

# Probabilistic word pre-activation during language comprehension inferred from electrical brain activity

Katherine A DeLong<sup>1</sup>, Thomas P Urbach<sup>1</sup> & Marta Kutas<sup>1,2</sup>

Despite the numerous examples of anticipatory cognitive processes at micro and macro levels in many animal species, the idea that anticipation of specific words plays an integral role in real-time language processing has been contentious. Here we exploited a phonological regularity of English indefinite articles ('an' precedes nouns beginning with vowel sounds, whereas 'a' precedes nouns beginning with consonant sounds) in combination with event-related brain potential recordings from the human scalp to show that readers' brains can pre-activate individual words in a graded fashion to a degree that can be estimated from the probability that each word is given as a continuation for a sentence fragment offline. These findings are evidence that readers use the words in a sentence (as cues to their world knowledge) to estimate relative likelihoods for upcoming words.

Despite the variety of real-time processing domains across the phylogenetic spectrum in which anticipatory processing has been observed (at both micro and macro levels), the concept of anticipation has played a relatively minor role in language processing theories. Human languages offer unlimited possibilities not only for saying new things but also for saying old things in new ways—far too many ways, some have argued, to make prediction of words a viable and effective strategy except when contextual constraint is unusually high<sup>1</sup>. Accordingly, early language processing models often included some form of memory buffer wherein sentential elements were temporarily stored for later integration at phrasal, clausal or sentence boundaries<sup>2–6</sup>. Since the 1970s, however, the consensus view has been that sentence processing is continuous and incremental, with provisional commitments made that at least temporarily resolve linguistic ambiguities as each word is processed upon its occurrence and rapidly integrated into the sentence representation<sup>7–16</sup>.

More recently, a few researchers have argued for the predictive power of context in generating expectancies during sentence processing, but it has proven difficult to distinguish prediction from integration. What some researchers take as evidence for neural pre-activation (prediction, at a psychological level), others take as a sign of the ease or difficulty in integrating words into message-level representations upon, but not before, their occurrence. A case in point is the N400 component observed in event-related brain potential (ERP) studies, in which cortical neuronal ensembles generate potentials measurable at the scalp.

The N400 (~200–500 ms post-item onset) is the brain's neural response to any potentially meaningful item. Its amplitude is sensitive to word frequency, repetition and concreteness, among other factors. The N400 is especially large to nouns that do not meaningfully fit with their preceding contexts<sup>17</sup>. However, N400s also characterize responses to all but the most highly expected nouns, even when they fit contextually, with amplitudes inversely related ( $r \approx -0.9$ ) to their

offline cloze probabilities<sup>18</sup>. An item's cloze probability is the percentage of individuals that continue a sentence fragment with that item in an offline sentence completion task. Despite the sensitivity of the N400 in response to offline semantic expectancy, it is impossible to determine whether variation in N400 amplitude in response to eliciting words during online, real-time sentence processing means that comprehenders are using context to generate expectancies for upcoming items (prediction view) or whether they are forced by the words themselves to devote more or fewer resources to integrating words into sentence representations (integration view).

Tracking prediction in sentence processing requires a measure that has high temporal resolution (such as ERPs, magnetoencephalogram or eye movements) and does not alter the comprehension process under study. Additionally, it calls for a design that precludes interpretation in terms of integrative difficulty. A few recent ERP and eye-tracking studies have demonstrated contextually generated expectancies for semantic or syntactic features of upcoming words<sup>19–27</sup>. None, however, has demonstrated contextual generation of expectancies for specific word forms in semantically meaningful, syntactically well-formed sentences.

To this end, we designed an experiment which capitalized on the phonological regularity in English whereby the singular indefinite article meaning 'some one thing' is phonologically realized as 'an' before words beginning with a vowel sound and 'a' before words beginning with a consonant sound (for example, 'an airplane' and 'a kite'). To determine whether comprehenders pre-activate specific articles and nouns before their occurrence, we used sentences of varying constraint that led to expectations for particular consonant- or vowel-initial nouns. Across sentences, target nouns ranged from highly probable to unlikely, based on offline cloze probability norming. For instance, given 'The day was breezy so the boy went outside to fly...', the most likely continuation was 'a kite' (cloze of 'a' = 86%, 'kite' =

<sup>1</sup>Department of Cognitive Science and <sup>2</sup>Department of Neurosciences, University of California San Diego, La Jolla, California 92093, USA. Correspondence should be addressed to K.A.D. (kadelong@cogsci.ucsd.edu).

