Language and Linguistics Compass 8/12 (2014): 631–645, 10.1111/lnc3.12093

Pre-Processing in Sentence Comprehension: Sensitivity to Likely Upcoming Meaning and Structure

Katherine A. DeLong[1]*, Melissa Troyer[1] and Marta Kutas[1,2,3,4]
[1]*Department of Cognitive Science, University of California (UCSD)*
[2]*Center for Research in Language, University of California (UCSD)*
[3]*Department of Neurosciences, University of California (UCSD)*
[4]*Kavli Institute for Brain and Mind, University of California (UCSD)*

Abstract

For more than a decade, views of sentence comprehension have been shifting toward wider acceptance of a role for linguistic pre-processing – that is, anticipation, expectancy, (neural) pre-activation, or prediction – of upcoming semantic content and syntactic structure. In this survey, we begin by examining the implications of each of these 'brands' of predictive comprehension, including the issue of potential costs and consequences to not encountering highly constrained sentence input. We then describe a number of studies (many using online methodologies) that provide results consistent with prospective sensitivity to various grains and levels of semantic and syntactic information, acknowledging that such pre-processing is likely to occur in other linguistic and extralinguistic domains, as well. This review of anticipatory findings also includes some discussion on the relationship of priming to prediction. We conclude with a brief examination of some possible limits to prediction and with a suggestion for future work to probe whether and how various strands of prediction may integrate during real-time comprehension.

1. *Introduction*

When it comes to language comprehension, as the saying goes, everyone's an expert. Although not quite true (e.g., at various developmental stages or with certain clinical populations or disorders), the average language comprehender, with seemingly little effort, is able to rapidly detect and utilize diverse and subtle patterns of linguistic input. Although she can never know exactly what will come next, experience and knowledge in combination with current context serve as constant guides to constructing meaning out of novel combinations of words. In addition to the fundamental roles that previous and current input play in language processing, there is increasing evidence that *the future* also plays a part; namely, that pre-activation, prediction, anticipation, predictability and expectancy of various types and grains of linguistic information are part and parcel of comprehension.

For decades, notions of prediction were generally not included in language processing accounts, presumably due to the unboundedness of linguistic combinations (e.g., Jackendoff 2002; Morris 2006), modular views of language processing (e.g., Fodor 1983; Forster 1989), limited evidence of costs for mispredicting (e.g., Gough, Alford, and Holley-Wilcox 1981), and an initial reliance on offline methods lacking real-time temporal resolution. However, over the past 15 years or so, evidence for predictive language processing in spoken and written communication has become more widespread. In fact, Van Petten and Luka (2012) suggest that the conversation has shifted away from *whether* there is prediction in language processing to *how* such prediction occurs, including questions about the forms predictions can take, what it means to predict, and prediction's neural mechanisms. With sentence processing researchers increasingly acknowledging predictive processing models, and as we focus on experimental evidence supporting such interpretations in the current article, we feel obliged to point out that language processing is unlikely to proceed in either a strictly anticipatory (context exerting its influence prior to the receipt of target input) or strictly integrative (context exerting its influence only following receipt of target input) manner. In other words, under various circumstances, sentence comprehension likely arises through some combination of predictive and integrative mechanisms, which, from a practical standpoint, are difficult to disentangle experimentally (see Kutas, DeLong, and Smith 2011, for a discussion). Indeed, isolating prediction effects from ones indicating facilitated integration during online sentence comprehension requires both careful experimental design and temporally sensitive measures – both of which have been utilized more frequently in recent years to explicitly, rather than incidentally, explore pre-processing.

Although there are some obvious circumstances under which language processing would seem to benefit from look-ahead (e.g., ambiguity resolution, in noisy or degraded input situations, when input rate is uncontrollable, when attentional demands are high, for poor readers, etc.), prediction may naturally 'fall out' of how meaning is constructed. In any case, given that much of language's richness stems from encountering the unexpected (hilarious punch lines, revealing confessions, concepts enlightened through simple phrases), any full account of language processing must consider how the brain handles unexpected but informative input. Moreover, we think it likely that prediction 'foiled' by unexpected input has some utility in contributing to learning, redirecting attention, or tuning executive processes, among other possibilities.

In this review, we first lay out some terms and concepts related to prediction and then focus on semantic and syntactic prediction effects in sentence comprehension, cognizant of potential predictions of other language-relevant information (e.g., phonological, orthographic, pragmatic, speech pattern, or emotional information). We conclude with evidence for consequences of less 'successful' prediction (when supplemental processing may be recruited) and with a brief description of some situations in which, and individuals for whom, prediction may not function as reliably.

2. Concepts and Terminology

Above, we mention several prediction-related terms: *prediction*, *pre-activation*, *anticipation*, *predictability*, and *expectancy*. These terms are often used interchangeably; however, they are not necessarily understood synonymously among researchers, with each having different processing implications. Although *prediction* is probably the term used most generically, in its traditional sense, it was understood as the all-or-none process of activating a linguistic item (a word) in advance of perceptual input. This predictive process is presumably active, conscious, attention-demanding, explicit, and intentional, affording processing benefits but also incurring 'costs' when wrong (Posner and Snyder 1975). Dating back to early behavioral evidence, benefits to predicting (speeded lexical decision times and reduced naming times) were abundant, supplemented subsequently by eye-tracking, event-related brain potential (ERP), and magnetoencephalography (MEG) evidence: in contrast, effects of 'cost' have been inconsistent if not elusive.

More recently, the term (*neural*) *pre-activation* (Kutas and Federmeier 2000) has been used to refer to a type of prospective language processing that is faster-acting, less (if at all) conscious, and importantly, graded, with the possibility for diverse information (e.g., lexical word forms, categorical information, related items, and pragmatically linked information) to be activated to various degrees over varying time frames (e.g., immediately preceding a word, further back in a sentence, or over wider discourses). 'Landscapes' of pre-activation are dynamically recalculated and reshaped potentially by all available linguistic (and extralinguistic) information sources. There also may be costs to (strongly) pre-activating but not receiving input, although costs are not an essential part of this proposal.

Anticipation is often used interchangeably with *prediction* but also could be aligned with a *pre-activation* view in that more than one item may be anticipated from a given context. Van Petten and Luka (2012) explicitly distinguish *prediction* for specific upcoming lexical items from *anticipation/expectation* for semantic content generally. *Anticipation* may imply shaping of contextual representations at a coarse-grained level (e.g., for conceptual or situational information), with the potential for this to lead to *predictions* about upcoming linguistic input. Van Berkum (2009, p. 304) frames *anticipation* in terms of activations in long-term memory that 'increase the availability (readiness) of conceptually associated information' in response to context, perhaps via more passive, even 'dumb', processing mechanisms.

