CHAPTER 15

A Look around at What Lies Ahead: Prediction and Predictability in Language Processing

Marta Kutas, Katherine A. DeLong, and Nathaniel J. Smith

Traditionally, prediction has been considered an inefficient and cognitively expensive processing mechanism in the domain of language comprehension, where there are many possible ways for relaying a single thought, meaning or desire and the chances of mispredicting are accordingly high. Predictive linguistic processing, however, does not seem untenable given its similarity to other neural processing domains that are contextually grounded and appear to implement knowledge- and experience-based mental representations anticipatorily. Here, we examine linguistic prediction from multiple perspectives, ranging from theoretical models that analyze predictability at the level of ambiguity resolution, to experimental evidence primarily from event-related brain potentials (ERPs) that supports a "strong" model of prediction in which items are not just incrementally integrated, but are wholly or featurally pre-activated via accruing mental sentential representations. We also explore possible consequences of a neural language parser (aka, brain) that may be prone to mispredicting, and what electrophysiological evidence for such processing may look like. We conclude by arguing for the importance of investigating such linguistic effects as yet another example of a neural system in which probability estimation is inherent, with a proposal to move beyond the debate of *whether* there is linguistic prediction, toward focusing research efforts on *how* pre-activation may occur and *what* is pre-activated.

Potential benefits of linguistic prediction

Within a variety of neural domains, perhaps including some described in other chapters of this volume, the understanding of how predictive processing benefits a system may be relatively straightforward. With the study of language processing, however, such benefits could be various and perhaps not immediately clear. Linguistic comprehension requires processing a noisy sensory input to recognize complex structures (e.g., phonemes, syllables, words, syntax) and integrating these structures with physical and social contexts, with general/world knowledge, and with each other, to construct meaning. Amazingly, this whole process occurs over mere hundreds of milliseconds. As such, there are a number of ways that linguistic comprehension could be facilitated by predicting upcoming material.

Given the time constraints under which comprehension operates, one clear benefit of being able to predict upcoming material is that it may allow a listener or reader to produce an overt response more quickly, without waiting for the material itself to become available. For instance, Marslen-Wilson (1973) argued that experimental participants performing verbal shadowing (i.e., listening to a stream of speech and simultaneously repeating it aloud) were—given the form of the shadowers' speech errors in which they supplied structurally appropriate, semantically and syntactically congruent continuations—using predictions about upcoming speech to achieve faster response latencies. Similarly, Sacks, Schegloff, and Jefferson (1974) observed that during ordinary dialogue, one interlocutor commonly begins to talk precisely at the moment that the previous speaker finishes, and they argued that listeners must be predicting the timing of such "transition-relevance places" to properly time their own responses.

Another potential benefit to linguistic prediction relates to a challenge faced by the linguistic processor: The appropriate interpretation of the current input often depends on future input. Consider the classic "garden-path" sentence, "*The horse raced past the barn fell*" (Bever, 1970). When the processor has encountered only part of this sentence—"*The horse raced past…*"—then it is ambiguous whether *raced* is the sentence's main verb (referring to what the horse did), or the beginning of a relative clause (describing the horse). In principle, the language processor should therefore wait until the end of the sentence to interpret *raced*. Only when it reaches *fell* would it discover that the relative clause reading ("*The horse, (that was) raced past the barn, fell*") is correct, but by this point, most people have already committed to a main verb reading and thus are at least briefly stymied. This early commitment can be seen as a prediction about the form of the remainder of the sentence, and the confusion the result of an inaccurate prediction. But garden-path sentences are relatively rare in real life, and when this early prediction is accurate, it should allow substantial benefits in speed of processing, reduced memory load, and so forth. In this sense, any language processor that proceeds incrementally, without waiting for the end of each utterance before beginning its work, could be seen as inherently predictive. However, this does not mean that such a processor is forming explicit expectations about upcoming material. Many classic models of such effects instead rely on various bottom-up heuristics that—as a kind of happy coincidence, from the processor's point of view—often result in choices that later turn out to have been correct (e.g., Frazier & Fodor, 1978).

The difficulties caused by garden-path sentences can also be seen as a special case of a more general problem in language comprehension: the problem of ambiguity. In addition to the temporary ambiguity created by temporal sequencing, there is ambiguity created by (at least) sensory noise (*letter* and *ladder* may be acoustically indistinguishable in a noisy room, or compare "*wreck a nice*" and *recognize*), lexical items with multiple senses (*palms*, the tree versus the body part), and syntactic ambiguity ("*I saw the girl in the park with the telescope*"—who had the telescope?). Just as comprehenders might resolve temporary ambiguities by predicting upcoming input, they may resolve these more durable ambiguities by "predicting" the speaker's intent. While such cases are, on a surface level, quite different from prediction as it is usually understood, these two kinds of ambiguity resolution are unified in probabilistic models of language comprehension, where both reduce to the problem of picking the most likely interpretation (Jurafsky, 1996).

A final possible role for prediction in language is that, by forming explicit predictions of upcoming material, the parser may not only ease the processing of current material (as described earlier), but get a head start on future material, leading to increased efficiency if this anticipatory processing turns out to be accurate (though it undoubtedly will not always be). Broadly construed, then, this "strongly" predictive view of language comprehension posits that comprehenders utilize a variety of sources of constraint as they become available to pre-activate upcoming items and/or their features, with aspects of linguistic input being processed both incrementally and in parallel. This type of presumably rapid, online pre-activation that most likely does not reach the level of consciousness has (until recently) been difficult to tap into experimentally, and thus, it has historically also been the most controversial.

Historical opposition to predictive language processing

Despite general considerations about biological continuity, informal intuitions stemming from our experiences "taking the words out of" a speaker's mouth, and evidence from specialized tasks or instances (such as the shadowing procedure and garden-path sentences described earlier), prediction has played a relatively minor role in theories of language processing. The idea that individuals might predict linguistic features or content has not been part of the generative grammar tradition. Indeed, linguists and psycholinguists alike (e.g., Jackendoff, 2002; Morris, 2006) have argued that with infinite options available as each new word of an unfolding sentence is encountered, predicting what comes next is not just improbable but nonviable as a strategy, except perhaps on the rare occasions when contextual constraint is unusually high (Stanovich & West, 1979).