89%). However, the sentence could continue with a plausible, though less likely, alternative such as 'an airplane'. Based on previous studies we knew that the N400 in response to 'kite' would be smaller than that to 'airplane', and more generally that noun N400 amplitude would be highly inversely correlated with cloze probability. However, as previously mentioned, the pattern of noun effects may be a consequence either of the brain's 'surprise' at encountering an item different than what it expects ('prediction view'), or greater difficulty integrating the received word into the sentence representation ('integration view').

Indeed, on the basis of our experiences, 'kite' may be easier to integrate into the developing sentence representation than 'airplane'. Given the difference in their meanings, it is likely that 'kite' and 'airplane' also differ in how well they fit with event schemas that the sentence 'brings to mind' through semantic memory processes. However, whereas 'kite' and 'airplane' differ in meaning, 'a' and 'an' do not, being distinguished only by their phonological forms. Since their semantics are identical and they differ only in frequency of usage and length, there is no reason for the articles to be differentially difficult to integrate into a given sentence representation unless (i) 'a' is always easier to integrate, because it is shorter and/or more frequent than 'an' in everyday usage, or, as we will maintain, (ii) comprehenders have already (unconsciously) formed a higher, non-trivial expectation for 'kite' than for 'airplane'.

If anticipation is an integral part of language processing, then it should be reflected in the brain activity probed by the more and less expected indefinite articles. If the amount of pre-activation is driven strictly by word length or frequency, then whatever the ERP effect, it would be context independent, with all examples of 'a' (versus 'an') patterning together. Even if pre-activation is context dependent, the brain may react to the anticipated article with one response and to anything else with a different response, in a binary rather than a graded fashion. Finally, we hypothesize that the language processor exploits sentence context to probabilistically pre-activate possible continuations, consistent with constraint-based models; if so, the N400 should respond to a degree that can be estimated from the article's offline cloze probability. In sum, no observable difference in the brain's response to more- versus less-expected articles would be a sharp blow to predictive processing accounts, whereas a graded N400 effect correlated with the article's offline cloze probability would support incremental, predictive processing that is sensitive to meaning-based constraints.

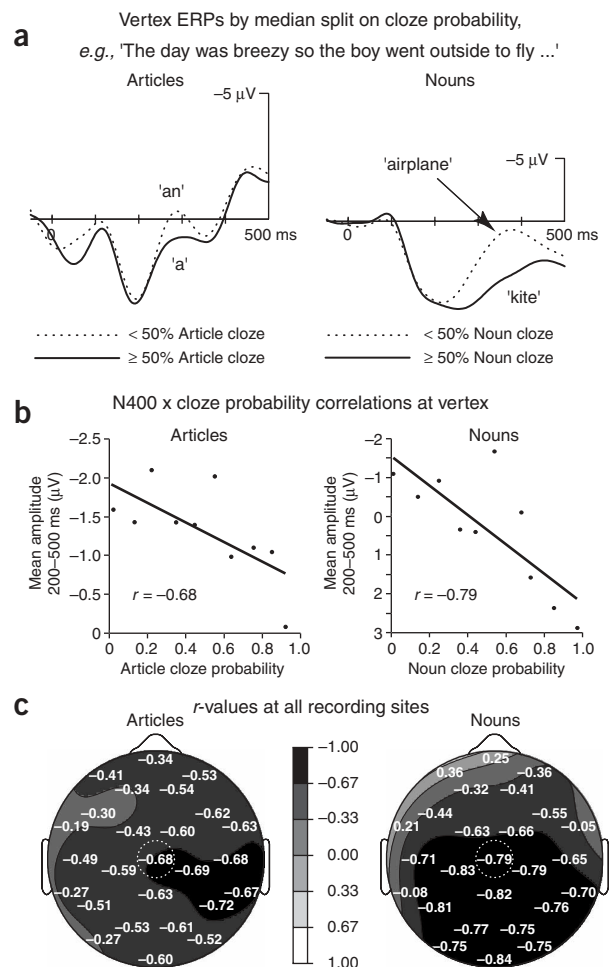
## RESULTS

We obtained offline probabilities for all article and noun targets. Participants were asked to provide the best continuations for sentences truncated before the article or noun. Article cloze ranged from 0–96%; noun cloze ranged from 0–100%. These broad ranges of expectancy allowed for analysis of the correlations between the ERP effects and the offline probabilities of the relevant continuations.