Predictability (sometimes *expectancy*) also can be divorced from prediction, in that an item can be *predictable* even if it is not *predicted*. *Predictability* can be understood as the likelihood that an element might have been predicted from a given context. A typical method for assessing a word's *predictability* is through offline (non-speeded) cloze probability tasks (Taylor 1953), although Levy (2008) discusses a related measure of *surprisal*, which is the negative log probability of a word given the preceding context, with *surprisal* being high when predictability is low.

To this prospective language processing terminology list, we might add *priming* – when exposure to a certain stimulus (prime) passively influences (e.g., via spreading activation) the response to a subsequent stimulus (target). Fast-acting priming effects rely on implicit memory processing, occurring outside of awareness for short durations. Traditionally, *priming* was considered to rely on mechanisms distinct from *prediction*, for which there was a proposed cost to processing unrelated word targets. *Priming* (and possibly *pre-activation*), on the other hand, both may lead to facilitation without costs. The major distinction between them seems to be that *pre-activation* is more generally considered to result from higher-level contextual representations built up from multi-word contexts on the fly, while *priming* is more frequently cited as resulting from 'pre-stored' linguistic units (e.g., words). Ultimately, their separability hinges on whether they are supported by the same, overlapping, or different mechanisms.

In this review, wherever possible, we adhere to the terminology used by individual authors themselves. However, going forward, our recommendation is for researchers to be as explicit as possible about the nature of the predictive processes under study, with the hope that such transparency may lead not only to more consistent cross-study labels but also to greater understanding of the precise neural/cognitive mechanisms involved.

3. *Predictive Processing of Lexico-Semantic Information*

In sentence constraint work, building on some of the early ERP studies of Kutas and colleagues (e.g., Kutas and Hillyard 1984; Kutas, Lindamood, and Hillyard 1984) and behavioral experiments of Schwanenflugel and colleagues (e.g., Schwanenflugel and Shoben 1985; Schwanenflugel and LaCount 1988), Federmeier and Kutas (1999) used online measures to explicitly argue for predictive semantic sentence processing. Utilizing a 'related anomaly' ERP paradigm, they inferred pre-activation of semantic features during sentence reading. They found that when high-constraint sentences targeted particular completions (e.g., *football*), categorically related but contextually implausible completions (e.g., *baseball*) showed reduced N400s (an ERP component whose amplitude decrease reflects an easing of semantic processing for that item) relative to implausible categorically unrelated words (e.g., *Monopoly*). In contrast, categorically related items in low-constraint contexts did not show similar N400 amplitude reductions. In constraining contexts, facilitation of the incongruent completion *baseball* was argued to occur via pre-activation of the semantic features of the expected but never encountered completion *football* that overlapped with *baseball*. It was inferred from the N400 results that word processing can benefit from *pre-activation* of semantic features. Here, however, there was no direct evidence of costs.

At around the same time, Altmann, Kamide, and colleagues (Altmann and Kamide 1999; Kamide, Altmann, and Haywood 2003) conducted influential eye-tracking studies utilizing the 'visual world' paradigm, showing that given a visual display of several objects, listeners moved their eyes in advance to those that satisfied selectional restrictions of sentential verbs. For instance, while hearing the sentence, '*The boy will eat/move the cake.*', at *eat* but not *move*, comprehenders moved their eyes to the only edible object (a cake) in the scene well before they encountered the word *cake*. Knoeferle, Crocker, Scheepers, and Pickering (2005), relying on ambiguous German case marking, similarly demonstrated preferential looks to upcoming depicted characters, even though their roles and relations in spoken sentences were atypical, indicating that observed anticipatory effects did not necessarily result from stored long-term representations of stereotyped relations. Though the presence of candidate entities in these 'visual worlds' weakens the case for strictly anticipatory sentence processing, these studies nonetheless indicate that listeners rapidly integrate linguistic information with visual scenes to establish thematic dependencies, which facilitate the pre-selection of likely continuations. These studies do not address the issue of costs.

Evidence for pre-activation at the level of specific words comes from another 'related anomaly' paradigm, this one utilizing non-word sentence completions. Laszlo and Federmeier (2009) used high-constraint sentence contexts (e.g., '*Before lunch he has to deposit his paycheck at the…*') with expected endings (e.g., *BANK*) that were alternated with pseudowords and illegal letter strings that were either (a) orthographic neighbors (e.g., *PANK*, *BXNK*) of the expected word or (b) not (e.g., *HORM*, *RQCT*). Although all non-words showed greater amplitude N400s relative to the expected words, orthographic neighbors – regardless of word status – showed reduced N400s relative to their non-neighbor counterparts. Pre-activated information for specific words thus seems to begin influencing bottom-up processing of input prior to word recognition and before the irregular completions are processed for meaning.

Additional evidence for pre-activation of semantic word content as well as phonological word forms comes from ERP studies isolating predictive effects to a pre-critical word, absent concurrent candidate continuations, as in visual world eye-tracking studies. DeLong, Urbach, and Kutas (2005), for example, adapted the grammatical gender-based paradigm of Wicha, Moreno, and Kutas (2004), similarly used by Van Berkum, Brown, Zwitserlood, Kooijman, and Hagoort (2005) and Szewczyk and Schriefers (2013), in Spanish, Dutch, and Polish, respectively. They capitalized on the *a/an* distinction in English (e.g., '*Anna did not like to say words with "s" because she spoke with a lisp/ an accent…*') to show that readers form graded expectations for upcoming nouns. For sentences with ranges in constraint, N400 amplitude to not only the more and less expected nouns (*lisp/accent*) but also the preceding indefinite articles (*a/an*), varied linearly with the words' offline cloze probabilities (*predictability*). The graded N400s to the articles – seen both in grand average and individual subject data – constituted strong evidence that readers were probabilistically pre-activating specific likely upcoming noun word forms, arguing for *pre-activation* in terms of dynamically shifting probability distributions for not-yet-encountered information (though see Van Petten and Luka 2012, who offer an account by which these results could be the by-product of multi-trial averaging instead[1]). Follow-up analyses of these data (DeLong, Groppe, Urbach, and Kutas (2013); DeLong, Urbach, Groppe, and Kutas 2011) also revealed a distinct post-N400 brain response associated with encountering unexpected continuations to constraining sentence contexts.