Results from controlled experiments seem to support this proposal: When participants are asked to continue a phrase or complete a sentence, their responses routinely converge on the same word when contextual constraint is strong but show greater variance under weak constraints (e.g., Bloom & Fischler, 1980; Schwanenflugel, 1986). In off-line language tasks, then, it is widely acknowledged that with *sufficient time* individuals do use sentential context to select the most probable linguistic completions. Most language, however, is not so highly constrained. In either case, there is a question as to whether this more deliberate, conscious (post-lexical) offline strategy translates to the rapid, less conscious processing that seems to characterize real-time language comprehension. Or is the real-time use of contextual constraint of a qualitatively different nature? Indeed, a major outstanding question in the sentence processing literature has been whether information about particular words or their features gets pre-activated during on-line sentence processing as a result of top-down contextual processing, or whether word processing is stimulus driven (bottom up), triggered by the input, that is, initiated only *after* the physical stimulus has been received. Perhaps surprisingly to some, there is no clear consensus. In fact, most theories of sentence processing have not addressed this issue directly, even though some stance about linguistic prediction is often implicit.

Why the paucity of research on this question? As already mentioned, natural language is not considered to be very constraining, certainly not constraining enough for a predictive system to be accurate a majority of the time. In principle, such errors should result in processing costs, but such costs have rarely been evidenced (Gough, Alford, & Holley-Wilcox, 1981). Even those who grant some word-level prediction from sentential context argue that word recognition can only benefit from context under special circumstances such as when the target input is degraded, when targets are temporally offset with respect to the context (allowing time for prediction), or when readers are unskilled, and thus slow (Mitchell, 1982). Furthermore, the ecological validity of the paradigms from which much of the existing evidence for prediction comes has been questioned. Rayner, Ashby, Pollatsek, and Reichle (2004), for example, point to the typically nonnaturalistic stimulus presentation rates (e.g., artificial timing of sentences or context-target word delays), which they believe provide readers with ample time to form conscious predictions, and thus bear little resemblance to natural language. We suggest that a more insidious (and perhaps unconscious) component to anti-prediction biases in language research may have to do with the long history of modular views of language processing (e.g., Fodor, 1983; Forster, 1979) and their inherently bottom-up emphasis. On the view that language comprehension is a context-invariant, data-driven process, with information from one neural domain unlikely to preemptively influence processing in another, anticipation of upcoming input based on contextual cues would seem untenable. In sum, some combination of these reasons, among others, has likely contributed to the general lack of exploration of anticipatory comprehension in language historically.

Fortunately, the tides are changing, and in the following sections we review evidence demonstrating that both incremental processing and estimation of linguistic probabilities are pervasive. Building on this work, we then present recent results which argue that the language processor implements even the "strong form" of prediction, regularly pre-activating material in an anticipatory fashion. In this discussion, we provide some evidence for how such anticipatory processing may be instantiated in the brain as well as examine potential consequences of predicting incorrectly.

From buffering to incrementality

Early models of sentence processing included some form of memory buffer by which sentential elements were temporarily stored for later integration at phrasal, clausal, or sentence boundaries (Abney, 1989; Carroll & Bever, 1978; Daneman & Carpenter, 1983; Just & Carpenter, 1980; Kintsch & van Dijk, 1978; Mitchell & Green, 1978; Pritchett, 1991). Eventually, such delayed processing models became difficult to reconcile with accumulating evidence for context updating on a word-by-word basis. Notions of buffering gradually gave way to the more widely held view that words are incorporated successively into the sentential context as they are received and identified, with gradual accretion of meaning in the mind of the comprehender—a more *incremental* view (e.g., Altmann & Steedman, 1988; Boland, Tanenhaus, Garnsey, & Carlson, 1995; Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995; Kutas & Hillyard, 1983; Marslen-Wilson, 1975; Marslen-Wilson & Tyler, 1980; Pickering, 1994; Steedman, 1989; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Traxler, Bybee, & Pickering, 1997; Tyler & Marslen-Wilson, 1977).

The view that language comprehension is largely incremental is rooted in evidence from a wide variety of studies with different methodologies, including off-line and on-line techniques. For instance, Boland et al. (1995) asked readers to indicate via a button press when a sentence stopped making sense. Presented with identically structured "*wh*"-questions such as "*Which military base did Hank deliver the machine guns to __ last week?*" vs. "*Which preschool did Hank deliver the machine guns to__ last week?*" participants pushed the button at *machine guns* in the *preschool* version but not in the *military base* version. These results indicate that the thematic role of the questioned element (*military base/preschool*), displaced from its location in a declarative sentence (e.g., *Hank delivered the machine guns to which military base/preschool*), is assigned as soon as the verb (*deliver*) is encountered, rather than buffered for interpretation at sentence end; in other words, it is processed incrementally. Using a different behavioral method, Marslen-Wilson (1975) showed that so-called fast shadowers—individuals that repeated back-recorded speech with very little delay (sometimes as short as 250 ms)—corrected pronunciation errors in the recorded speech signal, indicating that they were processing the shadowed text at a semantic level, *as* it was received.

Though clear evidence for incrementality has emerged from such behavioral work, the strongest evidence to date comes from investigations of real-time language processing using techniques like eye-tracking and event-related brain potentials (ERPs) that *(a)* do not require additional tasks (e.g., button presses or spoken responses) other than listening or reading, and *(b)* are able to continuously track sentence processing with high temporal resolution throughout the course of a sentence, thereby making it easier to isolate the precise processing stage implicated. In eye-tracking studies, for instance, the time-locked characteristics of eye movements provide information about the processes that underlie spoken language comprehension. In the visual world paradigm, for example, individuals' eye movements are monitored as they look at a visual scene while simultaneously listening to a sentence or set of instructions that refers to objects in that scene (Tanenhaus et al., 1995). Many such studies have found that participants make saccadic eye movements to depicted objects without delay upon hearing the relevant input (often prior to the referent word itself), supporting models of word-by-word integration (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002).

Similarly, ERPs—scalp recordings of the synaptic potentials generated primarily by synchronously firing multiple pyramidal cells in the neocortex—also afford a continuous but time-stamped look at ongoing neural activity during written and spoken language processing. Kutas and Hillyard (1980), for example, showed that a semantically anomalous word within a certain context elicits an enhanced ERP component known as the N400 at the point at which it is encountered—indeed, within 200 ms of the word's onset—relative to a semantically congruent continuation in the same context. The N400, as a neural response to any potentially meaningful item, has sometimes been considered to index semantic fit of an item in a particular context, and as such, provides additional evidence for context updating before the end of an utterance. Thus, these and other studies (e.g., van Berkum, Brown, & Hagoort, 1999; van Berkum, Koornneef, Otten, & Nieuwland, 2007; van Berkum, Zwitserlood, Hagoort, & Brown, 2003; Van Petten & Kutas, 1990) indicate that sentence processing is incremental.