In the ERP experiment, different participants read sentences of varying contextual constraint that included target articles and nouns with large ranges of cloze. Across participants the same sentence context appeared with both higher- and lower-probability articles and nouns. Notably, although some continuations were more probable than others, none was nonsensical, barring participants from developing a strategy (conscious or unconscious) whereby an improbable article was taken to signal an impending semantic anomaly.

### ERP recording and analyses

To measure brain activity associated with prediction effects, we recorded electroencephalograms at 26 scalp locations as 32 participants read sentences word-by-word from a CRT at a rate of 2 words/s (200-ms duration each). ERPs were analyzed for target articles and nouns in 160 sentences, with 16 participants viewing each item. The 160 articles and nouns were sorted into ten equal-width bins as a function of each item's cloze probability, from highest (90–100%) to lowest (0–10%). ERPs for each 10% bin were averaged first within, then across, participants. The average numerical cloze probability of each



**Figure 1** ERP waveforms and correlations between N400 amplitude and cloze probability showing that specific words were predicted during language comprehension. Articles and nouns were analyzed separately. **(a)** Illustrative ERPs at the midline central (vertex) recording site according to median splits on cloze probabilities. Negative amplitudes are plotted upwards. Both articles and nouns with <50% cloze elicited greater negativity between 200–500 ms post-stimulus onset (N400) than those with  $\geq$ 50% cloze. Although the ERPs include responses to both article types ('a', 'an') and both noun types ('kite', 'airplane'), a single sample sentence is provided for simplicity. **(b)** Mean N400 amplitudes were inversely correlated with items' cloze probabilities. Scatter diagrams show strong inverse relations between cloze and N400 amplitude at the vertex for both articles ( $r = -0.68$ ,  $P < 0.05$ ) and nouns ( $r = -0.79$ ,  $P < 0.01$ ). Best-fitting regression lines are also plotted. **(c)** The  $r$ -values for all 26 electrode sites plotted on an idealized head, looking down, nose at the top. Darker shading indicates larger negative correlations, with  $r$ -values between sites estimated by spherical spline interpolation. The dotted circle demarcates the vertex. Although ERP prediction effects were slightly larger over the right hemisphere for both articles and nouns and over posterior sites for nouns, the correlations for both articles and nouns showed maximal values over centroparietal sites, with a right hemisphere bias only for the articles (see **Supplementary Note** for detailed distributional analyses).

bin was then calculated and correlated with mean ERP amplitude in the N400 time window (200–500 ms) for articles and nouns separately. Correlation coefficients ( $r$ -values) and percentage of variance explained by offline probability ( $r^2$ ) were then calculated separately for all 26 electrode sites.

### N400 effects and correlations

As expected, N400 amplitude decreased (became less negative) as noun cloze probabilities increased (Fig. 1a). We replicated the well-known correlation between N400 amplitude and offline cloze of the target nouns (Fig. 1b), with correlation coefficients ranging from  $r = 0.36$  (not significant) to  $r = -0.84$  ( $P < 0.01$ ) at various scalp sites (Fig. 1c). Noun cloze probability thus accounted for up to 71% of variance in brain activity between 200–500 ms after a noun's appearance. Moreover, correlations peaked over posterior sites, where N400 amplitudes are typically largest, whereas anterior sites (where visual N400s are usually less prevalent) showed little if any evidence of similarly correlated brain activity (Fig. 1c). These results were an important precondition for analysis of the articles because they demonstrated that the different degrees of constraint in these materials were reflected in offline expectancies and N400 amplitude modulations in the usual way. However, as previously noted, the noun correlation pattern does not settle the question of prediction because high correlations could reflect either the degree of pre-activation or the variance in the integrability of the noun with the mental representation of the sentence up to that moment.

