and because some of these effects have failed to replicate. As such, evidence of facilitated processing for predictable words across eye-tracking and brain-imaging methodologies could be compatible with an integration account given that they involve measuring online processes that occur at, and not before, the critical target. More generally, the fact that predictability affects “early” eye-movement measures such as skipping and first fixation duration does not provide unequivocal evidence for an early locus of predictability effects because a substantial amount of lexical processing can occur when the word is in the parafovea (see Veldre et al., 2020, for E-Z Reader simulations demonstrating that skipping effects can be attributed to postlexical integration). Indeed, even if predictability effects capture genuine prediction processes, the fact that some of these effects emerge on relatively late measures suggests that predictability effects may *also* have some postlexical impact. Given the key contributions of both prediction and integration processes, it is plausible that both are involved in the manifestation of predictability effects during reading (Brouwer et al., 2021; Ferreira & Chantavarin, 2018; Onnis et al., 2022).

Probabilistic inference processes

One final account of predictability effects is that they reflect a processing cost, specifically, the cost of *probabilistic inference*. This view derives from information theory (Hale, 2001; Levy, 2008), which, as described in section Sentence-processing models, posits that the primary role of the language processor is to assign a probability distribution over all possible continuations of a sentence. Under this account, processes such as prediction play no role in explaining predictability effects during reading because these effects, as indexed by surprisal, are simply a natural consequence of the probability distribution being updated with each upcoming word. In other words, in contrast to the view of sentence processing adopted by this article so far, which focuses on understanding how prediction contributes to building a mental representation of language structure and meaning (see Shain, 2024), the inferential view of sentence processing focuses only on the problem of how the language processor computes probabilistic inferences given limited evidence. Empirical support for this account comes from findings that, firstly, increasing surprisal has been linked to more effortful behavioral and neural processing (see section Surprisal and entropy); secondly, the relationship between predictability and processing effort appears to be logarithmic (see section What is the functional relationship between predictability and processing difficulty?); and thirdly, there is no specific cost when an unpredictable word is encountered in a context that predicts another completion (see section Graded preactivation processes; see Staub, 2024, for a review). At the same time, however, the findings that,

firstly, predictability does not interact with frequency and, secondly, these two factors are dissociable across methodologies (see section Predictability and word frequency effects) remain a challenge to the inferential account because information theory hypothesizes that frequency effects function as predictability effects when context is absent (Hale, 2001; Levy, 2008; Norris, 2006). Thus, while probabilistic inference may provide an explanation for predictability effects to some extent, there appear to be limits on the scope of this account of cognitive processing. Indeed, this account does not deal with limitations in human attention and memory, which might lead to imperfect or “lossy” access to the linguistic context required by the language processor to make probabilistic inferences (Futrell et al., 2020; Hahn et al., 2022).

Future directions

Although the most intuitive explanation of predictability effects is that they reflect lexical prediction processes, this account cannot be definitively distinguished from processes of graded preactivation, integration, or probabilistic inference, which also predict processing benefits for predictable words in context. Indeed, as suggested throughout this section, it is likely that predictability effects reflect some combination of these processes. It will thus be important for ongoing research to clarify the processes that underlie predictability effects during reading. For example, the link between lexical prediction and graded preactivation processes could be strengthened by examining the factors and time course that underpin how readers transition from the partial preactivation to the specific prediction of upcoming words. These prediction-based processes may also be disentangled from general integration processes by using computational estimates generated before a target word is reached, which must reflect prediction processes, compared with when a target word is reached, which is more likely to capture some aspect of integration processes. Finally, it will also be a goal for future research to reconcile the underlying assumptions of differing views of sentence processing that have been used to explain predictability effects in order to provide a more integrated theory of language comprehension.

Summary, conclusions, and future directions

This article provides an overview of the current literature relating to prediction processes during real-time language comprehension, with a focus on evidence from predictability effects during reading. The overall picture that emerges based on eye-tracking and brain-imaging research to date is that, to some extent, effects of predictability can be taken as demonstrating evidence that the language processor engages in genuine linguistic prediction during online processing. In particular, the investigation of several prominent theoretical

issues lends support to the emerging view that prediction does not always entail the activation of a specific lexical item, but rather graded preactivation of multiple sources of information. Moreover, predictability effects do not solely reflect integration processes. Rather, they appear to operate at a very specific early stage of processing, possibly even before a word has been fixated. However, there is also accumulating evidence to suggest that factors like individual differences among readers, age, and different task demands can modulate readers’ use of prediction to varying degrees. Taken together then, readers’ ability to anticipate upcoming information does play a helping role in online processing; however, evidence that, in its strongest form, prediction depends on the availability of resources and time suggests that it may not always be absolutely necessary for successful language comprehension.

As mentioned throughout this article, however, there remain several unresolved theoretical issues which are important for understanding how prediction allows real-time language comprehension to unfold as rapidly and effortlessly as it does. The most important questions are summarized below:

1. Do computational estimates of predictability provide a viable alternative to cloze probability estimates in capturing the predictions of human comprehenders?
2. Is there compelling evidence that the effects of predictability arise during early stages of processing or even before the critical word has been encountered which would provide the clearest demonstration of prediction?
3. Is there evidence from factors including, but not limited to, individual differences, age, and task demands that prediction is a flexible strategy that depends on the availability of resources and time?
4. How can the processes underlying predictability effects be understood within theories of language comprehension that incorporate inferential views of sentence processing?
5. How can computer models of reading be used to account for how predictability facilitates the reading of text at both behavioral and neural levels?

Answering these questions will be key to progressing the literature from investigating whether prediction is an important component of real-time language comprehension towards understanding the cognitive architecture and mechanisms that support these computations.

One approach, in particular, that will facilitate future research into prediction processes during real-time language comprehension is the ongoing advancement of underutilized methodologies such as the co-registration of eye-movement and brain activity data. The recording of continuous brain activity under natural reading conditions affords a more comprehensive assessment of the behavioral and neural correlates underpinning online cognitive processing. Although these

techniques face significant methodological and analytical challenges (see Degno et al., 2021; Himmelstoss et al., 2020, for reviews), they have the potential to contribute to resolving the issues outlined above and to reconcile the apparent discrepancies in existing conclusions of eye-tracking and brain-imaging research. Co-registration methods will also play a role in the development of computer models of reading as they attempt to simulate the time course of lexical and contextual factors on both behavioral and neural data with minimal additional assumptions.

To conclude, linguistic prediction during online processing is consistent with the growing realization that the brain constantly functions as a “prediction machine,” as posited by general predictive accounts of cognitive functioning (Clark, 2013; Friston, 2010). Although some aspects of how these processes unfold are still being investigated, debated, and refined, it is increasingly apparent that anticipatory prediction should no longer be considered a peripheral question in the literature but rather understood as a natural part of the way language unfolds.

Author note We are grateful to Sally Andrews for her enduring contributions to this research and to studies of reading.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions This research was supported under Australian Research Council’s *Discovery Projects* funding scheme (project numbers DP18102705, DP190100719) and The University of Sydney’s Postgraduate Research Support Scheme.

Availability of data and materials Not applicable.

Code availability Not applicable.

Declarations

Conflicts of interest/competing interests Not applicable.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abbott, M. J., Angele, B., Ahn, Y. D., & Rayner, K. (2015). Skipping syntactically illegal the previews: The role of predictability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(6), 1703–1714. <https://doi.org/10.1037/xlm0000142>
- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264. [https://doi.org/10.1016/S0010-0277\(99\)00059-1](https://doi.org/10.1016/S0010-0277(99)00059-1)
- Alwin, D. F., & McCammon, R. J. (2001). Aging, cohorts, and verbal ability. *The Journals of Gerontology: Series B*, 56(3), S151–161. <https://doi.org/10.1093/geronb/56.3.s151>
- Amenta, S., Hasenäcker, J., Crepaldi, D., & Marelli, M. (2023). Prediction at the intersection of sentence context and word form: Evidence from eye-movements and self-paced reading. *Psychonomic Bulletin & Review*, 30(3), 1081–1092. <https://doi.org/10.3758/s13423-022-02223-9>
- Andrews, S., & Reynolds, G. (2013). Why it is easier to wreak havoc than unleash havoc: The Role of lexical co-occurrence, predictability and reading proficiency in sentence reading. In M. A. Britt, S. R. Goldman, & J.-F. Rouet (Eds.), *Reading—From words to multiple texts* (pp. 72–91). Routledge.
- Andrews, S., & Veldre, A. (2019). What is the most plausible account of the role of parafoveal processing in reading? *Language and Linguistics Compass*, 13(7), e12344. <https://doi.org/10.1111/lnc3.12344>
- Andrews, S., Veldre, A., Wong, R., Yu, L., & Reichle, E. D. (2022). How do task demands and aging affect lexical prediction during online reading of natural texts? *Journal of Experimental Psychology: Learning, Memory, and Cognition*. <https://doi.org/10.1037/xlm0001200>
- Angele, B., Schotter, E. R., Slattery, T. J., Tenenbaum, T. L., Bicknell, K., & Rayner, K. (2015). Do successor effects in reading reflect lexical parafoveal processing? Evidence from corpus-based and experimental eye movement data. *Journal of Memory and Language*, 79(80), 76–96. <https://doi.org/10.1016/j.jml.2014.11.003>
- Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review*, 105(4), 678–723. <https://doi.org/10.1037/0033-295X.105.4.678-723>
- Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology: A, Human Experimental Psychology*, 58(6), 1065–1086. <https://doi.org/10.1080/02724980443000476>
- Aurnhammer, C., & Frank, S. L. (2019). Evaluating information-theoretic measures of word prediction in naturalistic sentence reading. *Neuropsychologia*, 134, 107198–107198. <https://doi.org/10.1016/j.neuropsychologia.2019.107198>
- Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, 17(3), 364–390. [https://doi.org/10.1016/0010-0285\(85\)90013-1](https://doi.org/10.1016/0010-0285(85)90013-1)
- Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language*, 59(4), 495–523. <https://doi.org/10.1016/j.jml.2007.10.004>
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280–289. <https://doi.org/10.1016/j.tics.2007.05.005>
- Barber, H. A., Ben-Zvi, S., Bentin, S., & Kutas, M. (2011). Parafoveal perception during sentence reading? An ERP paradigm using rapid serial visual presentation (RSVP) with flankers.