Anticipatory language processing

In the past decade or so, however, there has been a wave of both empirical and modeling work suggesting that language processing may not *just* be incremental; instead, evidence from various sources has supported the view that sentential constraint begins to exert its influence *before* words have been uniquely identified. We might even consider this somewhat of a "revival," since a few early language processing models argued specifically for more anticipatory processing in which available contextual factors are used to activate words in advance of receiving them (McClelland & Elman, 1986; McClelland & Rumelhart, 1981). More recent investigations of anticipatory processing have been spurred, in part, by interest in questions of how generalized prediction might be, what it might be used for, what information is available to predictive processing, and what aspects of upcoming input are being predicted.

Eye-tracking methods, for instance, have successfully employed the visual world paradigm to detect preferential looks to visual targets before complete auditory information is available. As an example, Eberhard et al. (1995) found that when participants were given an instruction "*Pick up the candle*," they not only immediately initiated a saccade to an actual candle in front of them, but did so before the second syllable of *candle* had been fully articulated. Furthermore, this predictive process was sensitive to visual context; when the same instruction was given to participants whose work area contained both a candle and candy, so that the first syllable *can-* was no longer sufficient to identify the reference, the average saccade initiation occurred only after the disambiguating second syllable was spoken.

Other visual world studies have demonstrated looks to candidate entities that entirely preceded the relevant input (e.g., Altmann & Kamide, 1999; 2007; Kamide, 2008; Kamide, Scheepers & Altmann, 2003; Kamide, Scheepers, Altmann, & Crocker, 2002; Knoeferle, Crocker, Scheepers, & Pickering, 2005; Sussman & Sedivy, 2003). For instance, upon hearing a sentence fragment such as "*The girl will ride the*…" while viewing a scene depicting a man, a girl, a motorcycle, and a carousel, comprehenders looked toward the depiction of the carousel during the verb *ride*; conversely, upon hearing "*The man will ride the*…," they looked toward the motorcycle during the verb *ride* (Kamide, Altmann, & Haywood, 2003). Based on such research, we can conclude that the language parser is capable of combining visual context with noun/verb semantics to quickly narrow the possibilities for upcoming input.

Studies like these demonstrate one approach to examining prediction, although it might be argued that this is prediction in a very restricted sense (where candidates for upcoming input are limited, visually present, and highly constrained). We therefore turn our attention to studies using written stimuli, which—while forgoing the visual world paradigm's naturalistic environment—allow the examination of a broad range of linguistic structures beyond concrete referential phrases.

Effects of predictability

In our discussion so far, we have been careful to use the word *prediction* only for cases in which upcoming, unseen input in some way alters current processing. There is also, however, a substantial literature on the effects of "predictability" on linguistic processing. The basic intuition here is that some words could be predicted in advance from context; and even if the brain does not make this prediction in advance, once the word is encountered it might still be processed differently depending on its predict*ability*—how well it *could* have been predicted from context. Obviously, if the brain does make an explicit prediction in advance, then that may well alter its processing in a way that later creates predictability effects; but other mechanisms that do not involve prediction per se are also possible, and several have been proposed. In this section we briefly review theoretical and experimental evidence for predictability effects, as well as the potential mechanisms underlying them.

In the context of language, the most compelling theoretical reason for the brain to compute predictabilities is, as mentioned earlier, for use in disambiguation. Ambiguities of various sorts are pervasive in language, and resolving them requires combining evidence from semantic, syntactic, and other sources; experiments show that comprehenders are able to quickly bring all of these to bear (Hanna & Tanenhaus, 2004; MacDonald, 1993; MacDonald, Pearlmutter, & Seidenberg, 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Spivey et al., 2002; Tanenhaus et al., 1995; Trueswell, Tanenhaus, & Garnsey, 1994). This requires a common mechanism for evidence combination across different parts of the linguistic processing system, and probability theory (with Bayes' rule) is a natural, even optimal, fit (Jurafsky, 1996). In general, the system must presumably be flexible about what probabilities it can compute; picking the most likely sense of a word has different requirements from, for instance, picking the most likely parse of a sentence. For our discussion of prediction, the most relevant computation would be of the probabilities of current or upcoming individual words. But are these among the probabilities that the brain computes? Arguably, yes. Work on automatic speech recognition has shown that such single-word probabilities are exactly what are needed to accurately decode the noisy (thus ambiguous) acoustic speech signal.

Furthermore, experimental evidence shows that the processor is sensitive to such single-word probabilities. In ERP studies, the N400 component to semantically congruous words shows a graded sensitivity to those words' predictability in context (Kutas & Hillyard, 1984), while the P600 component responds differentially to high- versus low-likelihood syntactic forms (even when those forms are all acceptable; Hsu, 2009; Kaan, Harris, Gibson, & Holcomb, 2000; Osterhout, Holcomb, & Swinney, 1994), and even earlier effects of context on lexical access may be present as well (Sereno, Brewer, & O'Donnell, 2003). Such effects could potentially be glimpses of different aspects of the system described earlier at work. We return to the question of whether the N400's sensitivity to predict*ability* is truly predict*ive* in a later section.

Reading time studies are another source of evidence that the brain is sensitive to predictability. Less predictable words—those with lower probability given context—are read more slowly (Ehrlich & Rayner, 1981). This effect is sensitive to both semantic (Bicknell, Elman, Hare, McRae, & Kutas, 2008; Duffy, Henderson, & Morris, 1989; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Hare, McRae, & Elman, 2003; Morris, 1994) and syntactic (Levy, 2008; Staub & Clifton, 2006; Trueswell, Tanenhaus, & Kello, 1993) manipulations of predictability, and its underlying cause is not yet clear. For instance, Norris (2006) suggests that the reading time effect may arise in the sensory system, as a side effect of decoding noisy visual input. The idea is that to recognize words that have less top-down support (relative to their visual competitors), the brain must acquire more bottom-up sensory evidence before it can reach an acceptable level of certainty that the word has been correctly identified, and gathering more evidence requires looking longer. Levy (2008), on the other hand, proposes that the reading time effect arises because as each word is processed, each potential whole-sentence interpretation has some shift in likelihood, becoming more or less supported by the available data. If there is a cost to shifting the internal representation of these likelihoods, as measured in reading time, then words which cause a greater total likelihood shift will take longer to read. Finally, he shows mathematically that the size of this shift (measured in KL divergence; Kullback & Leibler, 1951) is determined by the word's predictability. These two theories are nonpredictive accounts of predictability effects: All of the affected processing takes place after the word itself is seen.

Another approach is to model the linguistic processor as maintaining a sorted, limited size list of candidate whole-sentence parses; then, increased reading time may occur either when it turns out that the appropriate parse was pruned from the list for having too low probability (Jurafsky, 1996; Narayanan & Jurafsky, 1998), or when new evidence turns a less preferred parse into the most likely possibility, triggering a shift in attention (Crocker & Brants, 2000; Narayanan & Jurafsky, 2002). Such models are weakly predictive in the sense that they explicitly marginalize over potential continuations to determine which parses are most likely and which can be pruned to reduce memory load. However, pruning can only explain effects in somewhat special cases (i.e., garden-path sentences), while attention shifts only occur at some relatively special points in a sentence.