Per our descriptions of *priming* and *pre-activation*, it might be assumed that the simplest path to predictive semantic effects would be via passive, automatic lexical priming. Through associative neural networks of lexical information, automatic spreading activation (e.g., Collins and Loftus

1975) from a prime triggers information that – despite the absence of additional linguistic context – eases processing of a subsequent target. Posner and Snyder's (1975) theory includes such a fast unconscious facilitatory priming process; it also includes a slower-acting, strategic mechanism for semantic word priming under conscious control via an 'expectancy' route (in line with the traditional sense of *prediction*). Their model predicts facilitation for semantically related items with either route, but inhibition for unrelated prime-target pairs, resulting from the limited capacity processor devoting its resources to activating selectively related information, only on the expectancy route. Neely (1977) demonstrated that for prime-target stimulus onset asynchronies (SOAs) of <250 ms (in line with natural speech rates), lexical decision times to targets were facilitated for related word primes (*body-arm*) relative to neutral primes (*XXXX-arm*), but not significantly slowed for non-related pairs (*body-door*). For longer SOAs (e.g., >750 ms), however, there were facilitation effects for related pairs only when the overall relatedness proportion of word pairs was high – i.e., when a prediction strategy was more likely to be accurate. Additionally, at the longer SOA, improbable items were inhibited (slowed relative to neutral primes) even for semantically related prime-targets (e.g., *arm* is inhibited after *body*) when instructions were to expect 'building part' targets (e.g., *door*) following *body*. These findings (among others, e.g., Favreau and Segalowitz 1983) were taken to support dual prospective priming routes and generate testable predictions for other, online methodologies.

In a similar vein, Lau, Holcomb, and Kuperberg (2013) offer a view of prediction in which commitments (via consolidation in working memory) are made to not-yet-encountered semantic information activated via preceding contexts. This strategic mechanism, which is less automatic and more subject to immediate influences, is contrasted with one by which words, concepts or features are 'resonantly' (Myers and O'Brien 1998), or through passive spreading activation, triggered through stored associations in long-term memory, importantly *without* any such commitments. Critically, in line with Posner and Snyder (1975), only the strategic mechanism should incur costs for disconfirmed semantic commitments, presumably because the working memory representation requires updating. Lau et al. (2013) examined the contributions of these two mechanisms to N400 amplitude reduction in a word priming study using more (e.g., *bride-groom*) and less (e.g., *blubber-groom*) associated prime-target pairs in which the overall proportions of semantically related/unrelated word pairs (relatedness proportion) varied: either 10% or 50% of items in an experimental block were related pairs. Increasing global predictive validity (high relatedness, 50% block) led to larger N400 amplitude reductions to associated targets with different timing and topography than those in the globally less predictive condition (10% block). The greater N400 effect was attributed to a prediction-generating mechanism, even though no cost effects were observed.

It is unclear to what extent word pair priming studies like these help to understand predictive mechanisms of sentence comprehension, when notions of prediction outlined in Posner and Snyder (1975) – occurring strategically, at longer SOAs, and with costs (although not evidenced in Lau et al. 2013) – and notions of neural pre-activation – occurring rapidly, neither fully automatic nor fully strategic, and not necessarily cost-incurring – are so different. Sentences by their nature involve the rapid combination of novel sequences of words to form message-level representations not stored in semantic memory. Data from DeLong et al. (2005) argue for graded, probabilistic pre-activation, and findings from related anomaly paradigms (like Federmeier and Kutas 1999) show that pre-activation can occur outside of awareness. Neither example appears to reflect the conscious commitment of a predicted item to the contextual representation being constructed in working memory.

At the same time, results from DeLong et al. (2005) also indicate that sentential prediction effects are not simply a result of lexico-semantic priming (e.g., via spreading activation). None of the words in '*Anna did not like to say words with "s" because she spoke with a…*' is strongly associated with the most probable continuation, *lisp*, although in combination, they lead to its activation at least coincident with the article preceding it. Moreover, in a comparison of predictive discourse contexts including lexical associates of a probable critical noun to non-predictive 'prime control' sentences comprised of the same lexical associates, Otten and Van Berkum (2008) showed that only in the former was there a differential ERP response to gender-marked adjectives either consistent or inconsistent with upcoming critical nouns. Readers then, it seems, formed predictions based on message-level representations of unfolding text and not only from lower-level word-based priming (see also Van Petten 1993).

So where does this leave us? Early accounts of word priming propose that semantic information gets pre-activated (or inhibited) via distinct processing mechanisms which (a) are under more or less executive control, (b) have varying degrees of specificity, and (c) are more or less fast acting. Findings of graded, probabilistic predictive sentence processing using online methods appear more compatible with the automatic, fast-acting, less conscious, (at least sometimes) inhibition-free processing. However, other ERP studies of sentence processing have ruled out lexical priming as the sole basis for the observed semantic pre-activation. These different strands can be reconciled if, as Van Berkum (2009) notes, predictions are driven not only by individual words but also by the rapid convergence of different types of higher-order information. Under such a view, for instance, likely continuations for a sentence fragment such as '*The mother hoped the child might someday be…*' may be '*a doctor*' or '*a lawyer*'. However, candidate continuations are revised considerably if the *mother* is encountered in a 19th-century English novel and is of the landed gentry class, hoping for her daughter to one day be *married* or '*a lady*'. Or what if the *mother* is a convent's abbess (might '*a nun*' might be more probable)? Or the sentence is part of a rhyming couplet whose first line ends in *preacher* (such that a continuation like '*a teacher*' would be a good candidate)? In each of these instances, we suggest that the identical words of the target sentence effectively serve as cues to differential knowledge stores, with combinations of informational sources (including more comprehender-/speaker-internal, state-/trait-based influences, e.g., individuals' gender, class, age, accent, knowledge; mood, emotional or attentional state; event/situational knowledge; as well as more external factors, such as visual/sensory environment; task demands; available processing time) interacting rapidly to co-exert their influences in reducing or supplementing activation levels of representations under construction prior to receipt of all input. Such a proposal could entail a fast-acting predictive component far more complex than envisioned heretofore. This need not negate a deliberate, conscious, slower-acting predictive language mechanism, or even a spectrum of predictive processes engaged as a function of available limited capacity resources (for instance, under more or less attentional control). However, the path to empirically demonstrating such a range of processes is not necessarily a straightforward one, a case in point being the N400: as Kutas and Federmeier (2011) point out, the component's sensitivities have blurred the distinction between so-called automatic and controlled processes (i.e., those occurring outside of or only under conscious attention, respectively), with N400 effects evident under both conditions. With this in mind, it is worth considering what additional types of evidence support the notion of prediction during language processing. One avenue to explore is whether, under certain circumstances at least, there may be 'costs' to encountering unexpected information (as proposed under more explicit views of prediction). Although we know that the N400 is not always, if ever, sensitive to such costs, there is some evidence, which we will discuss in Section 5, that other ERP effects may be.