- Psychophysiology*, 48(4), 523–531. <https://doi.org/10.1111/j.1469-8986.2010.01082.x>
- Barber, H. A., Otten, L. J., Kousta, S.-T., & Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain and Language*, 125(1), 47–53. <https://doi.org/10.1016/j.bandl.2013.01.005>
- Barber, H., Vergara, M., & Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: Evidence from ERPs. *Neuroreport*, 15(3), 545–548. <https://doi.org/10.1097/00001756-200403010-00032>
- Baumgaertner, A., Weiller, C., & Büchel, C. (2002). Event-related fMRI reveals cortical sites involved in contextual sentence integration. *NeuroImage*, 16(3), 736–745. <https://doi.org/10.1006/nimg.2002.1134>
- Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental psychology: Human Perception and Performance*, 5(2), 252–259. <https://doi.org/10.1037/0096-1523.5.2.252>
- Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 389–401.
- Besner, D., & Roberts, M. A. (2003). Reading nonwords aloud: Results requiring change in the dual route cascaded model. *Psychonomic Bulletin & Review*, 10(2), 398–404. <https://doi.org/10.3758/BF03196498>
- Bloom, P. A., & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory & Cognition*, 8(6), 631–642. <https://doi.org/10.3758/bf03213783>
- Blythe, H. I., Juhasz, B. J., Tbaily, L. W., Rayner, K., & Liversedge, S. P. (2019). Reading sentences of words with rotated letters: An eye movement study. *Quarterly Journal of Experimental Psychology (Hove)*, 72(7), 1790–1804. <https://doi.org/10.1177/1747021818810381>
- Borowsky, R., & Besner, D. (1993). Visual word recognition: A multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(4), 813–840. <https://doi.org/10.1037/0278-7393.19.4.813>
- Brothers, T., & Kuperberg, G. R. (2021). Word predictability effects are linear, not logarithmic: Implications for probabilistic models of sentence comprehension. *Journal of Memory and Language*, 116, 104174. <https://doi.org/10.1016/j.jml.2020.104174>
- Brothers, T., Swaab, T. Y., & Traxler, M. J. (2015). Effects of prediction and contextual support on lexical processing: Prediction takes precedence. *Cognition*, 136, 135–149. <https://doi.org/10.1016/j.cognition.2014.10.017>
- 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. <https://doi.org/10.1016/j.jml.2016.10.002>
- Brothers, T., Wlotko, E. W., Warnke, L., & Kuperberg, G. R. (2020). Going the extra mile: Effects of discourse context on two late positivities during language comprehension. *Neurobiology of Language*, 1(1), 135–160. https://doi.org/10.1162/nol_a_00006
- Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. J. (2017). A neurocomputational model of the N400 and the P600 in language processing. *Cognitive Science*, 41, 1318–1352. <https://doi.org/10.1111/cogs.12461>
- Brouwer, H., Delogu, F., Venhuizen, N. J., & Crocker, M. W. (2021). Neurobehavioral correlates of surprisal in language comprehension: A neurocomputational model. *Frontiers in Psychology*, 12, 615538–615538. <https://doi.org/10.3389/fpsyg.2021.615538>
- Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 53–78). Oxford University Press.
- Burnsky, J., Kretschmar, F., Mayer, E., & Staub, A. (2022). The influence of predictability, visual contrast, and preview validity on eye movements and N400 amplitude: Coregistration evidence that the N400 reflects late processes. *Language, Cognition and Neuroscience*. <https://doi.org/10.1080/23273798.2022.2159990>. ahead-of-print. Advance online publication.
- Carr, T. H., & Pollatsek, A. (1985). Recognizing printed words: A look at current models. In D. Besner, T. G. Waller, & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (pp. 1–82). Academic Press.
- Carter, B. T., Foster, B., Muncy, N. M., & Luke, S. G. (2019). Linguistic networks associated with lexical, semantic and syntactic predictability in reading: A fixation-related fMRI study. *NeuroImage*, 189, 224–240. <https://doi.org/10.1016/j.neuroimage.2019.01.018>
- Cevoli, B., Watkins, C., & Rastle, K. (2022). Prediction as a basis for skilled reading: Insights from modern language models. *Royal Society Open Science*, 9(6), 211837. <https://doi.org/10.1098/rsos.211837>
- Chandra, J., Krügel, A., & Engbert, R. (2020). Modulation of oculomotor control during reading of mirrored and inverted texts. *Scientific Report*, 10(1), 4210. <https://doi.org/10.1038/s41598-020-60833-6>
- Chandra, J., Witzig, N., & Laubrock, J. (2023). Synthetic predictabilities from large language models explain reading eye movements. *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications* (19th ed., pp. 1–7). ACM.
- Cheimariou, S., Farmer, T. A., & Gordon, J. K. (2021). The effects of age and verbal ability on word predictability in reading. *Psychology and Aging*, 36(4), 531–542. <https://doi.org/10.1037/pag0000609>
- Cheyette, S. J., & Plaut, D. C. (2017). Modeling the N400 ERP component as transient semantic over-activation within a neural network model of word comprehension. *Cognition*, 162, 153–166. <https://doi.org/10.1016/j.cognition.2016.10.016>
- Choi, W., Lowder, M. W., Ferreira, F., Swaab, T. Y., & Henderson, J. M. (2017). Effects of word predictability and preview lexicality on eye movements during reading: A comparison between young and older adults. *Psychology and Aging*, 32(3), 232–242. <https://doi.org/10.1037/pag0000160>
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences*, 36(3), 181–204. <https://doi.org/10.1017/S0140525X12000477>
- Clifton, C., Jr., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 341–371). Elsevier. <https://doi.org/10.1016/B978-008044980-7/50017-3>
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407–428. <https://doi.org/10.1037/0033-295X.82.6.407>
- Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151–216). Academic Press.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204–256. <https://doi.org/10.1037/0033-295X.108.1.204>
- Cutter, M. G., Martin, A. E., & Sturt, P. (2020). The activation of contextually predictable words in syntactically illegal positions. *Quarterly Journal of Experimental Psychology*, 73(9), 1423–1430. <https://doi.org/10.1177/1747021820911021>
- Cutter, M. G., Paterson, K. B., & Filik, R. (2022). Syntactic prediction during self-paced reading is age invariant. *The British Journal of Psychology*. <https://doi.org/10.1111/bjop.12594>