To address the question of prediction directly we turned to the correlation pattern of the target articles. Although article ERP waveforms were significantly smaller than those elicited by nouns, the amplitude of the negativity in the N400 time window did indeed vary as a function of article expectancy (Fig. 1a). Just as for the nouns, the higher the article's cloze probability, the smaller the ERP negativity between 200–500 ms post-onset (Fig. 1b), with correlation coefficients ranging from  $r = -0.19$  (not significant) to  $r = -0.72$  ( $P < 0.05$ ) at various recording sites (Fig. 1c). Moreover, maximum correlations clustered over centroparietal sites, similar to nouns, albeit more lateralized to the right (Fig. 1c and **Supplementary Note** online). So at least over certain scalp areas, up to 52% of variance in article N400 amplitude was accounted for by the average probability that individuals would continue the sentence context with that particular article offline.

### DISCUSSION

By constructing sentence contexts that led to varying offline expectations for nouns beginning with vowel or consonant sounds, we could assess the extent to which such expectations were formed online by preceding the noun with the phonologically appropriate indefinite article or the other semantically identical, sentimentally congruent but phonologically inappropriate one. Similar to the nouns, the more contextually unexpected an indefinite article was, the more negative the ERP mean amplitude between 200–500 ms post-word onset (N400). In other words, the brain's response to the articles differed in a graded fashion as a function of contextual constraint. Our results thus demonstrate not only that readers can rapidly, incrementally integrate incoming words into evolving mental sentence representations, but that they do so in part by exploiting various constraining forces to form probabilistic predictions of which specific words will come next. Here, we clearly showed this for the target articles and nouns, although we have no reason to assume the same would not hold for every word in a sentence throughout the range of normal reading rates. Notably, maximum correlations for both nouns and articles were not randomly distributed across the scalp but rather clustered over centroparietal

scalp sites (Fig. 1c) where previous reading studies have shown the largest N400 effects. This topographical pattern indicated that the values were not simply the spurious outcome of multiple testing at the 26 electrode sites.

In conducting the correlational analyses, we established the functional relationship of the negativity between 200–500 ms post-article onset to the canonical N400 typically elicited by nouns and verbs. Notably, we also demonstrated that this negativity indexed expectancy for the eliciting article (and upcoming noun). Given that all articles were grammatically and semantically congruent within their contexts and that 'a' and 'an' have identical semantics, there was no reason for either article type to have been any more or less difficult to integrate into the sentence representation. Systematic variation in amplitude of the ERP negativity in relation to offline article cloze probability thus constitutes strong evidence that participants were indeed anticipating the phonological form of a particular noun and therefore had formed expectations for one article type relative to the other and seemed to experience some processing difficulty when the less-expected article appeared.

Articles are relatively short; highly frequent; highly predictable as a word class; not as semantically rich as nouns, verbs, adjectives or adverbs and are often skipped over during natural reading<sup>28</sup>. In addition to providing unequivocal evidence for lexically specific prediction, the article correlations are compelling evidence that articles, too, are predicted and integrated with context in qualitatively similar ways as nouns. For reasons not yet known, the correlations with offline probability were on average lower for articles than for nouns (although at some electrode sites, the two were statistically similar). Nonetheless, the article correlations clearly demonstrated that prediction is not limited to highly constraining contexts. We believe that this sort of anticipation is an integral (perhaps inevitable) part of real-time language processing and is likely to have a functional role, although this has yet to be demonstrated.

Our findings thus suggest that individuals can use linguistic input to pre-activate representations of upcoming words in advance of their appearance. Exactly what informs these predictions, as well as the neural mechanism for predictive language processing, are matters for empirical and computational investigations. An open question, for instance, is how the human sentence comprehension system handles variation in natural input rates (for example, 2 to 3.5 words/s), and in particular, whether the same (or different) mechanisms are engaged. In line with most studies of comprehension that draw general conclusions without systematically varying input rate, we assume that basic language processing mechanisms do not vary fundamentally across the range of normal input rates. This parsimonious assumption is bolstered by results of N400 studies in which variation in presentation rate showed no evidence for the engagement of qualitatively different neural mechanisms<sup>29,30</sup>. Although the current study demonstrated graded prediction only at the slower end of natural input rates, we suggest that this conclusion may generalize to faster rates, given the aforementioned arguments and ERP evidence for binary prediction (expected versus unexpected) in natural speech<sup>26,27</sup>. Subsequent experimentation will undoubtedly shed more light on these issues.