4. Pre-activation of Structural and Syntactic Information

Behavioral and neurophysiological evidence also point to rapid access to syntactic properties of language, i.e., information about high-level sentential structure. It is an open question whether the same neurocognitive mechanisms are at play during structural compared to semantic pre-activation. The product of syntactic pre-activation is predominantly non-lexical in nature and typically defined over categories of words. In theory, language users could predict parts of speech (e.g., nouns or verbs), verb tense or aspect, animacy, and/or gender, among other aspects of structural information. They also may predict larger syntactic structures, incorporating, for instance, phrase structure rules (e.g., VP → V NP; Jurafsky and Martin 2009). Language comprehenders also may be awaiting the conclusion of a particular syntactic structure (see, for example, Gibson's 1998, Syntactic Prediction Locality Theory, in which unfilled syntactic roles create processing costs until fulfilled). Here, we will review syntactic pre-activation of (a) word category and other ostensibly structural word features and (b) larger syntactic structures, without necessarily subscribing to any particular theory of syntax or positing separability of structural and lexico-semantic information at a representational level (see, e.g., Troyer, O'Donnell, Fedorenko, and Gibson 2011, for evidence of mixed representations).

Findings from Marslen-Wilson's (1973, 1975) shadowing experiments provide early evidence for the rapid use, if not pre-activation, of structural information in linguistic processing. Ninety-seven percent of the errors participants made when tasked with immediately repeating speech input aloud (shadowing) were semantically and syntactically congruent continuations. Moreover, for errors, their shadowing latencies typically had shorter durations compared to their average latencies, '…as if [participants were] placing more reliance on the predictive properties of the higher order context' (Marslen-Wilson 1973, p.523) than on remaining faithful to the speech input.

Electrophysiological data likewise point to rapid access to high-level structural information about linguistic content. Phrase structure violations, for example, elicit relatively early, lateralized, anterior ERP negativities (ELANs) compared to controls (Neville, Nicol, Barss, Forster, and Garrett 1991; Friederici, Pfeifer, and Hahne 1993). ELANs (from 50–250 ms), for instance, are obtained to ungrammatical versus grammatical uses of the word *of* in: '*Max's of proof the theorem*' versus '*Max's proof of the theorem*'.[2] On opposing views, the ELAN indicates recognition of ungrammatical input within <250 ms after completing visual perceptual processes or violation of expectancy for particular words' forms (e.g., Rosenfelt 2012), which may be correlated with different parts of speech (for instance, verbs often end with *-ed* or *-ing* and nouns with derivational suffixes such as *-tion*).

MEG data likewise index early effects of processing non-canonical word category information (Dikker, Rabagliati, Farmer, and Pylkkanen 2010; Dikker, Rabagliati, and Pylkkänen 2009). Dikker and colleagues (2010), for example, contrasted MEG responses to well-formed versus anomalous sentences: target nouns were either expected (following adjectives) or not (following adverbs). Additionally, target nouns could (1a) end in a closed-class derivational morpheme indicating their noun identity (e.g., *-ess*), (1b) be word forms with canonical properties of nouns (e.g., ending in *-a*, as in *soda*, or *-ie*, as in *movie*), or (1c) be word forms with more neutral lexical properties (e.g., *infant*, a noun whose lexical properties are consistent with both nouns and verbs).

(1a) Bimorphemic noun	*The (beautiful / beautifully) princess was painted.*
(1b) Typical noun	*The (tasteless / tastelessly) soda was marketed.*
(1c) Neutral noun	*The (cute / cutely) infant was dressed.*

Crucially, Dikker and colleagues found that for the conditions where nouns had typical word forms, i.e., (1a) and (1b), but not (1c), an early visual MEG component peaking around 120 ms was larger to nouns following adverbs (ungrammatical) than adjectives. This study not only demonstrates the rapidity with which the brain accesses syntactic word form information and detects an anomaly therein but also implies a role for the predictability of an upcoming word's form.

A perhaps stronger claim for pre-activation of syntactic word category or word form information would involve effects of syntactic expectedness *before* a predicted word's appearance or possibly in the *absence* of a predicted word. To that end, Lau, Stroud, Plesch, and Phillips (2006) manipulated word category predictability, using properties of ellipsis to form sentences where nouns were more or less likely following genitives (e.g., *Mary's*). For instance, in (2a), the clause '*Although Erica kissed Mary's mother*' makes ellipsis of a noun possible; encountering *Dana's* (which can mean *Dana's mother*) obviates a subsequent noun. In (2b), however, the lack of a parallel genitive in '*Although the bridesmaid kissed Mary*' necessitates a noun phrase following *Dana's* for the sentence to remain grammatical, likely increasing the expectation for a noun in this position.

(2a)	Ungrammatical, ellipsis-possible:	Although Erica kissed Mary's mother, she did not kiss Dana's *of* the bride.
(2b)	Ungrammatical, No-ellipsis-possible:	Although the bridesmaid kissed Mary, she did not kiss Dana's *of* the bride.

They observed a larger ELAN in (2b) relative to (2a) at *of* following *Dana's* – i.e., to the no-ellipsis-possible condition where there was high constraint for a (not supplied) noun phrase. Such findings suggest that people are sensitive to properties of ongoing syntactic structure-building, including upcoming words' categories.

Structural features of highly predictable content words in sentences also may be available prior to encountering them. Similar to DeLong et al. (2005), several experiments have probed the availability of predictable structural cues, relying on constraining sentence contexts continued by articles or adjectives with inflected features (e.g., grammatical gender or animacy) consistent or not with upcoming predictable nouns. For spoken and read Spanish sentences, Wicha and colleagues (2003, 2004) found ERP differences to prenominal articles (e.g., un_{MASC}, 'a') whose grammatical gender was incompatible with that of not-yet-encountered best continuation (BC) nouns (e.g., $canasta_{\mathrm{FEM}}$, 'basket') relative to prediction-consistent articles (una_{FEM}), suggesting that grammatical gender was available prior to the presumably expected word. Van Berkum and colleagues (2005) employed a similar paradigm in spoken Dutch with critical nouns preceded by adjectives that matched either (a) BC nouns or (b) plausible but less probable and alternative gender nouns. They found an increased positivity (approximately 50–250 ms after the inflectional onset) at the adjectives for gender mismatches. Szewczyk and colleagues (2013) found that violations of inflection in Polish based on animacy yielded a greater negativity from 400–600 ms compared to gender- and animacy-consistent prenominal adjectives.