- Dambacher, M., Kliegl, R., Hofmann, M., & Jacobs, A. M. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain Research*, 1084(1), 89–103. <https://doi.org/10.1016/j.brainres.2006.02.010>
- Dambacher, Dimigen, O., Braun, M., Wille, K., Jacobs, A. M., & Kliegl, R. (2012). Stimulus onset asynchrony and the timeline of word recognition: Event-related potentials during sentence reading. *Neuropsychologia*, 50(8), 1852–1870. <https://doi.org/10.1016/j.neuropsychologia.2012.04.011>
- Dave, S., Brothers, T. A., & Swaab, T. Y. (2018a). 1/f neural noise and electrophysiological indices of contextual prediction in aging. *Brain Research*, 1691, 34–43. <https://doi.org/10.1016/j.brainres.2018.04.007>
- Dave, S., Brothers, T. A., Traxler, M. J., Ferreira, F., Henderson, J. M., & Swaab, T. Y. (2018b). Electrophysiological evidence for preserved primacy of lexical prediction in aging. *Neuropsychologia*, 117, 135–147. <https://doi.org/10.1016/j.neuropsychologia.2018.05.023>
- Davenport, T., & Coulson, S. (2011). Predictability and novelty in literal language comprehension: An ERP study. *Brain Research*, 1418, 70–82. <https://doi.org/10.1016/j.brainres.2011.07.039>
- Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, 117(3), 713–758. <https://doi.org/10.1037/a0019738>
- de Varda, A. G., Marelli, M., & Amenta, S. (2023). Cloze probability, predictability ratings, and computational estimates for 205 English sentences, aligned with existing EEG and reading time data. *Behavior Research*, 56, 5190–5213. <https://doi.org/10.3758/s13428-023-02261-8>
- Degno, F., Loberg, O., & Liversedge, S. P. (2021). Coregistration of eye movements and fixation-related potentials in natural reading: Practical issues of experimental design and data analysis. *Collabra. Psychology*, 7(1), 18032. <https://doi.org/10.1525/collabra.18032>
- Dell, G. S., & Chang, F. (2013). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions. Biological Sciences*, 369(1634), 20120394. <https://doi.org/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(8), 1117–1121. <https://doi.org/10.1038/nn1504>
- DeLong, K. A., Urbach, T. P., Groppe, D. M., & Kutas, M. (2011). Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology*, 48(9), 1203–1207. <https://doi.org/10.1111/j.1469-8986.2011.01199.x>
- DeLong, K. A., Groppe, D. M., Urbach, T. P., & Kutas, M. (2012). Thinking ahead or not? Natural aging and anticipation during reading. *Brain and Language*, 121(3), 226–239. <https://doi.org/10.1016/j.bandl.2012.02.006>
- DeLong, K. A., Quante, L., & Kutas, M. (2014a). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia*, 61, 150–162. <https://doi.org/10.1016/j.neuropsychologia.2014.06.016>
- DeLong, K. A., Troyer, M., & Kutas, M. (2014b). Pre-processing in sentence comprehension: Sensitivity to likely upcoming meaning and structure. *Language and Linguistics Compass*, 8(12), 631–645. <https://doi.org/10.1111/lnc3.12093>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 conference of the North American Chapter of the Association for Computational Linguistics: Human language technologies* (vol. 1, pp. 4171–4186). Long and Short Papers.
- Dien, J., Franklin, M. S., Michelson, C. A., Lemen, L. C., Adams, C. L., & Kiehl, K. A. (2008). fMRI characterization of the language formulation area. *Brain Research*, 1229, 179–192. <https://doi.org/10.1016/j.brainres.2008.06.107>
- Dikker, S., Rabagliati, H., Farmer, T. A., & Pykkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science*, 21(5), 629–634. <https://doi.org/10.1177/0956797610367751>
- Dimigen, O., Sommer, W., Hohlfield, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: Analyses and review. *Journal of Experimental Psychology: General*, 140(4), 552–572. <https://doi.org/10.1037/a0023885>
- Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology*, 44(6), 927–935. <https://doi.org/10.1111/j.1469-8986.2007.00593.x>
- Drieghe, D. (2008). Foveal processing and word skipping during reading. *Psychonomic Bulletin & Review*, 15, 856–860. <https://doi.org/10.3758/PBR.15.4.856>
- Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental psychology Human Perception and Performance*, 31(5), 954–969. <https://doi.org/10.1037/0096-1523.31.5.954>
- Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, 20(6), 641–655. [https://doi.org/10.1016/S0022-5371\(81\)90220-6](https://doi.org/10.1016/S0022-5371(81)90220-6)
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179–211. https://doi.org/10.1207/s15516709cog1402_1
- Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, 42(5), 621–636. [https://doi.org/10.1016/S0042-6989\(01\)00301-7](https://doi.org/10.1016/S0042-6989(01)00301-7)
- Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, 112(4), 777–813. <https://doi.org/10.1037/0033-295x.112.4.777>
- Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44, 491–505. <https://doi.org/10.1111/j.1469-8986.2007.00531.x>
- Federmeier, K. D. (2022). Connecting and considering: Electrophysiology provides insights into comprehension. *Psychophysiology*, 59(1), 1–32. <https://doi.org/10.1111/psyp.13940>
- Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, 41(4), 469–495. <https://doi.org/10.1006/jmla.1999.2660>
- Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, 42(2), 133–141. <https://doi.org/10.1111/j.1469-8986.2005.00274.x>
- Federmeier, K. D., Kutas, M., & Dickson, D. S. (2016). A common neural progression to meaning in about a third of a second. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 557–567). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-407794-2.00045-6>
- Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, 1146, 75–84. <https://doi.org/10.1016/j.brainres.2006.06.101>
- Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science*, 27(6), 443–448. <https://doi.org/10.1177/0963721418794491>
- Ferreira, F., & Lowder, M. W. (2016). Prediction, information structure, and good-enough language processing. In B. H. Ross (Ed.), *Psychology of learning and motivation* (vol. 65, pp.

- 217–247). Academic Press. <https://doi.org/10.1016/bs.plm.2016.04.002>
- Ferreira, F., & Qiu, Z. (2021). Predicting syntactic structure. *Brain Research*, 1770, 147632–147632. <https://doi.org/10.1016/j.brainres.2021.147632>
- Fischler, I. S., & Bloom, P. A. (1985). Effects of constraint and validity of sentence contexts on lexical decisions. *Memory & Cognition*, 13(2), 128–139. <https://doi.org/10.3758/BF03197005>
- Fitz, H., & Chang, F. (2019). Language ERPs reflect learning through prediction error propagation. *Cognitive Psychology*, 111, 15–52. <https://doi.org/10.1016/j.cogpsych.2019.03.002>
- Fitzsimmons, G., & Drieghe, D. (2013). How fast can predictability influence word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(4), 1054–1063. <https://doi.org/10.1037/a0030909>
- Flanagan, J. R., & Wing, A. M. (1997). Effects of surface texture and grip force on the discrimination of hand-held loads. *Perception & Psychophysics*, 59(1), 111–118. <https://doi.org/10.3758/BF03206853>
- Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and Language*, 27, 235–244. [https://doi.org/10.1016/0749-596X\(88\)90052-6](https://doi.org/10.1016/0749-596X(88)90052-6)
- Fleur, D. S., Flecken, M., Rommers, J., & Nieuwland, M. S. (2020). Definitely saw it coming? The dual nature of the pre-nominal prediction effect. *Cognition*, 204, 104335–104335. <https://doi.org/10.1016/j.cognition.2020.104335>
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT Press.
- Forster, K. I. (1976). Accessing the mental lexicon. In E. C. T. Walker & R. J. Wales (Eds.), *New approaches to language mechanisms* (pp. 257–287). North Holland.
- Forster, K. I. (1979). Basic issues in lexical processing. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process*. MIT Press.
- Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *The Quarterly Journal of Experimental Psychology Section A*, 33(4), 465–495. <https://doi.org/10.1080/14640748108400804>
- Forster, K. I., Guerreria, C., & Elliot, L. (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods*, 41(1), 163–171. <https://doi.org/10.3758/BRM.41.1.163>
- Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(5), 1461–1469. <https://doi.org/10.1037/a0036756>
- Frank, S. L. (2013). Uncertainty reduction as a measure of cognitive load in sentence comprehension. *Topics in Cognitive Science*, 5(3), 475–494. <https://doi.org/10.1111/tops.12025>
- Frank, S. L., Koppen, M., Noordman, L. G. M., & Vonk, W. (2003). Modeling knowledge-based inferences in story comprehension. *Cognitive Science*, 27(6), 875–910. <https://doi.org/10.1016/j.cogsci.2003.07.002>
- Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain and Language*, 140, 1–11. <https://doi.org/10.1016/j.bandl.2014.10.006>
- Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies [Doctoral dissertation]*. University of Connecticut.
- Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6, 291–325. [https://doi.org/10.1016/0010-0277\(78\)90002-1](https://doi.org/10.1016/0010-0277(78)90002-1)
- Freedman, S. E., & Forster, K. I. (1985). The psychological status of overgenerated sentences. *Cognition*, 19, 101–131. [https://doi.org/10.1016/0010-0277\(85\)90015-0](https://doi.org/10.1016/0010-0277(85)90015-0)
- Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics*, 33(6), 575–581. <https://doi.org/10.3758/BF03202940>
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138. <https://doi.org/10.1038/nrn2787>
- Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(5), 862–877. <https://doi.org/10.1037/0278-7393.31.5.862>
- Frisson, S., Harvey, D. R., & Staub, A. (2017). No prediction error cost in reading: Evidence from eye movements. *Journal of Memory and Language*, 95, 200–214. <https://doi.org/10.1016/j.jml.2017.04.007>
- Frith, C. D., & Frith, U. (2006). How we predict what other people are going to do. *Brain Research*, 1079(1), 36–46. <https://doi.org/10.1016/j.brainres.2005.12.126>
- Futrell, R., Gibson, E., & Levy, R. P. (2020). Lossy-context surprisal: An information-theoretic model of memory effects in sentence processing. *Cognitive Science*, 44(3), e12814. <https://doi.org/10.1111/cogs.12814>
- Glushko, R. J. (1979). The organization and activation of orthographic knowledge and reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 674–691. <https://doi.org/10.1037/0096-1523.5.4.674>
- Golden, R. M., & Rumelhart, D. E. (1993). A parallel distributed processing model of story comprehension and recall. *Discourse Processes*, 16(3), 203–237. <https://doi.org/10.1080/01638539309544839>
- Goldman, S. R., & Varma, S. (1995). CAPping the construction-integration model of discourse representation. In C. Weaver, S. Mannes, & C. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 337–358). Erlbaum.
- Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, 115, 577–601. <https://doi.org/10.1037/a0012667>
- Goodkind, A., & Bicknell, K. (2021). Local word statistics affect reading times independently of surprisal. *arXiv*. <https://doi.org/10.48550/arxiv.2103.04469>
- Gough, P. B. (1972). One second of reading. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Reading by ear and eye* (pp. 331–358). MIT Press.
- Grainger, J., Lopez, D., Eddy, M., Dufau, S., & Holcomb, P. J. (2012). How word frequency modulates masked repetition priming: An ERP investigation: Word frequency and masked repetition priming. *Psychophysiology*, 49(5), 604–616. <https://doi.org/10.1111/j.1469-8986.2011.01337.x>
- Hahn, M., Futrell, R., Levy, R., & Gibson, E. (2022). A resource-rational model of human processing of recursive linguistic structure. *Proceedings of the National Academy of Science of the United States of America*, 119(43), e2122602119. <https://doi.org/10.1073/pnas.2122602119>
- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In: *Proceedings of the second meeting of the North American chapter of the Association for Computational Linguistics on Language Technologies* (pp. 1–8). <https://doi.org/10.3115/1073336.1073357>
- Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research*, 32(2), 101–123. <https://doi.org/10.1023/A:1022492123056>