We propose that single words and combinations of words in a sentence tap into and differentially activate information via semantic memory, going beyond the immediate physical input. Semantic memory is presumed to include information about individual words as well as world knowledge built up from experience of people, places, things and events. We maintain that probabilistic pre-activation of particular word forms follows the access of this experiential knowledge as a result of linguistic input. Our observation of an ERP expectancy effect at the

article leads us to conclude that predictions can be for specific phonological forms—words beginning with either vowels or consonants. In this sense, we propose that prediction can be highly specific, at least under some circumstances.

Our results are in line with a growing list of empirical studies demonstrating that the brain's language parser projects probabilistic expectations about various aspects of linguistic processing during online sentence comprehension tasks. Several studies suggest that the parser uses constraints accruing as a sentence is analyzed word-by-word to (i) compute likely relationships among referents in linguistic and visual contexts (for example, upon hearing the word 'eat', a person is likely to scan the environment for something edible)<sup>19</sup>, (ii) pre-activate semantic features of categories (for example, expecting a particular kind of tree pre-activates features of trees, even when not all trees would be plausible in the sentence context)<sup>24</sup> or (iii) anticipate various syntactic aspects of to-be-presented material (for example, expecting the grammatical gender of upcoming items in gender-marked languages such as Spanish or Dutch) during word-by-word reading and natural speech<sup>25–27</sup>.

In particular, our study expands on findings from the aforementioned grammatical gender studies<sup>25,27</sup> in which nouns were preceded by words whose syntactic gender marking was inconsistent with that of the 'expected' noun. Although both the gender studies and our study used the same general logic, the differences in experimental manipulations, design and analyses lead to substantive differences in the justifiable conclusions. Whereas the Spanish and Dutch studies utilized pre-nominal gender marking on determiners and adjectives, respectively, our study relied upon a purely phonological (sound representation) relation between probed articles and upcoming nouns. Thus, our observation of article ERP variation provides a strong test of whether the language system predicts word forms with specific phonological content (lexemes), instead of simply representations specifying words' semantic and syntactic properties (lemmas). In addition, having tested prediction with semantically identical 'a'/'an' articles (function words) instead of words richer in meaning (content words) such as adjectives, we effectively counter the argument that the observed difference between more- versus less-predictable articles reflects difficulty interpreting them. And perhaps most notably, only our study compares brain activity elicited by a range of more- or less-predictable articles, not simply most- versus least-expected. The article correlation findings thus show for the first time that the language system does not simply pre-activate a single word when its representation exceeds some threshold given a highly constraining context. Instead, a gradient of pre-activation shows that the system makes graded predictions.

Our electrophysiological results extend previous prediction findings in several other notable ways. First, they demonstrate that a candidate entity (or its depiction) need not be physically present in order for the brain to narrow the possibilities for likely continuations; rather, predictions can emerge on the basis of associations that form as sentential context accrues. Second, our results illustrate that at least one subclass of function words (which generally provide more grammatical structure than lexical meaning), indefinite articles, can be important in building context and facilitating linguistic processing. This finding is particularly relevant given the paucity of evidence in the comprehension literature about semantic context effects on function words<sup>31,32</sup>. And finally, our findings unambiguously show that anticipatory processing can happen not only for conceptual or semantic features but also for specific phonological word forms. In sum, although natural language comprehension must occur over a range of input rates with a nearly infinite number of possible word combinations, these factors do not seem to prevent the brain from anticipating the most probable

continuations for sentences in advance of the actual input. In this regard, language comprehension appears to involve a special case of the anticipatory behavior observed in other biological systems.

## METHODS

**Experimental design and materials.** Stimuli consisted of 80 sentence contexts with two possible target types: relatively expected and unexpected indefinite article/noun pairs. Each article/noun pair served as a more- and less-expected target in different contexts. Targets were sentence medial and congruent (that is, no agreement violations such as 'a airplane'). The 160 stimuli were divided into two lists of 80 sentences, each participant viewing one list. Sentence contexts and article/noun targets were used only once per list. Each list contained equal numbers of relatively expected and unexpected targets as well as 'a' and 'an' targets. One-quarter of sentences were followed by yes/no comprehension questions, 94% of which were answered correctly by participants, on average (range, 88–100%).