Behavioral studies also have addressed how other syntactic properties may guide linguistic prediction. For instance, Ilkin and colleagues (2011) manipulated the likelihood for upcoming nouns to be plural, contrasting sentence beginnings such as '*It was to each other that the…*' (high plural likelihood) with ones such as '*It was to John and Mary that the…*' (low plural likelihood). Participants spent less time reading, and were more likely to skip over, plural noun phrases following reflexive constructions, suggesting that predictable syntactic features (such as number) were used to generate syntactic expectancies. Other eye-tracking and self-paced reading studies have found similar effects of syntactic predictors leading to facilitated processing for number (Lau, Rozanova, and Phillips 2007) and grammatical agreement (Vainio, Hyönä, and Pajunen 2003).

Evidence for pre-activation of larger syntactic structures comes primarily from behavioral data. Processing words in ambiguous syntactic structures, such as main verb vs. reduced relative clause readings of a verb, is difficult (Ferreira and Clifton 1986; Trueswell, Tanenhaus, and Garnsey 1994). One account is that sentence representations are not updated because multiple syntactic parses make prediction of a single syntactic representation impossible (Gibson 2006). Recent psycholinguistic models that quantify prediction in terms of *surprisal* (Hale 2001) make the counterintuitive hypothesis that sometimes ambiguity *facilitates* processing because an upcoming word may be successfully integrated in multiple parses of a sentence (Levy 2008). For instance, for '*The son of the colonel who shot himself…*', participants are faster to read the ambiguous reflexive pronoun *himself* (which could refer to the *son* or the *colonel*) than in '*The daughter of the colonel who shot himself…*', where the pronoun is disambiguated (Traxler, Pickering, and Clifton 1998). Such findings extend to unambiguous sentences in which syntactic expectations are constrained. For example, Staub and Clifton (2006) found that participants were faster to read regions immediately following the word *or* in a coordinate noun or sentence structure preceded by *either*, demonstrating a sensitivity to a predictable syntactic structure beyond the word.

Priming of syntactic structure in comprehension is closely related to the notion of pre-activation. Although syntactic priming has been examined more closely in production (Bock 1986; for a review, see Ferreira and Bock 2006), priming has also been observed in comprehension, both when prime and target structures share lexical items in addition to structure (see Tooley and Traxler 2010, for a review), as well as when there is no lexical overlap between primes and targets (Thothathiri and Snedeker 2008). For example, hearing a double-object (DO) construction (e.g., '*Feed the zebra the candy*') leads to facilitated processing for other DO constructions (e.g., '*Show the horse the book*') compared to a prepositional-object construction ('*Show the horn to the dog*') and vice versa (Thothathiri and Snedeker 2008).

Various accounts of syntactic priming place different levels of importance on transient activation (Branigan, Martin, and Cleland 1999) or longer-term implicit learning (Bock and Griffin 2000; Fine and Jaeger 2013). Importantly, however, both accounts posit changes in baseline activation levels of structures after recently encountering them. Fine and colleagues (Fine, Jaeger, Farmer, and Qian 2013) had participants read, at their own pace, sentences containing low-frequency relative clauses (RCs) which were signaled unambiguously (e.g., '*Several angry workers who were told about low wages…*') or where the verb was ambiguous, in that it could be interpreted either as a RC or main verb (e.g., '*Several angry workers warned about low wages…*'). Across the experiment, reaction time differences between ambiguous and unambiguous RCs decreased, supporting what was called *syntactic adaptation*, or convergence of expectations toward local probabilities. Further, a similar ambiguity effect *increased* for the alternate syntactic structure (main verb interpretation) after many RCs had been encountered, suggesting that comprehenders adapted to local statistics and that typically more-frequent structures became less preferred. Such findings suggest that local and extended contexts can perturb ongoing activation states of upcoming syntactic 'chunks'.

In sum, although there is growing evidence for pre-activation of structural features/components, demonstrations of pre-activation of larger syntactic units have been less forthcoming. Understanding what might be pre-activated in syntax necessitates an understanding of what constitutes psycholinguistically real syntactic units – a goal that remains a challenge. Chunks of syntactic information stored in memory may, for example, depend on their frequency of usage (Troyer et al. 2011), and very frequent multi-word strings may also be 'chunked' for storage (e.g., Arnon and Snider 2010; Bannard and Matthews 2008; Tremblay, Derwing, Libben, and Westbury 2011). Just as words take on different meanings in different contexts, pieces of syntactic structure are tied to the contexts in which they occur (as proposed in, for instance, construction grammars; Goldberg, 2006). Therefore, structural prediction may operate over different sized 'chunks'.

5. Consequences and Limitations to Predictive Processing

As we have mentioned, processing costs for disconfirmed predictions are an important part of some anticipatory comprehension models. The paucity of slowed reading or increased naming times for constrained input substituted with alternative sentence continuations was one reason why predictive comprehension models were traditionally dismissed (e.g., Stanovich and West 1983; Fischler and Bloom 1979; see Neely 1991, for a review of similar findings; but c.f., Schwanenflugel and Shoben 1985, for evidence of slowed lexical decision times to delayed-presentation unexpected endings of highly constraining sentences). So-called ‘inhibition’ effects – slowed overt responses to unlikely sentence continuations – however, appear fickle in their elicitation: they are subject to both target word delay from the sentence fragment (e.g., Traxler and Foss 2000) as well as the choice of ‘neutral’ condition used for comparison (e.g., Stanovich and West 1983). More current *pre-activation* views allow that graded, incremental, unconscious, multilevel predictions may be dynamically recomputed *without* consequence or penalty, such that lexical inhibition is not required in such models. Nonetheless, some recent data, using alternative methods, have indeed uncovered possible costs, at least under some circumstances.

For instance, as syntactic expectations change, behavioral consequences for processing less- or more-expected information also may change. A case in point, Fine et al. (2013) observed increased reading times for a frequent syntactic structure compared to a globally infrequent structure when proportions of these constructions within the experiment were reversed. Readers thus were able to capitalize on the local frequencies of complex syntactic structures to rapidly shift their expectations to conform to local information.