- Hale, J. (2006). Uncertainty about the rest of the sentence. *Cognitive Science*, 30(4), 643–672. https://doi.org/10.1207/s15516709cog0000_64
- Hale, J. (2016). Information-theoretical complexity metrics. *Language and Linguistics Compass*, 10(9), 397–412. <https://doi.org/10.1111/lnc3.12196>
- Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., & Dale, A. M. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *NeuroImage*, 17(3), 1101–1116. <https://doi.org/10.1006/nimg.2002.1268>
- Hand, C. J., Miellet, S., O'Donnell, P. J., & Sereno, S. C. (2010). The frequency-predictability interaction in reading: It depends where you're coming from. *Journal of Experimental Psychology: Human Perception and Performance*, 36(5), 1294–1313. <https://doi.org/10.1037/a0020363>
- Hand, C. J., O'Donnell, P. J., & Sereno, S. C. (2012). Word-initial letters influence fixation durations during fluent reading. *Frontiers in Psychology*, 3, 85–85. <https://doi.org/10.3389/fpsyg.2012.00085>
- Hartwigsen, G., Henseler, I., Stockert, A., Wawrzyniak, M., Wendt, C., Klingbeil, J., ..., & Saur, D. (2017). Integration demands modulate effective connectivity in a fronto-temporal network for contextual sentence integration. *NeuroImage*, 147, 812–824. <https://doi.org/10.1016/j.neuroimage.2016.08.026>
- Hauk, O., & Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*, 115(5), 1090–1103. <https://doi.org/10.1016/j.clinph.2003.12.020>
- Helenius, P., Salmelin, R., Service, E., & Connolly, J. F. (1998). Distinct time courses of word and context comprehension in the left temporal cortex. *Brain*, 121(6), 1133–1142. <https://doi.org/10.1093/brain/121.6.1133>
- Herwig, U., Baumgartner, T., Kaffenberger, T., Brühl, A., Kottlow, M., Schreier-Gasser, U., ..., & Rufer, M. (2007). Modulation of anticipatory emotion and perception processing by cognitive control. *NeuroImage*, 37(2), 652–662. <https://doi.org/10.1016/j.neuroimage.2007.05.023>
- Himmelstoss, N. A., Schuster, S., Hutzler, F., Moran, R., & Hawelka, S. (2020). Coregistration of eye movements and neuroimaging for studying contextual predictions in natural reading. *Language, Cognition and Neuroscience*, 35(5), 595–612. <https://doi.org/10.1080/23273798.2019.1616102>
- Hofmann, M. J., Remus, S., Biemann, C., Radach, R., & Kuchinke, L. (2022). Language models explain word reading times better than empirical predictability. *Frontiers in Artificial Intelligence*, 4, 730570. <https://doi.org/10.3389/frai.2021.730570>
- Hohwy, J. (2013). *The predictive mind*. Oxford University Press.
- Hohwy, J. (2020). New directions in predictive processing. *Mind & Language*, 35(2), 209–223. <https://doi.org/10.1111/mila.12281>
- Hubbard, R. J., & Federmeier, K. D. (2021). Dividing attention influences contextual facilitation and revision during language comprehension. *Brain Research*, 1764, 147466. <https://doi.org/10.1016/j.brainres.2021.147466>
- Hubbard, R. J., & Federmeier, K. D. (2023). The impact of linguistic prediction violations on downstream recognition memory and sentence recall. *Journal of Cognitive Neuroscience*, 36(1), 1–23. https://doi.org/10.1162/jocn_a_02078
- Hubbard, R. J., Rommers, J., Jacobs, C. L., & Federmeier, K. D. (2019). Downstream behavioral and electrophysiological consequences of word prediction on recognition memory. *Frontiers in Human Neuroscience*, 13, 291–291. <https://doi.org/10.3389/fnhum.2019.00291>
- Huetting, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135. <https://doi.org/10.1016/j.brainres.2015.02.014>
- Huetting, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31(1), 80–93. <https://doi.org/10.1080/23273798.2015.1047459>
- Huetting, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, 31(1), 19–31. <https://doi.org/10.1080/23273798.2015.1072223>
- Huetting, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137(2), 151–171. <https://doi.org/10.1016/j.actpsy.2010.11.003>
- Ihara, A., Hayakawa, T., Wei, Q., Munetsuna, S., & Fujimaki, N. (2007). Lexical access and selection of contextually appropriate meaning for ambiguous words. *NeuroImage*, 38(3), 576–588. <https://doi.org/10.1016/j.neuroimage.2007.07.047>
- Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, 40, 431–439. <https://doi.org/10.3758/BF03208203>
- Ito, A., & Pickering, M. J. (2021). Automaticity and prediction in non-native language comprehension. In E. Kaan & T. Grüter (Eds.), *Prediction in second language processing and learning* (pp. 25–46). John Benjamins.
- Ito, A., Corley, M., & Pickering, M. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism*, 21(2), 251–264. <https://doi.org/10.1017/S1366728917000050>
- Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language*, 86, 157–171. <https://doi.org/10.1016/j.jml.2015.10.007>
- Ito, A., Martin, A. E., & Nieuwland, M. S. (2017). How robust are prediction effects in language comprehension? Failure to replicate article-elicited N400 effects. *Language, Cognition and Neuroscience*, 32(8), 954–965. <https://doi.org/10.1080/23273798.2016.1242761>
- Jackendoff, R. (2002). *Foundations of language brain, meaning, grammar, evolution*. Oxford University Press.
- Johansson, R. S., & Cole, K. J. (1992). Sensory-motor coordination during grasping and manipulative actions. *Current Opinion in Neurobiology*, 2(6), 815–823. [https://doi.org/10.1016/0959-4388\(92\)90139-C](https://doi.org/10.1016/0959-4388(92)90139-C)
- Jongman, S. R., Copeland, A., Xu, Y., Payne, B. R., & Federmeier, K. D. (2022). Older adults show intraindividual variation in the use of predictive processing. *Experimental Aging Research*, 1–24. <https://doi.org/10.1080/0361073X.2022.2137358>. Advance online publication. 1–24
- Juhasz, B. J., White, S. J., Liversedge, S. P., & Rayner, K. (2008). Eye movements and the use of parafoveal word length information in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 1560–1579. <https://doi.org/10.1037/a0012319>
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20(2), 137–194. [https://doi.org/10.1016/S0364-0213\(99\)80005-6](https://doi.org/10.1016/S0364-0213(99)80005-6)
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1), 122–149. <https://doi.org/10.1037/0033-295x.99.1.122>
- Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133–156. [https://doi.org/10.1016/S0749-596X\(03\)00023-8](https://doi.org/10.1016/S0749-596X(03)00023-8)
- Karimi, H., Weber, P., & Zinn, J. (2024). Information entropy facilitates (not impedes) lexical processing during language

- comprehension. *Psychonomic Bulletin & Review*. <https://doi.org/10.3758/s13423-024-02463-x>
- Keller, F. (2010). Cognitively plausible models of human language processing. *Proceedings of the ACL 2010 Conference Short Papers* (pp. 60–67). Association for Computational Linguistics.
- Kennedy, A., Pynte, J., Murray, W. S., & Paul, S. A. (2013). Frequency and predictability effects in the Dundee Corpus: An eye movement analysis. *Quarterly Journal of Experimental Psychology* (2006), 66(3), 601–618. <https://doi.org/10.1080/17470218.2012.676054>
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363–394. <https://doi.org/10.1037/0033-295X.85.5.363>
- Kleiman, G. M. (1980). Sentence frame contexts and lexical decisions: Sentence-acceptability and word-relatedness effects. *Memory & Cognition*, 8(4), 336–344. <https://doi.org/10.3758/BF03198273>
- Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, 16(1/2), 262–284. <https://doi.org/10.1080/09541440340000213>
- Kliegl, R., Hohenstein, S., Yan, M., & McDonald, S. A. (2013). How preview space/time translates into preview cost/benefit for fixation durations during reading. *Quarterly Journal of Experimental Psychology* (2006), 66(3), 581–600. <https://doi.org/10.1080/17470218.2012.658073>
- Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, 135(1), 12–35. <https://doi.org/10.1037/0096-3445.135.1.12>
- Kretzschmar, F., Bornkessel-Schlesewsky, I., & Schlewsky, M. (2009). Parafoveal versus foveal N400s dissociate spreading activation from contextual fit. *NeuroReport*, 20(18), 1613–1618. <https://doi.org/10.1097/WNR.0b013e328332c4f4>
- Kretzschmar, F., Schlewsky, M., & Staub, A. (2015). Dissociating word frequency and predictability effects in reading: Evidence from coregistration of eye movements and EEG. *Journal of Experimental Psychology Learning, Memory, and Cognition*, 41(6), 1648–1662. <https://doi.org/10.1037/xlm0000128>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Kukona, A., Fang, S.-Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition*, 119(1), 23–42. <https://doi.org/10.1016/j.cognition.2010.12.002>
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59. <https://doi.org/10.1080/23273798.2015.1102299>
- Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience*, 32(1), 12–35. https://doi.org/10.1162/jocn_a_01465
- Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes*, 8, 533–572. <https://doi.org/10.1080/01690969308407587>
- Kutas, M., & Federmeier, K. D. (2007). *Event-related brain potential (ERP) studies of sentence processing*. Oxford University Press. <https://doi.org/10.1093/oxfordhpb/9780198568971.013.0023>
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203–205. <https://doi.org/10.1126/science.7350657>
- Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161–163. <https://doi.org/10.1038/307161a0>
- Kutas, M., Lindamood, T. E., & Hillyard, S. A. (1984). Word expectancy and event-related brain potentials during sentence processing. In S. Kornblum & J. Requin (Eds.), *Preparatory States & Processes* (pp. 217–237). Taylor & Francis.
- Kutas, M., Smith, N. J., & DeLong, K. A. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain* (pp. 190–207). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195395518.003.0065>
- Lai, M. K., Rommers, J., & Federmeier, K. D. (2021). The fate of the unexpected: Consequences of misprediction assessed using ERP repetition effects. *Brain Research*, 1757, 147290. <https://doi.org/10.1016/j.brainres.2021.147290>
- Lai, M. K., Payne, B. R., & Federmeier, K. D. (2023). Graded and ungraded expectation patterns: Prediction dynamics during active comprehension. *Psychophysiology*, 61(1), e14424. <https://doi.org/10.1111/psyp.14424>
- Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions: Biological Sciences*, 352(1358), 1231–1239. <https://doi.org/10.1098/rstb.1997.0105>
- Langston, M. C., & Trabasso, T. (1999). Modeling causal integration and availability of information during comprehension of narrative texts. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 29–69). Erlbaum.
- Laszlo, S., & Armstrong, B. C. (2014). PSPs and ERPs: Applying the dynamics of post-synaptic potentials to individual units in simulation of temporally extended event-related potential reading data. *Brain and Language*, 132, 22–27. <https://doi.org/10.1016/j.bandl.2014.03.002>
- Laszlo, S., & Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language*, 61(3), 326–338. <https://doi.org/10.1016/j.jml.2009.06.004>
- Laszlo, S., & Plaut, D. C. (2012). A neurally plausible parallel distributed processing model of event-related potential word reading data. *Brain and Language*, 120(3), 271–281. <https://doi.org/10.1016/j.bandl.2011.09.001>
- Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, 98(1), 74–88. <https://doi.org/10.1016/j.bandl.2006.02.00>
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177. <https://doi.org/10.1016/j.cognition.2007.05.006>
- Lewis, R. L., & Vasisht, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29, 375–419. https://doi.org/10.1207/s15516709cog0000_25
- Li, X., & Pollatsek, A. (2020). An integrated model of word processing and eye-movement control during Chinese reading. *Psychological Review*, 127(6), 1139–1162. <https://doi.org/10.1037/rev0000248>
- Li, H., Warrington, K. L., Pagán, A., Paterson, K. B., & Wang, X. (2021). Independent effects of collocation strength and contextual predictability on eye movements in reading. *Language, Cognition and Neuroscience*, 36(8), 1001–1009. <https://doi.org/10.1080/23273798.2021.1922726>
- Lindborg, A., & Rabovsky, M. (2021). Meaning in brains and machines: Internal activation update in large-scale language model partially reflects the N400 brain potential. *Proceedings*