Norming for articles and nouns was done with different groups of student volunteers. Informed written consent was obtained from all norming (and ERP) participants. For articles, cloze ratings were obtained from 30 participants for 80 sentence contexts truncated before the target article. For nouns, sentences were truncated after target articles, with two versions of each context (160 sentences total): one with the more-probable article supplied, the other with the less-probable article. Individual participants saw only one version, with each normed by 30 participants.

**Participants.** Thirty-two volunteers (23 women) participated in the ERP experiment for course credit or for cash. All were right-handed, native English speakers with normal or corrected-to-normal vision, between 18–37 years (mean, 21 years). Seven participants reported a left-handed parent or sibling.

**Procedure.** Volunteers were tested in a single session, with visual sentences presented centrally, one word at a time (200-ms duration, 500-ms stimulus onset asynchrony). The instructions were to read sentences for comprehension and answer yes/no comprehension questions by pressing hand-held buttons.

The electroencephalogram (EEG) was recorded from 26 electrodes arranged geodesically in an Electro-cap, each referenced online to the left mastoid. Blinks and eye movements were monitored from electrodes placed on the outer canthi and under each eye, referenced to the left mastoid. Electrode impedances were kept below 5 K $\Omega$ . The EEG was amplified with Grass amplifiers with a band-pass of 0.01 to 100 Hz, continuously digitized at a sampling rate of 250 Hz.

**Data analysis.** Trials contaminated by eye movements, excessive muscle activity or amplifier blocking were rejected offline before averaging; on average, 10.7% of article trials and 11.4% of noun trials were rejected. Data with excessive blinks were corrected using a spatial filter algorithm. A digital band-pass filter set from 0.2 to 15 Hz was used on all data to reduce high-frequency noise. Data were re-referenced offline to the algebraic sum of left and right mastoids and averaged for each experimental condition, time-locked to the target article and noun onsets.

*Note: Supplementary information is available on the Nature Neuroscience website.*

## ACKNOWLEDGMENTS

Many thanks to E. De Ochoa and J. Cagle for help with data collection and to J. Elman for helpful comments on this manuscript. This research was supported by US National Institute of Child Health and Human Development grant HD22614 and National Institute on Aging grant AG08313 to M.K. and by a Center for Research in Language Training Fellowship to K.A.D.

## COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 21 April; accepted 21 June 2005

Published online at <http://www.nature.com/natureneuroscience/>

1. Stanovich, K.E. & West, R.F. Mechanisms of sentence context effects in reading: automatic activation and conscious attention. *Mem. Cognit.* **7**, 77–85 (1979).
2. Kintsch, W. & van Dijk, T.A. Towards a model of text comprehension and production. *Psychol. Rev.* **85**, 363–394 (1978).