Pre-activating but not encountering sentence input, especially in highly constraining contexts, sometimes has led to a class of ERP effects that has been linked to costs. Van Petten and Luka (2012) catalog post-N400 late ERP positivities associated with reanalysis or repair following impaired interpretation due to syntactically or semantically unexpected input. In particular, they describe a late anterior positive ERP deflection dissociable from the more established posterior/parietal late positive component (LPC or P600). Support for this component’s sensitivity to prediction violations comes from its elicitation by plausible unexpected continuations of highly constraining contexts. Kutas (1993) noted a larger left hemisphere frontal post-N400 positivity to low relative to high cloze congruent sentence endings in high-constraint frames (75% and higher), suggesting that it might index inhibition of expected, but not presented, words. Federmeier, Wlotko, De Ochoa-Dewald, and Kutas (2007), as well, observed a similar anterior potential to congruent low cloze continuations of high but not low constraint sentences. Thornhill and Van Petten (2012) observed the anterior positivity to plausible words both semantically similar and dissimilar to expected lexical sentence continuations. DeLong, Urbach, Groppe, and Kutas (2011) reported frontal positivities to congruent low relative to high cloze continuations in predictive sentences (see also Moreno, Federmeier, and Kutas 2002). In each of these cases, a prediction-violating word plausibly continues the prior context; thus, an alternative continuation’s congruency/contextual interpretability may be a delimiting factor for eliciting this potential and may explain why the effect has not been observed to semantic anomalies.

The prediction-related functionality of these late anterior ERP effects is an open question. Such patterns need not reflect inhibition or revision, but might. Alternatively, they may represent resource allocation to ‘shift frames’ (e.g., Coulson 2001; Wlotko and Federmeier 2012) to bolster weakly activated representations or activate/integrate new contextual representations. Other candidate processes include conflict monitoring, attentional switching, lexical suppression, reanalysis or revision of contextual representations, or updating of a learning mechanism. Minimally, these data patterns suggest recruitment of additional processing when linguistic information is contextually pre-activated but not encountered.

Finally, we note that predictive language processing may be limited across the lifespan and by various individual factors. Older adults (age 60+ years) generally appear less likely than their younger counterparts to engage in pre-activating information during sentence processing (e.g., Federmeier, Kutas, and Schul 2010; Federmeier, McLennan, de Ochoa, and Kutas 2002; Wlotko, Federmeier, and Kutas 2012; DeLong, Groppe, Urbach, and Kutas 2012). Older adults who maintain relatively high verbal fluency, however, exhibit predictive patterns indistinguishable from younger adults (Federmeier et al. 2002, 2010; DeLong et al. 2012). At the opposite end of the lifespan, Borovsky, Elman, and Fernald (2012) and Mani and Huettig (2012) show that children with smaller vocabulary size are less likely to rapidly anticipate probable sentence continuations, as inferred from eye tracking, than those with larger vocabularies. There also seem to be limits to predictive processing in second languages, e.g., Martin et al. (2013) found that native Spanish speaking late learners of English do not show predictive ERP effects at prenominal articles (*a/an*) as in DeLong et al. (2005).

In sum, evidence for consequences to prediction has been less consistent than for predictive processing itself. This might be due, in part, to 'consequences' or 'costs' manifesting further downstream, affecting memory, learning, or other higher level processes. Such effects may also be subject to more individual variation (even in literacy) or to experimental design or task factors whose influence is not yet understood. A lack of evidence for 'costs', however, is no longer considered fatal to predictive accounts of language comprehension.

6. Conclusions

Herein, we have detailed evidence for different types and levels of predictive processing during sentence processing. One of our goals was to emphasize that studies describing *prediction* should be read with an eye toward the particular theories behind different research groups' use of the term, particularly in regard to proposals for cost-free versus cost-incurring processes. Although we focused on predictive processing of semantic and syntactic information, we believe that other types of linguistic information also may concurrently be pre-activated and that wide-ranging internal (stored memory representations) and external information sources influence these processes. Indeed, a question going forward is how, when, or under which circumstances predictions at different levels interact (or not) during real-time language comprehension. Smith and Levy (2013), for instance, point out how a word completing a statement of an obvious fact (e.g., '*The sun is in the sky*.') may, based on cloze probability ratings, reading times or ERPs, be considered highly predictable; however, this continuation is so obvious that unless an interlocutor is an early or second language learner, it is pragmatically unlikely. With the recognition that theories of language comprehension will benefit from a better understanding of how predictions are made and revised and how they facilitate language processing, one important goal will be to determine how different strands of predictive processing integrate online, grounded in a variety of contexts (word pairs, phrases, sentences, discourses, conversations, learning environments, etc.). With this, we may also better grasp the timing and conditions under which potential consequences of prediction emerge.

Acknowledgement

Support from NICHD grant HD22614 to M. K. and from the NSF Graduate Research Fellowship and UCSD Kroner Fellowship to M.T. is gratefully acknowledged. Special thanks to Tom Urbach for helpful comments on an earlier version of this manuscript and to Nathaniel Smith for insightful discussions.

Short Biographies

Katherine DeLong's research primarily utilizes event-related brain potential measures to explore questions relating to semantic language comprehension, and in particular, the anticipatory nature of sentence processing and its consequences. In these areas, she has co-authored papers for journals such as Nature Neuroscience, Brain and Language, and Psychophysiology, as well as several scholarly book chapters. She obtained her MS and PhD in Cognitive Science from the University of California, San Diego, where she currently continues her research as a postdoctoral scholar in Professor Marta Kutas' laboratory.

Melissa Troyer's research interests include human memory, language comprehension, and, most importantly, how they work together. Her work has appeared in the proceedings of the Cognitive Science Society and in the Journal of Undergraduate Neuroscience Education. Melissa is currently a PhD student working in Dr. Marta Kutas' cognitive electrophysiology lab at the University of California, San Diego, where she holds a Dean's Fellowship, the Kroner Award. She is also the recipient of both the National Science Foundation Graduate Research Fellowship and the National Defense Science and Engineering Fellowship. She holds a master's degree in cognitive science from MIT. Prior to graduate studies, Melissa obtained a BA in French and in Linguistics, a BS in Psychology, and a BS in Cognitive Science at Indiana University, Bloomington.

Marta Kutas is Professor and Chair of the Cognitive Science Department and Director of the Center for Research in Language at the University of California, San Diego. She received her PhD in Biological Psychology from the University of Illinois, Champaign-Urbana. Kutas uses electrophysiological (event-related brain potentials) and behavioral techniques to investigate sentence processing, and human information processing more generally. More specifically she has been working on pre-activation, quantification and negation, reading preview effects, word learning, hemispheric asymmetries, metaphor processing, and the role of mood on language processing.