- of the annual meeting of the Cognitive Science Society, 43, 1049–1055.
- Linzen, T., & Jaeger, T. F. (2016). Uncertainty and expectation in sentence processing: Evidence from subcategorization distributions. *Cognitive Science*, 40, 1382–1411. <https://doi.org/10.1111/cogs.12274>
- Lowder, M. W., Choi, W., Ferreira, F., & Henderson, J. M. (2018). Lexical predictability during natural reading: Effects of surprisal and entropy reduction. *Cognitive Science*, 42(Suppl 4 (Suppl 4)), 1166–1183. <https://doi.org/10.1111/cogs.12597>
- Luke, S. G. (2018). Influences on and consequences of parafoveal preview in reading. *Attention, Perception, & Psychophysics*, 80(7), 1675–1682. <https://doi.org/10.3758/s13414-018-1581-0>
- Luke, S. G., & Christianson, K. (2012). Semantic predictability eliminates the transposed-letter effect. *Memory & Cognition*, 40, 628–641. <https://doi.org/10.3758/s13421-011-0170-4>
- Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology*, 88, 22–60. <https://doi.org/10.1016/j.cogpsych.2016.06.002>
- Luke, S. G., & Christianson, K. (2018). The Provo Corpus: A large eye-tracking corpus with predictability norms. *Behavior Research Methods*, 50(2), 826–833. <https://doi.org/10.3758/s13428-017-0908-4>
- Lupyan, G., & Clark, A. (2015). Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychological Science*, 24(4), 279–284. <https://doi.org/10.1177/0963721415570732>
- Maess, B., Herrmann, C. S., Hahne, A., Nakamura, A., & Friederici, A. D. (2006). Localizing the distributed language network responsible for the N400 measured by MEG during auditory sentence processing. *Brain Research*, 1096(1), 163–172. <https://doi.org/10.1016/j.brainres.2006.04.037>
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W. H Freeman and Company.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1), 71–102. [https://doi.org/10.1016/0010-0277\(87\)90005-9](https://doi.org/10.1016/0010-0277(87)90005-9)
- Marsman, J. B. C., Renken, R., Velichkovsky, B. M., Hooymans, J. M. M., & Cornelissen, F. W. (2012). Fixation based event-related fMRI analysis: Using eye fixations as events in functional magnetic resonance imaging to reveal cortical processing during the free exploration of visual images. *Human Brain Mapping*, 33, 307–318. <https://doi.org/10.1002/hbm.21211>
- Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, 69(4), 574–588. <https://doi.org/10.1016/j.jml.2013.08.001>
- Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin*, 81(12), 899–917. <https://doi.org/10.1037/h0037368>
- McClelland, J. L. (1987). *The case for interactionism in language processing*. Erlbaum.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86(4), 287–330. <https://doi.org/10.1037/0033-295X.86.4.287>
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375–407. <https://doi.org/10.1037/0033-295X.88.5.375>
- McClelland, J. L., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes*, 4(4), SI287–SI335. <https://doi.org/10.1080/01690968908406371>
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676–703. <https://doi.org/10.1037/0033-295X.101.4.676>
- McDonald, S. A., & Shillcock, R. C. (2003a). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, 14(6), 648–652. <https://doi.org/10.1046/j.0956-7976.2003.psci.1480.x>
- McDonald, S. A., & Shillcock, R. C. (2003b). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research*, 43(16), 1735–1751. [https://doi.org/10.1016/S0042-6989\(03\)00237-2](https://doi.org/10.1016/S0042-6989(03)00237-2)
- McDonald, S. A., Carpenter, R. H. S., & Shillcock, R. C. (2005). An anatomically constrained, stochastic model of eye movement control in reading. *Psychological Review*, 112(4), 814–840. <https://doi.org/10.1037/0033-295X.112.4.814>
- McGowan, V. A., & Reichle, E. D. (2018). The “risky” reading strategy revisited: New simulations using E-Z Reader. *Quarterly Journal of Experimental Psychology*, 71(1), 179–189. <https://doi.org/10.1080/17470218.2017.1307424>
- Merkx, D., & Frank, S. L. (2021). Human sentence processing: Recurrence or attention? *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics* (pp. 12–22). Association for Computational Linguistics.
- Mesulam, M. (2008). Representation, inference, and transcendent encoding in neurocognitive networks of the human brain. *Annals of Neurology*, 64(4), 367–378. <https://doi.org/10.1002/ana.21534>
- Metusalem, R., Kutas, M., Urbach, T. P., Hare, M., McRae, K., & Elman, J. L. (2012). Generalized event knowledge activation during online sentence comprehension. *Journal of Memory and Language*, 66, 545–567. <https://doi.org/10.1016/j.jml.2012.01.001>
- Meyer, D., Schvaneveldt, R., & Ruddy, M. (1975). Loci of contextual effects on visual word recognition. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 998–118). Academic Press.
- Michaelov, J. A., Coulson, S., & Bergen, B. K. (2021). So cloze yet so far: N400 amplitude is better predicted by distributional information than human predictability judgements. *IEEE Transactions on Cognitive and Developmental Systems*, 15(3), 1033–1042. <https://doi.org/10.1109/TCDS.2022.3176783>
- Morris, R. K. (2006). Lexical processing and sentence context effects. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 377–401). Elsevier. <https://doi.org/10.1016/B978-012369374-7/50011-0>
- Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental psychology: Human Perception and Performance*, 10(5), 667–682. <https://doi.org/10.1037/0096-1523.10.5.667>
- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76(2), 165–178. <https://doi.org/10.1037/h0027366>
- Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, 26(2/3), 131–157. <https://doi.org/10.1080/01638539809545042>
- Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, 106(3), 226–254. <https://doi.org/10.1037/0096-3445.106.3.226>
- Ness, T., & Meltzer-Asscher, A. (2018). Lexical inhibition due to failed prediction: Behavioral evidence and ERP correlates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(8), 1269–1285. <https://doi.org/10.1037/xlm0000525>
- Nour Eddine, S., Brothers, T., & Kuperberg, G. R. (2022). The N400 in silico: A review of computational models. In K. Federmeier