Smith and Levy (2008), however, demonstrate predictability effects on reading time at arbitrary points in naturalistic text, and they argue that sensory factors (as per Norris) cannot explain all variations in processing time (Rayner, Liversedge, White, & Vergilino-Perez, 2003). They propose instead a motor control–inspired model of "optimal preparation," in which the processor does a graded amount of anticipatory processing for each potential continuation; the actual amount of processing is determined as a trade-off between the resources spent on this anticipatory processing (preparing to process a word quickly is costly) versus the increased processing efficiency achieved if this continuation is actually encountered (sometimes preparation is worth it). Such a model naturally expends more effort on preparing for those continuations which are more likely to be encountered, thus explaining the predictability effect on subsequent reading time of the continuation. By our classifications, this is a strongly predictive model, but one with little direct evidence to support it—we cannot observe this postulated anticipatory processing directly via behavioral measures. We therefore turn to discussion of recent electrophysiological work that bears on this question.

Event-related potential evidence for a "strong" form of prediction

As should be evident from the previous sections, linguistic prediction and predictability can be identified in a variety of ways. We will now present evidence for the "strong" form of prediction—that is, experimental evidence showing that specific linguistic items or their features are pre-activated at some time point prior to encountering the confirmatory bottom-up input. Such evidence turns out to be somewhat tricky to obtain, due to the challenge of identifying processing related to an item that has not yet occurred. In particular, one difficulty centers on the *prediction/integration* debate—in other words, the challenge of distinguishing facilitation effects at a target word as being due to that word being predicted versus it being simply easier to integrate upon receipt. A case in point is the N400 ERP component. N400 amplitude is sensitive to a variety of factors—including word frequency, repetition, concreteness, and orthographic neighborhood size, among others—and is especially large to nouns that do not semantically fit with their preceding context (e.g., the word *dog* in "*I take my coffee with cream and dog*"; Kutas & Hillyard, 1980). N400s are also evident in responses to all but the most highly expected of nouns, even when they do fit with a prior sentence context. However, despite the sensitivity of the N400 to offline semantic expectancy, it is impossible to determine whether variation in N400 amplitude to the eliciting word during online sentence processing means that readers are using context to generate expectancies for that upcoming item (*prediction* view) or whether they are forced by the word itself to devote more or fewer resources to integrating the word into the existing sentence representation (*integration* view). Clearly, an argument for information getting pre-activated would be strengthened if it could be demonstrated that predictions were being formulated prior to target words.

In one of the earlier ERP studies to argue for predictive language comprehension, Federmeier and Kutas (1999a) found that in highly constraining contexts (e.g., "*He caught the pass and scored another touchdown. There was nothing he* *enjoyed more than a good game of…*"), unexpected and implausible target nouns (e.g., *baseball*) that were categorically related to the contextually expected target endings (e.g., *football*) were processed more similarly at a semantic level (as reflected in N400 amplitudes) to the expected endings than they were to categorically unrelated unexpected nouns (e.g., *monopoly*). *Baseball* and *monopoly* are equally implausible endings in this context, and yet the brain response to these two endings in the N400 region (200–500 ms post-target onset) is different. Critically, this facilitation decreases as sentential constraint becomes weaker even as the plausibility of these categorically related anomalies increases, for example, "*Eleanor wanted to fix her visitor some coffee. Then she realized she did not have a clean cup/bowl/fork.*" Why might the within category violation (*baseball*) behave like a more plausible ending, eliciting a brain response closer to the expected ending? Federmeier and Kutas argue that it is because the language processing system is predicting in the strong sense; in other words, pre-activating perceptual-conceptual features of the expected ending, which is more likely to share these features with the within category violation (*baseball*) than the between category one (*monopoly*). Thus, while norming shows that these words are equally incongruous with the linguistic context, *baseball* is more congruous with the brain's predictions. Importantly, similar results obtain whether the sentences were read one word at a time at relatively slow rates or presented as natural speech (Federmeier, Mclennan, De Ochoa, & Kutas, 2002).

Van Petten, Coulson, Plante, Rubin, and Parks (1999) also found N400-based evidence for prediction during naturalistic speech processing in a study designed to determine whether semantic integration processes began before or only after complete identification of a spoken word. To that end, participants listened to sentences (e.g., "*It was a pleasant surprise to find that the car repair bill was only seventeen…*") that were completed by a highly constrained, expected, congruous completion (*dollars*), a semantically incongruous word that began with the same initial sound (phonemic similarity) as the expected congruous completion (*dolphins*), and a semantically incongruous word (*scholars*) that ended with the same final sound and thus rhymed with the expected congruous completion. The critical contrast is between the N400s elicited by the two incongruous endings, which are equally nonsensical in the sentence context. Although equivalent in amplitude, the N400s differed in their latencies, with the N400 to the incongruous rhyming endings (*scholars*) diverging much earlier from the relative positivity to the congruous ending (*dollars*) than the incongruous ending with an overlapping initial sound (*dolphins*). Moreover, the effect of context preceded a word's isolation point—that is, before sufficient acoustic information had accrued to determine the word's identity (by about 200 ms). This demonstrates that not only do listeners use context to disambiguate partial auditory input—confirming the visual world results described earlier—but, in addition, their N400 response to incongruity begins when they detect the deviation from the expected word, potentially before the incongruous word could be identified.

Results from these studies are difficult to reconcile with a purely integrative model of comprehension, implicating some form of neural pre-activation instead. In all these cases, however, the observed ERP effects were evident at the target words of interest, leaving them at least superficially open to the "oh, it's just integration" criticism. Perhaps more compelling evidence of pre-activation comes from designs in which the electrophysiological sign of prediction precedes the target word that was presumably being predicted. For instance, work by Wicha and colleagues (Wicha, Bates, Moreno, & Kutas, 2003; Wicha, Moreno, & Kutas, 2003; Wicha, Moreno, & Kutas, 2004) investigated linguistic expectation in a series of studies with Spanish sentences by focusing on ERPs to gender-marked articles preceding target nouns of particular gender classes. In separate experiments, native Spanish speakers either listened to or read moderately to highly constraining sentences that contained a gender-marked article followed by either a written word or an embedded line drawing. The word or line drawing target was either the expected (highly probable) continuation or a semantically incongruent continuation of the same gender class as the expected continuation. In half of the sentences, the gender of the article mismatched the gender of the following noun or picture, although participants were not explicitly informed about this. A set of sample Spanish stimuli with their English glosses follows:

> *Caperucita Roja llevaba la comida para su abuela en…(una/un) canasta/corona.*
> Little Red Riding Hood carried the food for her grandmother in a…basket/crown.
> (a) …*una*$_{[\text{feminine}]}$ *CANASTA*$_{[\text{feminine}]}$ *muy bonita* (gender match/semantically congruous)
> (b) …*una*$_{[\text{feminine}]}$ *CORONA*$_{[\text{feminine}]}$ *muy bonita* (gender match/semantically incongruous)
> (c) …*un*$_{[\text{masculine}]}$ *CANASTA*$_{[\text{feminine}]}$ *muy bonita* (gender mismatch/semantically congruous)
> (d) …*un*$_{[\text{masculine}]}$ *CORONA*$_{[\text{feminine}]}$ *muy bonita* (gender mismatch/semantically incongruous)

It is important to note that in the experimental sentences there was always a reasonable word continuation regardless of the grammatical gender of the (manipulated) article; in this case, *una CANASTA*$_{(\text{feminine})}$/BASKET vs. *un COSTAL*$_{(\text{masculine})}$/SACK. The interpretation of any ERP effect in response to the article hinges on the fact that there was nothing semantically or syntactically wrong with an article of either gender. Accordingly, if any article ERP effect was obtained, it must have reflected the language system's discord at receiving an article of one gender when it was expecting a noun (and accompanying article) of the other gender. The pattern of ERPs for both word and picture targets clearly confirmed this hypothesis, even if the specifics of the ERP effects for words and pictures differed. The language processing system had expectations, and noncompliance with those was reflected in a differential ERP pattern for the articles of one gender versus the other.

van Berkum, Brown, Zwitserlood, Kooijman, and Hagoort (2005) sought evidence of prediction for spoken words in Dutch with similar logic by manipulating gender marking (neuter or common) on prenominal adjectives—a feature controlled by the gender of the upcoming noun. Participants heard sentences with moderately predictable noun continuations, such as "*The burglar had no trouble at all locating the secret family safe. Of course, it was situated behind a big[noun-appropriate gender marking] but rather unobtrusive painting[neuter]/bookcase[common].*" The gender-marked inflectional suffixes on the target adjectives were either congruent with the gender of the highly expected target noun or were not, being instead of the opposite gender category and then followed by a less expected but still semantically congruent noun. In contrast to the Wicha studies, they did not include any gender mismatches between adjectives and their upcoming nouns. Within high constraint contexts, there was a differential ERP effect for adjectives with versus without the expected gender, which emerged at the point the predicted and unpredicted inflectional suffixes of the adjectives first diverged from each other (e.g., the Dutch word for "big" marked with neuter gender is *groot* vs. with common gender *grote*; the ERP effect began when their pronunciations begin to diverge). They concluded that the ERP effect at the gender-marked adjectives was primarily a syntactic one, and as this study and other work by their group (e.g., Otten, Nieuwland, & Van Berkum, 2007; Otten & van Berkum, 2008, 2009) attest, comprehenders do indeed use sentence context to form predictions.

DeLong, Urbach, and Kutas (2005) employed a logically similar experimental design in English to test for prediction of yet another feature of language. Their design took advantage of a phonological feature of the English language in which different indefinite articles, *a* and *an*, are used depending on the initial phoneme of the immediately following word. Devoid of gender, case marking, and specific semantic content, English indefinite articles offer a means for exploring linguistic prediction at the level of phonological word forms. Participants read sentences that ranged in constraint and were continued by a range of more or less expected indefinite article/noun pairings; for example, "*The day was breezy so the boy went outside to fly…a kite/an airplane…in the park*" in which *a kite* is highly expected and *an airplane*, although plausible, is less so. The primary focus, as in the Wicha et al. and van Berkum et al. studies, was not on the nouns but on the prenominal words—in this case, the articles. As expected based on the literature, there was a significant inverse correlation between N400 amplitudes in the ERPs to the target nouns (*kite/airplane*) and their offline cloze probabilities (with cloze defined as the percentage of respondents supplying a particular continuation for a context in an offline norming task; Taylor, 1953): the higher a word's cloze probability, the smaller its N400 amplitude. This correlation, however, is equally consistent with theoretical accounts based on integration difficulty or prediction. The same pattern of reliable inverse correlations in the N400 region of the ERP to the articles, however, is less theoretically accommodating. Although the article correlations were slightly lower (maximal *r*-values in high –.60's to low –.70's at posterior electrode sites), they could only be explained by a predictive account. Otherwise, would the brain respond differentially to *a* versus *an*? Certainly, they mean the same thing, and they are in principle equally easy to integrate with respect to the ongoing context. What they differ in is their phonological form and critically what that says about the initial sound of the upcoming noun. We maintain that these correlations demonstrate that readers were anticipating the most likely noun given the prior context (e.g., *kite*), and encountering the phonologically wrong article (*an*) for that noun affected ongoing sentence processing.

Similar to the nouns, the maximum N400 article-cloze correlations were not randomly distributed, but rather were clustered over the posterior regions of the head where N400 effects are typically the largest. So over these scalp sites, a large percentage of the variance of brain activity (N400 amplitude) for the indefinite articles was accounted for by the average probability that individuals would continue the sentence context with that article. In short, these data are clear evidence of prediction in language at a phonological level. These results indicate that people do use sentence context to form graded predications for specific upcoming words, and not just in highly constraining contexts. Even in moderately constraining sentence contexts at least one and perhaps multiple items seem to be pre-activated albeit to varying degrees.

Studies of ERPs in written and spoken language, then, have offered up relatively strong evidence for linguistic prediction at featural and lexical levels. Arguably even more striking findings about linguistic prediction have come from studies in which ERPs are recorded in combination with the visual hemifield technique. In the visual hemifield technique, target stimuli are presented a few degrees to the right or left of subject's fixation to expose only the contralateral hemisphere to that stimulus for the first approximately 10 ms (Banich, 2003). This mode of presentation provides the receiving hemisphere a brief head start that remarkably results in temporally extended lateralized processing differences in the two hemispheres, which by inference has been taken to reflect the different ways in which the two hemispheres deal with various linguistic factors. In studies of this type, Federmeier and colleagues (e.g., Federmeier, 2007; Federmeier & Kutas, 1999b, 2002; Wlotko & Federmeier, 2007) found that only with right visual field presentations (when target words were initially presented to the left hemisphere, or LH) did the pattern of ERPs resemble those with central visual presentation; ERP patterns with left visual field presentations (when targets were initially presented to the right hemisphere) resulted in a different ERP pattern. These results thus led them to propose that left hemisphere processing was biased toward semantic feature pre-activation (i.e., prediction) via top-down cues, whereas right hemisphere (RH) processing was characterized by more bottom-up processing in combination with a wait-and-see approach, operating via integration with working memory. These findings dovetail nicely with a longstanding and more general view of the LH functioning as the brain's "interpreter" (sometimes "confabulator"), hypothesizing patterns even when there are none; in contrast, the RH maintains a more veridical record of the world it perceives (Gazzaniga, 1985; Metcalfe, Funnell, & Gazzaniga, 1995; Wolford, Miller & Gazzaniga, 2000). The idea that the LH may respond more to perceived event probabilities (even if the event's occurrence is not immediate), while the RH may be less likely to generalize away from the input (Gazzaniga, 2000), is one that undoubtedly requires more exploration within the domain of linguistic pre-activation. It is of particular relevance if one takes the view that comprehension does not employ a special language processor, but rather is a "new machine built of old parts" (Bates & Goodman, 1997).