Notes

* Correspondence address: Katherine A. DeLong, Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive La Jolla, CA 92093-0515, USA. E-mail: kadelong@cogsci.ucsd.edu

[1] Urbach and Kutas (2013) address this issue directly in a timed cloze-norming study, examining the relationship of group contextual constraint with individuals' response times to single trials. Their findings offer support that group average relative frequencies and response speeds adequately model the majority of individual responses. Smith and Levy (2013) also address this point, with results that offer strong evidence specifically for gradation within single trials.

[2] Note that some researchers have criticized several ELAN studies, suggesting that differences in processing words just prior to the target words may have resulted in baseline issues when comparing targets (see Steinhauer and Drury 2012, for a review).

Works Cited

Altmann, G., and Y. Kamide. 1999. Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 73(3) 247–64.

Arnon, I., and N. Snider. 2010. Frequency effects for multi-word phrases. *Journal of Memory and Language* 62. 67–82.

Bannard, C., and D. Matthews. 2008. Stored word sequences in language learning: the effect of familiarity on children's recognition of four-word combinations. *Psychological Science* 19(3) 241–8.

Bock, J. K. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18. 355–87.

Bock, J. K., and Z. M. Griffin. 2000. The persistence of structural priming: transient activation or implicit learning? *Journal of Experimental Psychology: General* 129(2) 177–92.

Borovsky, A., J. L. Elman, and A. Fernald. 2012. Knowing a lot for one's age: vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology* 112(4) 417–36.

Branigan, H. P., J. P. Martin, and A. A. Cleland. 1999. Syntactic priming in written production: evidence for rapid decay. *Psychonomic Bulletin & Review* 6(4) 635–40.

Collins, A. M., and E. F. Loftus. 1975. A spreading-activation theory of semantic processing. *Psychological Review* 82(6) 407.

Coulson, S. 2001. *Semantic leaps: frame-shifting and conceptual blending in meaning construction*. New York and Cambridge: Cambridge University Press.

DeLong, K. A., D. M. Groppe, T. P. Urbach, and M. Kutas. 2012. Thinking ahead or not? Natural aging and anticipation during reading. *Brain and Language* 121(3) 226–39.

DeLong, K. A., T. P. Urbach, D. M. Groppe, and M. Kutas. 2011. Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology* 48(9) 1203–7.

DeLong, K. A., T. P. Urbach, and M. Kutas. 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience* 8(8) 1117–21.

Dikker, S., H. Rabagliati, T. A. Farmer, and L. Pylkkanen. 2010. Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science* 21(5) 629–34.

Dikker, S., H. Rabagliati, and L. Pylkkänen. 2009. Sensitivity to syntax in visual cortex. *Cognition* 110(3) 293–321.

Favreau, M., and N. S. Segalowitz. 1983. Automatic and controlled processes in the first-and second-language reading of fluent bilinguals. *Memory & Cognition* 11(6) 565–74.

Federmeier, K. D., and M. Kutas. 1999. A rose by any other name: long-term memory structure and sentence processing. *Journal of Memory and Language* 41(4) 469–95.

Federmeier, K. D., M. Kutas, and R. Schul. 2010. Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language* 115(3) 149–61.

Federmeier, K. D., D. B. McLennan, E. de Ochoa, and M. Kutas. 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) 133–46.

Federmeier, K. D., E. W. Wlotko, E. De Ochoa-Dewald, and M. Kutas. 2007. Multiple effects of sentential constraint on word processing. *Brain Research* 1146. 75–84.

Ferreira, V. S., and K. Bock. 2006. The functions of structural priming. *Language and Cognitive Processes* 21(7–8) 1011–29.

Ferreira, F., and C. Clifton. 1986. The independence of syntactic processing. *Journal of Memory and Language* 25. 348–68.

Fine, A. B., and T. F. Jaeger. 2013. Evidence for implicit learning in syntactic comprehension. *Cognitive Science* 37. 578–91.

Fine, A. B., T. F. Jaeger, T. A. Farmer, and T. Qian. 2013. Rapid expectation adaptation during syntactic comprehension. *PLoS ONE* 8(10) 1–18.

Fischler, I., and P. A. Bloom. 1979. Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior* 18(1) 1–20.

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

Forster, K. I. 1989. Basic issues in lexical processing. *Lexical representation and process*, ed. by W. Marslen-Wilson, 75–107. Cambridge, MA: The MIT Press.

Friederici, A. D., E. Pfeifer, and A. Hahne. 1993. Event-related brain potentials during natural speech processing: effects of semantic, morphological, and syntactic violations. *Cognitive Brain Research* 1. 183–92.

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

——. 2006. The interaction of top-down and bottom-up statistics in the resolution of syntactic category ambiguity. Journal of Memory and Language 54. 363–88.

Goldberg, A. 2006. *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.

Gough, P. B., J. A. Alford, and P. Holley-Wilcox. 1981. Words and contexts. *Perception of print: reading research in experimental psychology*, ed. by O. J. L. Tzeng and H. Singer, 85–102. Hillsdale, NJ: Erlbaum.

Hale, J. 2001. A probabilistic Earley parser as a psycholinguistic model. *Proceedings of NAACL* 2. 159–66.

Ilkin, Z., and P. Sturt. 2011. Active prediction of syntactic information during sentence processing. *Dialogue and Discourse* 2(1) 35–58.

Jackendoff, R. 2002. *Foundations of language: brain, meaning, grammar, evolution*. New York: Oxford University Press.

Jurafsky, D. and J. H. Martin. 2009. Formal grammars of English. *Speech and language processing: an introduction to natural language processing, speech recognition, and computational linguistics*, 2nd edition, ed. by D. Jurafsky and J. H. Martin. Prentice-Hall. Upper Saddle River, NJ: Pearson Prentice Hall.

Kamide, Y., G. Altmann, and L. Haywood. 2003. The time-course of prediction in incremental sentence processing: evidence from anticipatory eye movements. *Journal of Memory and Language* 49(1) 133–56.

Knoeferle, P., M. W. Crocker, C. Scheepers, and M. J. Pickering. 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.

Kutas, M. 1993. In the company of other words: electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes* 8(4) 533–72.

Kutas, M., and K. D. Federmeier. 2000. Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences* 4(12) 463–70.

——. 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annual Review of Psychology 62. 621–47.

Kutas, M., K. A. DeLong, and N. J. Smith. 2011. A look around at what lies ahead: prediction and predictability in language processing. *Predictions in the Brain: Using our Past to Generate a Future*, Editor M. Bar, New York, NY: Oxford University Press, 2011, pp. 190–207.