- (Ed.), *Psychology of learning and motivation* (vol 76, pp. 123–206). Academic Press.
- Ng, S., Payne, B. R., Steen, A. A., Stine-Morrow, E. A. L., & Federmeier, K. D. (2017). Use of contextual information and prediction by struggling adult readers: Evidence from reading times and event-related potentials. *Scientific Studies of Reading*, 21(5), 359–375. <https://doi.org/10.1080/10888438.2017.1310213>
- Nieuwland, M. S. (2019). Do ‘early’ brain responses reveal word form prediction during language comprehension? A critical review. *Neuroscience and Biobehavioral Review*, 96, 367–400. <https://doi.org/10.1016/j.neubiorev.2018.11.019>
- Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., ..., & Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *eLife*, 7, e33468. <https://doi.org/10.7554/eLife.33468>
- Nieuwland, M., Barr, D., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D., ..., & von Grebmer zu Wolfsturn, S. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. *Philosophical Transactions of the Royal Society of London: Series B Biological Sciences*, 375(1791), 20180522. <https://doi.org/10.1098/rstb.2018.0522>
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3), 189–234. [https://doi.org/10.1016/0010-0277\(94\)90043-4](https://doi.org/10.1016/0010-0277(94)90043-4)
- Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, 113(2), 327–357. <https://doi.org/10.1037/0033-295X.113.2.327>
- Oh, B., & Schuler, W. (2023). Why does surprisal from larger transformer-based language models provide a poorer fit to human reading times? *Transactions of the Association for Computational Linguistics*, 11, 336–350. https://doi.org/10.1162/tacl_a_00548
- Onnis, L., Lim, A., Cheung, S., & Huettig, F. (2022). Is the mind inherently predicting? Exploring forward and backward looking in language processing. *Cognitive Science*, 46(10), e13201. <https://doi.org/10.1111/cogs.13201>
- Otten, M., & Van Berkum, J. J. A. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain Research*, 1291, 92–101. <https://doi.org/10.1016/j.brainres.2009.07.042>
- Owsley, C. (2011). Aging and vision. *Vision Research*, 51(13), 1610–1622. <https://doi.org/10.1016/j.visres.2010.10.020>
- Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word-superiority effect. *Psychological Review*, 89, 573–594. <https://doi.org/10.1037/0033-295X.89.5.573>
- Parker, A. J., & Slattery, T. J. (2019). Word frequency, predictability, and return-sweep saccades: Towards the modeling of eye movements during paragraph reading. *Journal of Experimental Psychology: Human Perception and Performance*, 45(12), 1614–1633. <https://doi.org/10.1037/xhp0000694>
- Parker, A. J., Kirkby, J. A., & Slattery, T. J. (2017). Predictability effects during reading in the absence of parafoveal preview. *Journal of Cognitive Psychology*, 29(8), 902–911. <https://doi.org/10.1080/20445911.2017.1340303>
- Paterson, K. B., McGowan, V. A., Warrington, K. L., Li, L., Li, S., Xie, F., ..., & Wang, J. (2020). Effects of normative aging on eye movements during reading. *Vision*, 4, 7. <https://doi.org/10.3390/vision4010007>
- Payne, B. R., & Federmeier, K. D. (2017a). Event-related brain potentials reveal age-related changes in parafoveal-foveal integration during sentence processing. *Neuropsychologia*, 106, 358–370. <https://doi.org/10.1016/j.neuropsychologia.2017.10.002>
- Payne, B. R., & Federmeier, K. D. (2017b). Pace yourself: Intraindividual variability in context use revealed by self-paced event-related brain potentials. *Journal of Cognitive Neuroscience*, 29(5), 837–854. https://doi.org/10.1162/jocn_a_01090
- Payne, B., & Federmeier, K. D. (2019). Individual differences in reading speed are linked to variability in the processing of lexical and contextual information: Evidence from single-trial event-related brain potentials. *Word*, 65(4), 252–272. <https://doi.org/10.1080/00437956.2019.1678826>
- Payne, B. R., & Silcox, J. W. (2019). Aging, context processing, and comprehension. *Psychology of Learning and Motivation*, 71, 215–264. <https://doi.org/10.1016/bs.plm.2019.07.001>
- Payne, B. R., Gao, X., Noh, S. R., Anderson, C. J., & Stine-Morrow, E. A. L. (2012). The effects of print exposure on sentence processing and memory in older adults: Evidence for efficiency and reserve. *Aging, Neuropsychology, and Cognition*, 19(1/2), 122–149. <https://doi.org/10.1080/13825585.2011.628376>
- Payne, B. R., Stites, M. C., & Federmeier, K. D. (2019). Event-related brain potentials reveal how multiple aspects of semantic processing unfold across parafoveal and foveal vision during sentence reading. *Psychophysiology*, 56(10), e13432. <https://doi.org/10.1111/psyp.13432>
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 67–86). John Benjamins.
- Perfetti, C. A., & Lesgold, A. M. (1979). Coding and comprehension in skilled reading and implications for reading instruction. In L. Resnick & P. Weaver (Eds.), *Theory and practice of early reading*. Erlbaum.
- Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, 114(2), 273–315. <https://doi.org/10.1037/0033-295X.114.2.273>
- Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, 144(10), 1002–1044. <https://doi.org/10.1037/bul0000158>
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–190. <https://doi.org/10.1017/S0140525X04000056>
- Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11, 105–110. <https://doi.org/10.1016/j.tics.2006.12.002>
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36, 329–347. <https://doi.org/10.1017/S0140525X12001495>
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, 62(2), 816–847. <https://doi.org/10.1016/j.neuroimage.2012.04.062>
- Rabe, M. M., Paape, D., Mertzen, D., Vasishth, S., & Engbert, R. (2023). SEAM: An integrated activation-coupled model of sentence processing and eye movements in reading. *Journal of Memory and Language*. <https://doi.org/10.48550/arxiv.2303.05221>
- Rabovsky, M. (2020). Change in a probabilistic representation of meaning can account for N400 effects on articles: A neural network model. *Neuropsychologia*, 143, 107466. <https://doi.org/10.1016/j.neuropsychologia.2020.107466>
- Rabovsky, M., & McRae, K. (2014). Simulating the N400 ERP component as semantic network error: Insights from a feature-based connectionist attractor model of word meaning. *Cognition*, 132(1), 68–89. <https://doi.org/10.1016/j.cognition.2014.03.010>

- Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2018). Modelling the N400 brain potential as change in a probabilistic representation of meaning. *Nature Human Behaviour*, 2(9), 693–705. <https://doi.org/10.1038/s41562-018-0406-4>
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8), 9.f.
- Rayner, K. (1975). Parafoveal identification during a fixation in reading. *Acta Psychologica*, 39(4), 271–281. [https://doi.org/10.1016/0001-6918\(75\)90011-6](https://doi.org/10.1016/0001-6918(75)90011-6)
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506. <https://doi.org/10.1080/17470210902816461>
- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14(3), 191–201. <https://doi.org/10.3758/BF03197692>
- Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review*, 3(2), 245–248. <https://doi.org/10.3758/BF03212426>
- Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, 3(4), 504–509. <https://doi.org/10.3758/BF03214555>
- Rayner, K., & Clifton, C. (2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological Psychology*, 80(1), 4–9. <https://doi.org/10.1016/j.biopsycho.2008.05.002>
- Rayner, K., & Liversedge, S. P. (2011). Linguistic and cognitive influences on eye movements during reading. In S. P. Liversedge (Ed.), *The Oxford handbook of eye movements* (pp. 752–766). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199539789.013.0041>
- Rayner, K., Binder, K. S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research*, 41(7), 943–954. [https://doi.org/10.1016/S0042-6989\(00\)00310-2](https://doi.org/10.1016/S0042-6989(00)00310-2)
- 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(4), 720–732. <https://doi.org/10.1037/0096-1523.30.4.720>
- Rayner, K., Li, X., & Juhasz, B. J. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review*, 12, 1089–1093. <https://doi.org/10.3758/BF03206448>
- Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, 21(3), 448–465. <https://doi.org/10.1037/0882-7974.21.3.448>
- Rayner, K., Pollatsek, A., Drieghe, D., Slattery, T. J., & Reichle, E. D. (2007). Tracking the mind during reading via eye movements: Comments on Kliegl, Nuthmann, and Engbert (2006). *Journal of Experimental Psychology. General*, 136(3), 520–529. <https://doi.org/10.1037/0096-3445.136.3.520>
- Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, 37(2), 514–528. <https://doi.org/10.1037/a0020990>
- Reichle, E. D. (2021). *Computational models of reading: A Handbook*. Oxford University Press.
- Reichle, E. D., & Yu, L. (2024). *The psychology of reading: Insights from Chinese*. Cambridge University Press.
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, 105(1), 125–157. <https://doi.org/10.1037/0033-295x.105.1.125>
- Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, 16(1), 1–21. <https://doi.org/10.3758/PBR.16.1.1>
- Reichle, E. D., Pollatsek, A., & Rayner, K. (2012). Using E-Z Reader to simulate eye movements in nonreading tasks: A unified framework for understanding the eye-mind link. *Psychological Review*, 119(1), 155–185. <https://doi.org/10.1037/a0026473>
- Reilly, R. (1993). A connectionist framework for modeling eye-movement control in reading. In G. d'Ydewalle & J. Van Rensbergen (Eds.), *Perception and cognition: Advances in eye movement research* (pp. 193–212). Elsevier.
- Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research*, 7(1), 34–55. <https://doi.org/10.1016/j.cogsys.2005.07.006>
- Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z Reader model. *Psychological Science*, 17(9), 742–746. <https://doi.org/10.1111/j.1467-9280.2006.01775.x>
- Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive Psychology*, 65(2), 177–206. <https://doi.org/10.1016/j.cogpsych.2012.03.001>
- Rich, S., & Harris, J. A. (2023). Global expectations mediate local constraint: Evidence from concessive structures. *Language, Cognition and Neuroscience*, 38(3), 302–327. <https://doi.org/10.1080/23273798.2022.2114598>
- Roark, B., Bachrach, A., Cardenas, C., & Pallier, C. (2009). Deriving lexical and syntactic expectation-based measures for psycholinguistic modeling via incremental top-down parsing. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing* (pp. 324–333). Association for Computational Linguistics.
- Rommers, J., & Federmeier, K. D. (2018a). Lingering expectations: A pseudo-repetition effect for words previously expected but not presented. *NeuroImage*, 183, 263–272. <https://doi.org/10.1016/j.neuroimage.2018.08.023>
- Rommers, J., & Federmeier, K. D. (2018b). Predictability's aftermath: Downstream consequences of word predictability as revealed by repetition effects. *Cortex*, 101, 16–30. <https://doi.org/10.1016/j.cortex.2017.12.018>
- Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition*, 18(4), 367–379. <https://doi.org/10.3758/BF03197126>
- Rumelhart, D. E. (1975). Understanding and summarizing stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 211–236). Academic Press.
- Ryskin, R., & Nieuwland, M. S. (2023). Prediction during language comprehension: What is next? *Trends in Cognitive Sciences*. <https://doi.org/10.1016/j.tics.2023.08.003>
- Ryskin, R., Levy, R. P., & Fedorenko, E. (2020). Do domain-general executive resources play a role in linguistic prediction? Re-evaluation of the evidence and a path forward. *Neuropsychologia*, 136, 107258. <https://doi.org/10.1016/j.neuropsychologia.2019.107258>