3. Carroll, J. & Bever, T. Sentence comprehension: a case study in the relation of knowledge and perception. in *Handbook of Perception*, Vol. 7 (eds. Carterette, E. & Friedman, M.) 299–317 (Academic, New York, 1978).
4. Just, M.A. & Carpenter, P.A. A theory of reading: From eye fixations to comprehension. *Psychol. Rev.* **87**, 329–354 (1980).
5. Daneman, M. & Carpenter, P.A. Individual differences in integrating information between and within sentences. *J. Exp. Psychol. Learn. Mem. Cogn.* **9**, 561–584 (1983).
6. Mitchell, D.C. & Green, D.W. The effects of context and content on immediate processing in reading. *Q. J. Exp. Psychol.* **30**, 609–636 (1978).
7. Marslen-Wilson, W.D. Sentence perception as an interactive parallel process. *Science* **189**, 226–228 (1975).
8. Altmann, G.T.M. & Steedman, M.J. Interaction with context during human sentence processing. *Cognition* **30**, 191–238 (1988).
9. Pickering, M.J. Processing local and unbounded dependencies: a unified account. *J. Psycholinguist. Res.* **23**, 323–352 (1994).
10. Tyler, L.K. & Marslen-Wilson, W.D. The on-line effects of semantic context on syntactic processing. *J. Verb. Learn. Verb. Behav.* **16**, 683–692 (1977).
11. Traxler, M., Bybee, M. & Pickering, M. Influence of connectives on language comprehension: eye-tracking evidence for incremental interpretation. *Q. J. Exp. Psychol. A* **50A**, 481–497 (1997).
12. Marslen-Wilson, W.D. & Tyler, L.K. The temporal structure of spoken language understanding. *Cognition* **8**, 1–71 (1980).
13. Steedman, M. Grammar, interpretation, and processing from the lexicon. in *Lexical Representation and Process* (ed. Marslen-Wilson, W.D.) 463–504 (MIT Press, Cambridge, Massachusetts, 1989).
14. Boland, J.E., Tanenhaus, M.K., Garnsey, S.M. & Carlson, G.N. Verb argument structure in parsing and interpretation: Evidence from wh-questions. *J. Mem. Lang.* **34**, 774–806 (1995).
15. Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. Integration of visual and linguistic information during spoken language comprehension. *Science* **268**, 1632–1634 (1995).
16. Kutas, M. & Hillyard, S.A. Event-related brain potentials to grammatical errors and semantic anomalies. *Mem. Cognit.* **11**, 539–550 (1983).
17. Kutas, M. & Hillyard, S.A. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* **207**, 203–204 (1980).
18. Kutas, M. & Hillyard, S.A. Brain potentials during reading reflect word expectancy and semantic association. *Nature* **307**, 161–163 (1984).
19. Altmann, G.T.M. & Kamide, Y. Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* **73**, 247–264 (1999).
20. Tanenhaus, M.K., Magnuson, J.S., Dahan, D. & Chambers, C. Eye movements and lexical access in spoken-language comprehension: evaluating a linking hypothesis between fixations and linguistic processing. *J. Psycholinguist. Res.* **29**, 557–580 (2000).
21. Kamide, Y., Scheepers, C. & Altmann, G.T.M. Integration of syntactic and semantic information in predictive processing: cross-linguistic evidence from German and English. *J. Psycholinguist. Res.* **32**, 37–55 (2003).
22. Chambers, C.G. & Smyth, R. Structural parallelism and discourse coherence: a test of centering theory. *J. Mem. Lang.* **39**, 593–608 (1998).
23. Sedivy, J.C., Tanenhaus, M.K., Chambers, C.G. & Carlson, G.N. Achieving incremental semantic interpretation through contextual representation. *Cognition* **71**, 109–148 (1999).
24. Federmeier, K.D. & Kutas, M. A rose by any other name: long-term memory structure and sentence processing. *J. Mem. Lang.* **41**, 469–495 (1999).
25. Wicha, N.Y.Y., Moreno, E.M. & Kutas, M. Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *J. Cogn. Neurosci.* **16**, 1272–1288 (2004).
26. Wicha, N.Y.Y., Bates, E.A., Moreno, E.M. & Kutas, M. Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neurosci. Lett.* **346**, 165–168 (2003).
27. Van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V. & Hagoort, P. Anticipating upcoming words in discourse: evidence from ERPs and reading times. *J. Exp. Psychol. Learn. Mem. Cogn.* **31**, 443–467 (2005).
28. O'Regan, K. Saccade size control in reading: evidence for the linguistic control hypothesis. *Percept. Psychophys.* **25**, 501–509 (1979).
29. Kutas, M. Event-related brain potentials (ERPs) elicited during rapid serial visual presentation of congruous and incongruous sentences. *Electroencephalogr. Clin. Neurophysiol.* **40** (suppl.), 406–411 (1987).
30. Gunter, T.C., Jackson, J.L. & Mulder, G. An electrophysiological study of semantic processing in young and middle-aged academics. *Psychophysiology* **29**, 38–54 (1992).
31. King, J.W. & Kutas, M. A brain potential whose latency indexes the length and frequency of words. *CRL Newsletter* **10**, 1–9 (1995).
32. Kluender, R. & Kutas, M. Subjacency as a processing phenomenon. *Lang. Cognitive Proc.* **8**, 573–633 (1993).