Taken together, these electrophysiological studies argue for the "strong" form of prediction by implicating a neural language parser that triggers word features (e.g., syntactic, conceptual, semantic) and forms in advance of their input. This work demonstrates that linguistic expectancies can emerge from contextual operators on semantic memory as sentential context accrues, in cases where candidate entities (or their depictions) are not physically present to aid the brain in narrowing the possibilities for likely continuations. And most importantly, these experiments are sensitive to the key factor for demonstrating that prediction is a routine part of real-time linguistic processing—that is, the pre-target *timing* of such effects.

Possible consequences to linguistic prediction

A flip side to anticipating upcoming language input is that there could (though according to some models, there needn't!) be some type of processing consequence—or even a landscape of processing consequences—for not encountering highly probable material. A possible example of such an effect, Federmeier, Wlotko, De Ochoa-Dewald, and Kutas (2007) observed a late positive (LP) ERP to low probability congruous sentence endings (relative to high probability ones) that completed highly but not weakly constraining contexts (e.g., "*He bought her a pearl necklace for her collection*" vs. "*He looked worried because he might have broken his collection.*") From these results they argued for a cost—perhaps reflecting inhibition or revision—upon encountering an unlikely, but plausible, word in a strongly predictive context. Similarly, DeLong, Urbach, and Kutas (2007) also observed a prolonged, late frontal positivity (500–1200 ms) to unexpected relative to expected nouns (e.g., *airplane* in "*The day was breezy so the boy went outside to fly…a kite/an airplane…*"). Moreover, they demonstrated that this graded late positivity to unexpected nouns varied as a function of the strength of expectancy (constraint) for the most predictable items. Taken together, these findings of late positive ERPs to unexpected nouns that increase with the degree of constraint violation strongly support the idea that when highly preactivated input is not received, some form of additional processing may be called for.

In contrast, there has been little behavioral evidence over the years for such sentence processing "costs" in terms of naming/lexical decision time latencies, a detail which has served for some as an important argument against linguistic prediction (e.g., Gough et al., 1981). However, a general problem with studies basing their "no prediction" arguments on "no cost" findings (e.g., a lack of an inhibitory effect) relates to the baseline conditions of comparison, that is, the supposedly "neutral" conditions against which the "cost" conditions are contrasted (e.g., Stanovich & West, 1983). The difficulty (if not impossibility) in determining what constitutes an appropriately neutral context, brings into question the weight of such conclusions, and indeed the specific binary contrasts typical of such studies may not be the only (or best) way to go about testing for "cost."

Comprehension theories that have included some type of processing "costs" have mainly posited them in terms of syntactic predictions. For instance, in Gibson's Dependency Locality Theory (Gibson, 1998) grammatical complexity in sentences is modeled in terms of memory and integration costs that arise from predictions generated by previous items, with such costs being a function of the distance between syntactically dependent elements. Additionally, various electrophysiological studies have proposed that at a minimum, the P600 (a late occurring positive-going ERP component) has an amplitude that increases as various syntactic aspects of sentence processing become more difficult (e.g., for syntactic disambiguation, Osterhout & Mobley, 1995; syntactic revision, Osterhout & Holcomb, 1992; syntactic integration difficulty, Kaan et al., 2000; or syntactic reanalysis, Hagoort, Brown, & Groothusen, 1993). Although "cost" may not be the right way of thinking about this component, it is certainly a different situation than, say, the N400 whose amplitude *decreases* as contextual facilitation from preceding semantic information increases.

Another hint that the P600 might be related to some sort of prediction violation comes from a number of recent studies that have found modulations of this ERP component to more semantic experimental manipulations. These "semantic P600s" have alternately been linked to costs for thematic role assignments (e.g., Hoeks, Stowe, & Doedens, 2004), detections of conflicts between semantic plausibility and syntactic requirements (e.g., Van Herten, Chwilla, & Kolk, 2006), and conflicting processing streams, including syntax versus semantics (e.g., Kuperberg, 2007). Though none of these lines of research directly implicates violation of general linguistic prediction as the possible source of such ERP patterns, an intriguing possibility is that these effects could potentially be related to those observed in the Federmeier et al. (2007) and DeLong et al. (2007) studies described earlier—an idea that undoubtedly warrants further investigation.

Whether referred to as a processing "cost" or a processing "consequence," we believe what is relevant in considering such ideas is that the brain may need to engage in some form of "extra" processing when, on the basis of constraint-based predictive comprehension, an accruing contextual representation must be overridden, revised, inhibited, or reanalyzed—at least, if not only, in cases where there is a strong lexical candidate that does not materialize. "Cost" also might not be the most apt term if an unexpected item triggers updating in a learning signal, where probability likelihoods are being adjusted for the future. Such learning might be considered a "cost" in the short term, but in the longer term, the comprehender would benefit by gaining an accurate model of their linguistic environment.

Our interpretation of a "misprediction consequence" is compatible with theories suggesting that such effects may be best accounted for in terms of cognitive control and conflict monitoring. Kolk and colleagues (Kolk, Chwilla, van Herten, & Oor, 2003; van Herten et al., 2006; Vissers, Kolk, van de Meerendonk, & Chwilla, 2008), for instance, have suggested that when conflicts emerge between incompatible sentential representations, reanalysis is initiated to check whether the conflict is due to processing error. Novick, Trueswell, and Thompson-Schill (2005) and Thompson-Schill (2005) suggest that such reanalysis might stem from the selection among competing representations based on task demands. Although none of these authors frame a hypothesis specifically in terms of pre-activation or prediction violation, we believe that our results, and others, are compatible with this proposal. Moreover, we propose that the relevant "conflicts" need not be ones of syntactic or even semantic violation, arising even when items or their features are pre-activated to varying degrees, and then disconfirmed by the actual (physical) input. This more domain-general proposal is also consistent with the view that the cognitive control mechanisms involved in sentence comprehension may be similar to those employed in more general conflict tasks like the Stroop task (e.g., Novick et al., 2005; Thompson-Schill, 2005; Ye & Zhou, 2008). In addition, the observation of the generally more frontal scalp distribution of Federmeier et al.'s and DeLong et al.'s LP effect is roughly consistent with imaging data implicating various frontal and prefrontal cortical areas in inhibition (e.g., Aron, Robbins, & Poldrack, 2004), error detection (e.g., Rubia, Smith, Brammer, & Taylor, 2003), and suppression of interfering memories (e.g., Anderson et al., 2004). While ERP patterns at the scalp do not allow for direct mappings to specific brain areas, these distributional similarities are nonetheless suggestive.