- Salthouse, T. A. (1993). Effects of aging on verbal abilities: Examination of the psychometric literature. In D. M. Burke & L. L. Light (Eds.), *Language, memory, and aging* (pp. 17–35). Cambridge University Press. <https://doi.org/10.1017/CBO9780511575020.003>
- Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, 74(1), 5–35. <https://doi.org/10.3758/s13414-011-0219-2>
- Schuster, S., Hawelka, S., Hutzler, F., Kronbichler, M., & Richlan, F. (2016). Words in context: The effects of length, frequency, and predictability on brain responses during natural reading. *Cerebral Cortex*, 26(10), 3889–3904. <https://doi.org/10.1093/cercor/bhw184>
- Schuster, S., Hawelka, S., Himmelstoss, N. A., Richlan, F., & Hutzler, F. (2020). The neural correlates of word position and lexical predictability during sentence reading: Evidence from fixation-related fMRI. *Language, Cognition and Neuroscience*, 35(5), 613–624. <https://doi.org/10.1080/23273798.2019.1575970>
- Schuster, S., Himmelstoss, N. A., Hutzler, F., Richlan, F., Kronbichler, M., & Hawelka, S. (2021). Cloze enough? Hemodynamic effects of predictive processing during natural reading. *NeuroImage*, 228, 117687–117687. <https://doi.org/10.1016/j.neuroimage.2020.117687>
- Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology Learning, Memory, and Cognition*, 14(2), 344–354. <https://doi.org/10.1037/0278-7393.14.2.344>
- Schwanenflugel, P. J., & Shoben, E. J. (1985). The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language*, 24(2), 232–252. [https://doi.org/10.1016/0749-596X\(85\)90026-9](https://doi.org/10.1016/0749-596X(85)90026-9)
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71(2), 109–147. [https://doi.org/10.1016/S0010-0277\(99\)00025-6](https://doi.org/10.1016/S0010-0277(99)00025-6)
- Seelig, S. A., Rabe, M. M., Malem-Shinitski, N., Risse, S., Reich, S., & Engbert, R. (2020). Bayesian parameter estimation for the SWIFT model of eye-movement control during reading. *Journal of Mathematical Psychology*, 95, 102313. <https://doi.org/10.1016/j.jmp.2019.102313>
- Senior, C., Barnes, J., Giampietroc, V., Simmons, A., Bullmore, E. T., Brammer, M., & David, A. S. (2000). The functional neuroanatomy of implicit-motion perception or 'representational momentum.' *Current Biology*, 10(1), 16–22. [https://doi.org/10.1016/S0960-9822\(99\)00259-6](https://doi.org/10.1016/S0960-9822(99)00259-6)
- Sereno, S. C., & Rayner, K. (2000). The when and where of reading in the brain. *Brain and Cognition*, 42(1), 78–81. <https://doi.org/10.1006/brcg.1999.1167>
- Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences*, 7(11), 489–493. <https://doi.org/10.1016/j.tics.2003.09.010>
- Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of word recognition: Evidence from eye movements and event-related potentials. *NeuroReport*, 9(10), 2195–2200. <https://doi.org/10.1097/00001756-199807130-00009>
- Sereno, S. C., Hand, C. J., Shahid, A., Yao, B., & O'Donnell, P. J. (2018). Testing the limits of contextual constraint: Interactions with word frequency and parafoveal preview during fluent reading. *Quarterly Journal of Experimental Psychology*, 71(1), 302–313. <https://doi.org/10.1080/17470218.2017.1327981>
- Sereno, S. C., Hand, C. J., Shahid, A., Mackenzie, I. G., & Leuthold, H. (2020). Early EEG correlates of word frequency and contextual predictability in reading. *Language, Cognition and Neuroscience*, 35, 625–640. <https://doi.org/10.1080/23273798.2019.1580753>
- Shain, C. (2019). A large-scale study of the effects of word frequency and predictability in naturalistic reading. In J. Burstein, C. Doran, & T. Solorio (Eds.), *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4086–4094). Association for Computational Linguistics.
- Shain, C. (2024). Word frequency and predictability dissociate in naturalistic reading. *Open Mind*, 5(8), 177–201. https://doi.org/10.1162/opmi_a_00119
- Shain, C., Meister, C., Pimentel, T., Cotterell, R., & Levy, R. P. (2024). Large-scale evidence for logarithmic effects of word predictability on reading time. *Proceedings of the National Academy of Sciences*, 121(10), e2307876121. <https://doi.org/10.31234/osf.io/4hyna>
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 623–656.
- Sheridan, H., & Reingold, E. M. (2012). The time course of predictability effects in reading: Evidence from a survival analysis of fixation durations. *Visual Cognition*, 20(7), 733–745. <https://doi.org/10.1080/13506285.2012.693548>
- Slattery, T. J., & Yates, M. (2018). Word skipping: Effects of word length, predictability, spelling and reading skill. *Quarterly Journal of Experimental Psychology*, 71(1), 250–259. <https://doi.org/10.1080/17470218.2017.1310264>
- Smith, N. J., & Levy, R. (2011). Cloze but no cigar: The complex relationship between cloze, corpus, and subjective probabilities in language processing. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (p. 33). Cognitive Science Society.
- Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, 128(3), 302–319. <https://doi.org/10.1016/j.cognition.2013.02.013>
- Snell, J., van Leipsig, S., Grainger, J., & Meeter, M. (2018). OB1-Reader: A model of word recognition and eye movements in text reading. *Psychological Review*, 125(6), 969–984. <https://doi.org/10.1037/rev0000119>
- Stanovich, K. E. (1984). The interactive-compensatory model of reading: A confluence of developmental, experimental, and educational psychology. *RASE: Remedial & Special Education*, 5(3), 11–19.
- Stanovich, K. E., & West, R. F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance*, 7(3), 658–672. <https://doi.org/10.1037/0096-1523.7.3.658>
- Stanovich, K. E., & West, R. F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General*, 112(1), 1–36. <https://doi.org/10.1037/0096-3445.112.1.1>
- Staub, A. (2011). The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin & Review*, 18(2), 371–376. <https://doi.org/10.3758/s13423-010-0046-9>
- Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass*, 9(8), 311–327. <https://doi.org/10.1111/lnc3.12151>
- Staub, A. (2020). Do effects of visual contrast and font difficulty on readers' eye movements interact with effects of word frequency or predictability? *Journal of Experimental Psychology: Human Perception and Performance*, 46(11), 1235–1251. <https://doi.org/10.1037/xhp0000853>
- Staub, A. (2024). Predictability in language comprehension: Prospects and problems for surprisal. *Annual Review of Linguistics*, 11, 1451–1470. <https://doi.org/10.1146/annurev-linguistics-011724-121517>