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

Kutas, M., T. E. Lindamood, and S. A. Hillyard. 1984. Word expectancy and event-related brain potentials during sentence processing. *Preparatory states and processes*, ed. by S. Kornblum and J. Requin, 217–37. Hillsdale, New Jersey: Lawrence Erlbaum.

Laszlo, S., and K. D. Federmeier. 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) 326–38.

Lau, E. F., P. J. Holcomb, and G. R. Kuperberg. 2013. Dissociating n400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience* 25(3) 484–502.

Lau, E. F., K. Rozanova, and C. Phillips. 2007. Syntactic prediction and lexical frequency effects in sentence processing. University of Maryland Working Papers.

Lau, E., C. Stroud, S. Plesch, and C. Phillips. 2006. The role of structural prediction in rapid syntactic analysis. *Brain and Language* 98. 74–88.

Levy, R. 2008. Expectation-based syntactic comprehension. *Cognition* 106(3) 1126–77.

Mani, N., and F. Huettig. 2012. Prediction during language processing is a piece of cake-but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance* 38(4) 843–7.

Marslen-Wilson, W. 1973. Linguistic structure and speech shadowing at short latencies. *Nature* 244. 522–3.

——. 1975. Sentence perception as an interactive parallel process. Science 189(4198) 226–8.

Martin, C. D., G. Thierry, J. R. Kuipers, B. Boutonnet, A. Foucart, and A. Costa. 2013. Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language* 69(4) 574–88.

Moreno, E., K. Federmeier, and M. Kutas. 2002. Switching languages, switching palabras (words): an electrophysiological study of code switching. *Brain and Language* 80(2) 188–207.

Morris, R. K. 2006. Lexical processing and sentence context effects. *Handbook of psycholinguistics*, 2nd edition. ed. by M. J. Traxler and M. A. Gernsbacher. London: Academic Press.

Myers, J. L., and E. J. O'Brien. 1998. Accessing the discourse representation during reading. *Discourse Processes* 26(2–3) 131–57.

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.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: a selective review of current findings and theories. *Basic Processes in Reading: Visual Word Recognition* 11, 264–336.

Neville, H., J. L. Nicol, A. Barss, K. I. Forster, and M. F. Garrett. 1991. Syntactically based sentence processing classes: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience* 3(2) 151–65.

Otten, M., and J. J. Van Berkum. 2008. Discourse-based word anticipation during language processing: prediction or priming? *Discourse Processes* 45(6) 464–96.

Posner, M. I., and C. R. R. Snyder. 1975. Attention and cognitive control. *Information processing and cognition: the Loyola symposium*, ed. by R. L. Solso. Hillsdale, NJ: Erlbaum.

Rosenfelt, L. 2012. Rethinking the functional significance of early negativity: investigations in the language/sensory-motor interface. PhD Dissertation, University of California, San Diego, La Jolla, CA.

Schwanenflugel, P. J., and K. L. LaCount. 1988. Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14(2) 344.

Schwanenflugel, P. J., and E. J. Shoben. 1985. The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language* 24(2) 232–52.

Smith, N. J., and R. Levy. 2013. The effect of word predictability on reading time is logarithmic. *Cognition* 128(3) 302–19.

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

Staub, A., and C. Clifton. 2006. Syntactic prediction in language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32(2) 425–36.

Steinhauer, K., and J. E. Drury. 2012. On the left-anterior early negativity (ELAN) in syntax studies. *Brain & Language* 120. 135–62.

Szewczyk, J. M., and H. Schriefers. 2013. Prediction in language comprehension beyond specific words: an ERP study on sentence comprehension in Polish. *Journal of Memory and Language* 68(4) 297–314.

Taylor, W. L. 1953. 'Cloze procedure': a new tool for measuring readability. *Journalism Quarterly* 30, 415–433.

Thornhill, D. E., and C. Van Petten. 2012. Lexical versus conceptual anticipation during sentence processing: frontal positivity and N400 ERP components. *International Journal of Psychophysiology* 83(3) 382–92.

Thothathiri, M. and J. Snedeker. 2008. Give and take: syntactic priming during spoken language comprehension. *Cognition* 108. 51–68.

Tooley, K. M., and M. J. Traxler. 2010. Syntactic priming effects in comprehension: a critical review. *Language and Linguistics Compass* 4(10) 925–37.

Traxler, M. J., and D. J. Foss. 2000. Effects of sentence constraint on priming in natural language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26(5) 1266.

Traxler, M. J., M. J. Pickering, and C. Clifton. 1998. Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language* 39. 558–92.

Tremblay, A., B. Derwing, C. Libben, and C. Westbury. 2011. Processing advantages of lexical bundles: evidence from self-paced reading and sentence recall tasks. *Language Learning* 61(2) 569–613.

Troyer, M., T. J. O'Donnell, E. Fedorenko, and E. Gibson. 2011. Storage and computation in syntax: evidence from relative clause priming. *Proceedings of the 33rd annual conference of the cognitive Science society*, ed. by L. Carlson, C. Hölscher and T. Shipley, 336–41. Austin, TX: Cognitive Science Society.

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

Urbach, T. P., and M. Kutas. 2013. A little context goes a long way: when did you expect? Manuscript submitted for publication.

Vainio, S., J. Hyönä, and A. Pajunen. 2003. Facilitatory and inhibitory effects of grammatical agreement: evidence from readers' eye fixation patterns. *Brain and Language* 85. 197–202.

Van Berkum, J. J. A. 2009. The neuropragmatics of 'simple' utterance comprehension: an ERP review. *Semantic and Pragmatics: From experiment to theory*, ed. U. Sauerland, K. Yatsushiro, pp. 276–316. Basingstoke, UK: Palgrave Macmillan.

Van Berkum, J. J., C. M. Brown, P. Zwitserlood, V. Kooijman, and P. Hagoort. 2005. Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31(3) 443.

Van Petten, C. V. 1993. A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes* 8(4) 485–531.

Van Petten, C., and B. J. Luka. 2012. Prediction during language comprehension: benefits, costs, and ERP components. *International Journal of Psychophysiology* 83(2) 176–90.

Wicha, N. Y., E. A. Bates, E. M. Moreno, and M. Kutas. 2003. Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters* 346(3) 165–8.

Wicha, N. Y., E. M. Moreno, and M. Kutas. 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) 1272–88.

Wlotko, E. W., and K. D. Federmeier. 2012. So that's what you meant! event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage* 62(1) 356–66.

Wlotko, E. W., K. D. Federmeier, and M. Kutas. 2012. To predict or not to predict: age-related differences in the use of sentential context. *Psychology and Aging* 27(4) 975.