Conclusions

In this chapter, we have offered some evidence for implicit, probabilistic anticipatory language processing, which we have argued may be cost incurring when continuations are highly anticipated but not received. These findings stand in contrast to a more classical view of language comprehension, and brain processing in general, as being essentially bottom up, waiting for sensory input that is processed and eventually recruited for action. The research reviewed herein is more compatible with the new wave of neural models proposing that a unifying principle for brain operation is one of being more "proactive," "prospective," or "pre-experiencing" (Bar, 2007; Gilbert & Wilson, 2007; Schacter, Addis, & Buckner, 2007;). Under active brain accounts like these, neural processors are assumed to constantly be predicting upcoming input and monitoring the consistency of the anticipated and actual outcomes. This default mode of operation is proposed to occur across all domains, at sensory, motor, and cognitive levels. With respect to a more cognitive domain, Schacter and Addis (2007) have proposed that a crucial function of memory is to make information available for simulating future events. Under this model, it is unclear what the exact role of a semantic component is in constructing future scenarios; however, it seems that prediction in language processing fits nicely with models of predicting upcoming language input based on our stored mental representations in combination with contextual factors. And co-opting another idea from vision research (Enns & Lleras, 2008), it seems possible that some recent findings indicating a "cost" for pre-activation may be compatible with the idea of processing information that is inconsistent with some prevailing, pre-activated schema or expectation; in turn, the information triggering such discrepancies may ultimately be processed relatively slowly because the parser must start over restructuring a new contextual representation. These few examples highlight our belief that as we "look ahead" to continued exploration of prediction issues in the language domain, it will also be beneficial to "look around" and let our research be informed, shaped, and spurred by examinations within a larger framework of general brain processing, incorporating proposals of prediction from theories of human motor control (so-called forward models), from a variety of aspects of vision research, from judgment and decision making, and from episodic and semantic memory studies; indeed, scientists within these various domains are already doing just this! Without denying the uniqueness and seeming specialization of the human brain for comprehending and producing language, it seems that the door has been cracked wider for investigations of how predictive linguistic processing might better be understood in terms of how the brain more generally predicts. We maintain that the studies we have described here have served to tip the scale, such that anticipatory processing should no longer be considered a lingering question in the literature, but rather should be understood as a natural part of the way language is comprehended, with future investigations targeting the nature and consequences of linguistic preactivation.

Acknowledgments

This work was supported by NIH grants HD-22614 and AG-08313 to M. K. and by a Center for Research in Language Training Fellowship to N. J. S. We thank Kara Federmeier and Roger Levy for their helpful comments on earlier drafts; as ever, we are influenced by and grateful for thought-provoking discussions with Jeff Elman.

References

Abney, S. P. (1989). A computational model of human parsing [Special issue]. *Journal of Psycholinguistic Research*, *18*(1), 129–144.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*(4), 419–439.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*, 247–264.

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*(4), 502–518.

Altmann, G. T. M., & Steedman, M. J. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238.

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., et al. (2004). Neural systems underlying the suppression of unwanted memories. *Science*, *303*(5655), 232–235.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*(4), 170–177.

Banich, M. T. (2003). Interaction between the hemispheres and its implications for the processing capacity of the brain. In K. Hugdahl & R. J. Davidson (Eds.), *The asymmetrical brain* (pp. 261–302). Cambridge, MA: MIT Press.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280–289.

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, *12*, 507–584.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. xx–xx). New York: John Wiley & Sons.

Bicknell, K., Elman, J. L., Hare, M., McRae, K., & Kutas, M. (2008). Online expectations for verbal arguments conditional on event knowledge. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 2220–2225). Austin, TX: Cognitive Science Society.

Bloom, P. A., & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory and Cognition*, *8*, 631–642.

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, *34*, 774–806.

Carroll, J., & Bever, T. (1978). The perception of relations: A case study in the relation of knowledge and perception. In E. Carterette & M. Friedman (Eds.), *Handbook of perception: Vol. 7. Speech and language*. New York: Academic Press.

Crocker, M. W., & Brants, T. (2000). Wide-coverage probabilistic sentence processing. *Journal of Psycholinguistic Research*, *29*(6), 647–669.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 561–584.

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.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2007, March 29-31). *A cost to mispredicting: Effects of sentential constraint violations*. Poster presented at the 20th Annual CUNY Conference on Human Sentence Processing, La Jolla, CA.

Duffy, S. A., Henderson, J. M., & Morris, R. K. (1989). The semantic facilitation of lexical access during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 791–801.

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, *24*, 409–436.

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.

Enns, J. T., & Lleras, A. (2008). What's next? New evidence for prediction in human vision. *Trends in Cognitive Sciences*, *12*(9), 327–333.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491–505.

Federmeier, K. D., & Kutas, M. (1999a). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*, 469–495.

Federmeier, K. D., & Kutas, M. (1999b). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*, *8*(3), 373–392.

Federmeier, K. D., & Kutas, M. (2002). Picture the difference: Electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia*, *40*(7), 730–747.

Federmeier, K. D., Mclennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, *39*(2), 133–146.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing [Special issue]. *Brain Research*, *1146*, 75–84.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. Cooper & E. Walker (Eds.), *Sentence processing* (pp. xx–xx). Hillsdale, NJ: Erlbaum.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*(4), 291–325.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, *37*(1), 58–93.

Gazzaniga, M. S. (1985). *The social brain*. New York: Basic Books.

Gazzaniga, M. S. (2000). Cerebral specialization and interhemispheric communication: Does the corpus callosum enable the human condition? *Brain*, *123*(7), 1293–1326.

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

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. *Science*, *317*(5843), 1351–1354.

Gough, P. B., Alford, J. A., & Holley-Wilcox, P. (1981). Words and contexts. In O. J. L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (pp. 85–102). Hillsdale, NJ: Erlbaum.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing [Special issue]. *Language and Cognitive Processes*, *8*(4), 439–483.

Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive Science*, *28*(1), 105–115.

Hare, M., McRae, K., & Elman, J. L. (2003). Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Language*, *48*(2), 281–303.

Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, *19*(1), 59–73.

Hsu, H. (2009). A neurophysiological study on probabilistic grammatical learning and sentence processing (Doctoral dissertation). Retrieved July 1, 2010, from http://ir.uiowa.edu/etd/243.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science: A Multidisciplinary Journal*, *20*(2), 137–194.

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

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201.

Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass*, *2*(4), 647–670.

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.

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, *32*(1), 37–55.

Kamide, Y., Scheepers, C., Altmann, G. T. M., & Crocker, M. W. (2002, March). *Integration of syntactic and semantic information in predictive processing: Anticipatory eye-movements in German*. Paper presented at the 14th Annual CUNY Conference on Human Sentence Processing, New York.

Kintsch, W., & van Dijk, T. A. (1978). Towards a model of text comprehension and productions. *Psychological Review*, *85*, 363–394.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, *95*(1), 95–127.

Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, *85*(1), 1–36.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, *22*(1), 79–86.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax [Special issue]. *Brain Research*, *1146*, 23–49.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition*, *11*, 539–550.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–163.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, *32*, 692–692.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676–703.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, *244*, 522–523.

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science*, 189, 226–228.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*, 1–71.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*, 375–407.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*(3), 283–312.

Metcalfe, J., Funnell, M., & Gazzaniga, M. S. (1995). Right-hemisphere memory superiority: Studies of a split-brain patient. *Psychological Science*, *6*(3), 157–164.

Mitchell, D. C. (1982). *The process of reading: A cognitive analysis of fluent reading and learning to read*. New York: Wiley.

Mitchell, D. C., & Green, D. W. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, *30*, 609–636.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(1), 92–103.

Morris, R. K. (2006). Lexical processing and sentence context effects. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 377–402).

Narayanan S., & Jurafsky, D. (1998). Bayesian models of human sentence processing. In *Proceedings of the 20th Cognitive Science Society Conference* (pp. 84–90).

Narayanan, S., & Jurafsky, D. (2002). A Bayesian model predicts human parse preference and reading time in sentence processing. In S. B. T. G. Dietterich & Z. Ghahramani (Eds.), *Advances in neural information processing systems 14* (pp. 59–65). Cambridge, MA: MIT Press.

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, *113*(2), 327–357.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective and Behavioral Neuroscience*, *5*(3), 263–281.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 786–803.

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *34*(6), 739–773.

Otten, M., Nieuwland, M., & van Berkum, J. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience*, *8*(1), 89.

Otten, M., & van Berkum, J. J. A. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes*, *45*(6), 464–496.

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.

Pickering, M. J. (1994). Processing local and unbounded dependencies: A unified account. *Journal of Psycholinguistic Research*, *23*, 323–352.

Pritchett, B. L. (1991). Head position and parsing ambiguity [Special issue]. *Journal of Psycholinguistic Research*, *20*(3), 251–270.

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.

Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye movements. *Psychological Science*, *14*, 385–389.

Rubia, K., Smith, A. B., Brammer, M. J., & Taylor, E. (2003). Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *NeuroImage*, *20*(1), 351–358.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4, Part 1), 696–735.

Schacter, D. L., & Addis, D. R. (2007). On the constructive episodic simulation of past and future events. *Behavioral and Brain Sciences*, *30*(3), 331–332.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience*, *8*(9), 657–661.

Schwanenflugel, P. J. (1986). Completion norms for final words of sentences using a multiple production measure. *Behavior Research Methods, Instrumentation, and Computers*, *18*, 363–371.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*, 109–148.

Sereno, S. C., Brewer, C. C., & O'Donnell, P. J. (2003). Context effects in word recognition: Evidence for early interactive processing. *Psychological Science*, *14*(4), 328–333.

Smith, N. J., & Levy, R. (2008, June 23–26). *Optimal processing times in reading: A formal model and empirical investigation*. Paper presented at the 30th Annual Meeting of the Cognitive Science Society, Washington, DC.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, *45*(4), 447–481.

Stanovich, K. E., & West, R. F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory and Cognition*, *7*(2), 77–85.

Stanovich, K. E., & West, R. F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General*, *112*(1), 1–36.

Staub, A., & Clifton, C., Jr. (2006). Syntactic prediction in language comprehension: Evidence from either…or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 425–436.

Steedman, M. (1989). Grammar, interpretation, and processing from the lexicon. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. xx–xx). Cambridge, MA: MIT Press.

Sussman, R. S., & Sedivy, J. C. (2003). The time-course of processing syntactic dependencies: Evidence from eye movements. *Language and Cognitive Processes*, *18*(2), 143–161.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science*, *268*, 1632–1634.

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

Thompson-Schill, S. L. (2005). Dissecting the language organ: A new look at the role of Broca's area in language processing. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones*. (pp. 173–189). Mahwah, NJ: Erlbaum.

Traxler, M., Bybee, M., & Pickering, M. (1997). Influence of connectives on language comprehension: Eye-tracking evidence for incremental interpretation. *Quarterly Journal of Experimental Psychology A*, *50*(3), 481–497.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*(3), 285–318.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 528–528.

Tyler, L. K., & Marslen-Wilson, W. D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 683–692.

van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, *41*(2), 147–182.

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.

van Berkum, J. J. A., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*, *1146*, 158–171.

van Berkum, J. J. A., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, *17*(3), 701–718.

van Herten, M., Chwilla, D. J., & Kolk, H. H. J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, *18*(7), 1181–1197.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition*, *18*(4), 380–393.

Van Petten, C., Coulson, S., Plante, E., Rubin, S., & Parks, M. (1999). Timecourse of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 394–417.

Vissers, C. T. W. M., Kolk, H. H. J., van de Meerendonk, N., & Chwilla, D. J. (2008). Monitoring in language perception: Evidence from ERPs in a picture-sentence matching task. *Neuropsychologia*, *46*(4), 967–982.

Wicha, N. Y. Y., Bates, E., Moreno, E. M., & Kutas, M. (2003). Potatoes not Pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, *346*, 165–168.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2003). 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*, 483–508.

Wicha, N. Y. Y., Moreno, E., & 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.

Wlotko, E. W., & Federmeier, K. D. (2007). Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*, *45*(13), 3001–3014.

Wolford, G., Miller, M. B., & Gazzaniga, M. (2000). The left hemisphere's role in hypothesis formation. *Journal of Neuroscience*, *20*(6), RC64.

Ye, Z., & Zhou, X. (2008). Involvement of cognitive control in sentence comprehension: Evidence from ERPs. *Brain Research*, *1203*(8), 103–115.