- Staub, A., & Benatar, A. (2013). Individual differences in fixation duration distributions in reading. *Psychonomic Bulletin & Review*, 20(6), 1304–1311. <https://doi.org/10.3758/s13423-013-0444-x>
- Staub, A., & Clifton, C. (2006). Syntactic prediction in language comprehension: Evidence from either...or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(2), 425–436. <https://doi.org/10.1037/0278-7393.32.2.425>
- Staub, A., & Goddard, K. (2019). The role of preview validity in predictability and frequency effects on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(1), 110–127. <https://doi.org/10.1037/xlm0000561>
- Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 1280–1293. <https://doi.org/10.1037/a0016896>
- Staub, A., Grant, M., Astheimer, L., & Cohen, A. (2015). The influence of cloze probability and item constraint on cloze task response time. *Journal of Memory and Language*, 82, 1–17. <https://doi.org/10.1016/j.jml.2015.02.004>
- Steen-Baker, A. A., Ng, S., Payne, B. R., Anderson, C. J., Federmeier, K. D., & Stine-Morrow, E. A. L. (2017). The effects of context on processing words during sentence reading among adults varying in age and literacy skill. *Psychology and Aging*, 32(5), 460–472. <https://doi.org/10.1037/pag0000184>
- Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 57(4), 421–457.
- Stites, M. C., Payne, B. R., & Federmeier, K. D. (2017). Getting ahead of yourself: Parafoveal word expectancy modulates the N400 during sentence reading. *Cognitive, Affective, & Behavioral Neuroscience*, 17(3), 475–490. <https://doi.org/10.3758/s13415-016-0492-6>
- Szewczyk, J. M., & Federmeier, K. D. (2022). Context-based facilitation of semantic access follows both logarithmic and linear functions of stimulus probability. *Journal of Memory and Language*, 123, 104311. <https://doi.org/10.1016/j.jml.2021.104311>
- Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes*, 12(2/3), 211–271. <https://doi.org/10.1080/016909697386853>
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science (American Association for the Advancement of Science)*, 268(5217), 1632–1634. <https://doi.org/10.1126/science.7777863>
- Taylor, W. L. (1953). “Cloze procedure”: A new tool for measuring readability. *Journalism Quarterly*, 30, 415–433. <https://doi.org/10.1177/107769905303000401>
- Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology*, 83(3), 382–392. <https://doi.org/10.1016/j.ijpsycho.2011.12.007>
- Traxler, M. J., & Foss, D. J. (2000). Effects of sentence constraint on priming in natural language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1266–1282. <https://doi.org/10.1037/0278-7393.26.5.1266>
- Urbach, T. P., DeLong, K. A., Chan, W.-H., & Kutas, M. (2020). An exploratory data analysis of word form prediction during word-by-word reading. *Proceedings of the National Academy of Sciences*, 117(34), 20483–20494. <https://doi.org/10.1073/pnas.1922028117>
- van den Broek, P., Ridsen, K., Fletcher, C. R., & Thurlow, R. (1996). A “landscape” view of reading: Fluctuating patterns of activation and the construction of a memory representation. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 165–187). Erlbaum.
- Van Berkum, J. J. A., 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(3), 443–467. <https://doi.org/10.1037/0278-7393.31.3.443>
- 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(3), 285–316. [https://doi.org/10.1016/S0749-596X\(03\)00081-0](https://doi.org/10.1016/S0749-596X(03)00081-0)
- Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, 18(4), 380–393. <https://doi.org/10.3758/BF03197127>
- Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory & Cognition*, 19(1), 95–112. <https://doi.org/10.3758/BF03198500>
- Van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language*, 97(3), 279–293. <https://doi.org/10.1016/j.bandl.2005.11.003>
- Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, 83(2), 176–190. <https://doi.org/10.1016/j.ijpsycho.2011.09.015>
- Van Rijn, H., & Anderson, J. R. (2003). Modeling lexical decision as ordinary retrieval. In F. Detje, D. Doerner, & H. Schaub (Eds.), *Paper presented at the Fifth International Conference on Cognitive Modeling*. Universitäts-Verlag Bamberg.
- van Schijndel, M., & Linzen, T. (2018). Can entropy explain successor surprisal effects in reading? In G. Jarosz & J. Pater (Eds.), *Proceedings of the Society for Computation in Linguistics (SCiL)* (pp. 1–7). Society for Computation in Linguistics.
- Vasilev, M. R., & Angele, B. (2017). Parafoveal preview effects from word N + 1 and word N + 2 during reading: A critical review and Bayesian meta-analysis. *Psychonomic Bulletin & Review*, 24, 666–689. <https://doi.org/10.3758/s13423-016-1147-x>
- Vasisht, S., von der Malsburg, T., & Engelmann, F. (2013). What eye movements can tell us about sentence comprehension: Eye movements and sentence comprehension. *Wiley Interdisciplinary Reviews Cognitive Science*, 4(2), 125–134. <https://doi.org/10.1002/wcs.1209>
- Veldre, A., & Andrews, S. (2018). How does foveal processing difficulty affect parafoveal processing during reading? *Journal of Memory and Language*, 103, 74–90. <https://doi.org/10.1016/j.jml.2018.08.001>
- Veldre, A., & Andrews, S. (2018). Parafoveal preview effects depend on both preview plausibility and target predictability. *Quarterly Journal of Experimental Psychology*, 71(1), 64–74. <https://doi.org/10.1080/17470218.2016.1247894>
- Veldre, A., Reichle, E. D., Wong, R., & Andrews, S. (2020). The effect of contextual plausibility on word skipping during reading. *Cognition*, 197, 104184. <https://doi.org/10.1016/j.cognition.2020.104184>
- Veldre, A., Wong, R., & Andrews, S. (2021). Reading proficiency predicts the extent of the right, but not left, perceptual span in older readers. *Attention, Perception, & Psychophysics*, 83, 18–26. <https://doi.org/10.3758/s13414-020-02185-x>
- Veldre, A., Wong, R., & Andrews, S. (2022). Predictability effects and parafoveal processing in older readers. *Psychology and Aging*, 37(2), 222–238. <https://doi.org/10.1037/pag0000659>
- Veldre, A., Reichle, E. D., Yu, L., & Andrews, S. (2023). Understanding the visual constraints on lexical processing: New

- empirical and simulation results. *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/xge0001295>
- Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging*, 18(2), 332–339. <https://doi.org/10.1037/0882-7974.18.2.332>
- Verhaeghen, P. (2013). Cognitive Aging. In D. Reisberg (Ed.), *The Oxford Handbook of Cognitive Psychology*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195376746.001.0001>
- Walsh, K. S., McGovern, D. P., Clark, A., & O'Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, 1464(1), 242–268. <https://doi.org/10.1111/nyas.14321>
- Wang, L., Hagoort, P., & Jensen, O. (2018a). Language prediction is reflected by coupling between frontal gamma and posterior alpha oscillations. *Journal of Cognitive Neuroscience*, 30(3), 432–447. https://doi.org/10.1162/jocn_a_01190
- Wang, L., Kuperberg, G., & Jensen, O. (2018b). Specific lexico-semantic predictions are associated with unique spatial and temporal patterns of neural activity. *eLife*, 7, e39061. <https://doi.org/10.7554/eLife.39061>
- Weiss, K.-L., Hawelka, S., Hutzler, F., & Schuster, S. (2023). Stronger functional connectivity during reading contextually predictable words in slow readers. *Scientific Reports*, 13(1), 5989. <https://doi.org/10.1038/s41598-023-33231-x>
- Weissman, B., Cohn, N., & Tanner, D. (2024). The electrophysiology of lexical prediction of emoji and text. *Neuropsychologia*, 198, 108881. <https://doi.org/10.1016/j.neuropsychologia.2024.108881>
- West, R. F., & Stanovich, K. E. (1982). Source of inhibition in experiments on the effect of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8(5), 385–399. <https://doi.org/10.1037/0278-7393.8.5.385>
- White, S. J., & Staub, A. (2012). The distribution of fixation durations during reading: Effects of stimulus quality. *Journal of Experimental Psychology: Human Perception and Performance*, 38(3), 603–617. <https://doi.org/10.1037/a0025338>
- White, S. J., Rayner, K., & Liversedge, S. P. (2005a). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A reexamination. *Psychonomic Bulletin & Review*, 12(5), 891–896. <https://doi.org/10.3758/BF03196782>
- White, S. J., Rayner, K., & Liversedge, S. P. (2005b). The influence of parafoveal word length and contextual constraint on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review*, 12, 466–471. <https://doi.org/10.3758/BF03193789>
- Whitford, V., & Titone, D. (2014). The effects of reading comprehension and launch site on frequency-predictability interactions during paragraph reading. *Quarterly Journal of Experimental Psychology* 2006, 67(6), 1151–1165. <https://doi.org/10.1080/17470218.2013.848216>
- Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, 8, 221–243. <https://doi.org/10.3758/BF03196158>
- Wicha, N. Y. Y., Bates, E. A., Moreno, E. M., & Kutas, M. (2003a). Potato not Pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, 346, 165–168. [https://doi.org/10.1016/S0304-3940\(03\)00599-8](https://doi.org/10.1016/S0304-3940(03)00599-8)
- Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2003b). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex*, 39(3), 483–508. [https://doi.org/10.1016/S0010-9452\(08\)70260-0](https://doi.org/10.1016/S0010-9452(08)70260-0)
- Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, 16(7), 1272–1288. <https://doi.org/10.1162/0898929041920487>
- Wilcox, E. G., Pimentel, T., Meister, C., Cotterell, R., & Levy, R. P. (2023). Testing the predictions of surprisal theory in 11 languages. *Transactions of the Association for Computational Linguistics*, 11, 1451–1470. https://doi.org/10.1162/tacl_a_00612
- Wlotko, E. W., & Federmeier, K. D. (2012). Age-related changes in the impact of contextual strength on multiple aspects of sentence comprehension. *Psychophysiology*, 49(6), 770–785. <https://doi.org/10.1111/j.1469-8986.2012.01366.x>
- Wlotko, E. W., & Federmeier, K. D. (2015). Time for prediction? The effect of presentation rate on predictive sentence comprehension during word-by-word reading. *Cortex*, 68, 20–32. <https://doi.org/10.1016/j.cortex.2015.03.014>
- Wlotko, E. W., Lee, C. L., & Federmeier, K. D. (2010). Language of the aging brain: Event-related potential studies of comprehension in older adults. *Language and Linguistics Compass*, 4(8), 623–638. <https://doi.org/10.1111/j.1749-818X.2010.00224.x>
- Wlotko, E. W., Federmeier, K. D., & Kutas, M. (2012). To predict or not to predict: Age-related differences in the use of sentential context. *Psychology and Aging*, 27(4), 975–988. <https://doi.org/10.1037/a0029206>
- Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, 11(18), R729–R732. [https://doi.org/10.1016/S0960-9822\(01\)00432-8](https://doi.org/10.1016/S0960-9822(01)00432-8)
- Wong, R., Veldre, A., & Andrews, S. (2022). Are there independent effects of constraint and predictability on eye movements during reading? *Journal of Experimental Psychology: Learning, Memory, and Cognition*. <https://doi.org/10.1037/xlm0001206>
- Wong, R., Veldre, A., & Andrews, S. (2024). Looking for immediate and downstream evidence of lexical prediction in eye movements during reading. *Quarterly Journal of Experimental Psychology*, 0(0). <https://doi.org/10.1177/17470218231223858>
- Woods, W. A. (1970). Transition network grammar for natural language analysis. *Communications of the ACM*, 13, 591–606.
- Wu, S., Bachrach, A., Cardenas, C., & Schuler, W. (2010). Complexity metrics in an incremental right-corner parser. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (pp. 1189–1198). Association for Computational Linguistics.
- Zhang, J., Warrington, K. L., Li, L., Pagán, A., Paterson, K. B., White, S. J., & McGowan, V. A. (2022). Are older adults more risky readers? Evidence from meta-analysis. *Psychology and Aging*, 37(2), 239–259. <https://doi.org/10.1037/pag0000522>
- Zirnsstein, M., van Hell, J. G., & Kroll, J. F. (2018). Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition*, 176, 87–106. <https://doi.org/10.1016/j.cognition.2018.03.001>
- Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance*, 24(4), 1131–1161. <https://doi.org/10.1037/0096-1523.24.4.1131>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.