Understanding Sentences in Context

What Brain Waves Can Tell Us

Jos J.A. Van Berkum

Max Planck Institute for Psycholinguistics and F.C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

ABSTRACT—*Language comprehension looks pretty easy. You pick up a novel and simply enjoy the plot, or ponder the human condition. You strike a conversation and listen to whatever the other person has to say. Although what you're taking in is a bunch of letters and sounds, what you really perceive—if all goes well—is meaning. But how do you get from one to the other so easily? The experiments with brain waves (event-related brain potentials or ERPs) reviewed here show that the linguistic brain rapidly draws upon a wide variety of information sources, including prior text and inferences about the speaker. Furthermore, people anticipate what might be said about whom, they use heuristics to arrive at the earliest possible interpretation, and if it makes sense, they sometimes even ignore the grammar. Language comprehension is opportunistic, proactive, and, above all, immediately context-dependent.*

KEYWORDS—*sentence comprehension; meaning; communication; prediction; brain potentials*

"*The deadline for this review paper was ten days ago. If I don't finish it soon they will start bugging me.*"

Suppose we just met in some university café, and I start telling you about my deadline. How would you make sense of the above? Words are obvious carriers of meaning, so here's a naive model: First you recognize each of the words, then you look up their meaning in your mental dictionary, and then, using syntax to guide the combination, you simply combine the meanings so that you know what I said. Presto.

Address correspondence to Jos Van Berkum, Max Planck Institute for Psycholinguistics, PO Box 310, 6500 AH Nijmegen, The Netherlands; e-mail: jos.vanberkum@mpi.nl.

NOT SO EASY

In fact, there is a lot more to utterance interpretation than just adding up word meanings in line with their syntactic arrangement. First, words like "paper" have multiple meanings, so it is not all that clear what needs to be combined. Next, although combining the right meanings may give you an abstract meaning for a phrase like "this review paper" (say, a particular scientific survey article, presumably salient to the listener), that doesn't by itself tell you what paper I have in mind. Furthermore, your understanding of the second sentence clearly depends on your understanding of the first, on background knowledge about science writing, and on what you know about me—for instance, whether I just enrolled in a freshman writing course or am an experienced scientist instead.

Of course, language researchers have long since realized that the naive model is too simple; how words, their arrangement, and the wider communicative context contribute to utterance meaning has been studied intensely by linguists. However, little is known about how the brain actually constructs an interpretation, in context, as the utterance unfolds in real time. For instance, when and how do listeners and readers relate the words of an unfolding sentence to the wider communicative context? How do they work out to what or whom a speaker refers with "they," or "this review paper"? Do people actually guess how the unfolding sentence might continue? Here I review studies from my lab in which my colleagues and I have used event-related brain potentials (ERPs) to address these questions.

An ERP is a patterned series of voltage deflections elicited by a critical stimulus, such as a tone or a word. It is obtained by averaging bits of raw electroencephalogram (EEG) data recorded at the scalp as the listener or reader encounters a large number of these stimuli. Because of their speed, ERPs allow us to trace the very rapid processes involved in language comprehension. Additional advantages are that ERPs can be recorded

without having participants perform an additional—and often unnatural—response task, and that the type of deflection (its shape, polarity, and location at the scalp) can help identify the specific process being engaged at that time. Some example ERPs, time-locked to different words in a text, are shown in Figure 1 (with negative voltage conventionally plotted upwards); the three differential ERP effects will be discussed below.

MAKING SENSE OF WORDS IN DISCOURSE

One thing we observe again and again in ERPs is that the brain very rapidly relates the words of an unfolding sentence to the wider discourse. Listeners and readers, for instance, need only a fraction of a second to discover that a word does not fit the preceding text (see Van Berkum, in press-a, for references). Figure 1c gives an example: If the contextually appropriate word "sack" is replaced by a word that might fit the local sentence in some happier discourse context but is inappropriate here, such as "promote," the latter word immediately elicits an *N400 effect*, an increase in the size (amplitude) of the N400 component. The N400 is a negative ERP deflection that begins to develop around 150 to 250 milliseconds, peaks around 400 milliseconds, and is strongest over the back of the head. It is associated with the analysis of meaning, and not directly sensitive to, for instance, the analysis of phrase structure (Kutas, Van Petten, & Kluender, 2006). Importantly, the N400 is not simply an anomaly detector;

it reflects something about normal language comprehension. For example, relative to a coherent word like "sack," a somewhat less expected yet equally acceptable word like "report" also elicits a larger N400 (Otten & Van Berkum, 2007; see Fig. 2b for an ERP example).

Strikingly, with spoken words, the brain responds to the fit between word and context well before people have actually heard the end of the word. So, in the example of Figure 1c, hearing something as short as "pro . . ." rather than "sa . . ." is already enough to generate an N400 effect, even though none of these strings by itself pins down a specific word. Also, discourse-dependent N400 effects cannot simply be reduced to basic word-to-word associative priming (Camblin, Gordon, & Swaab, 2007; Ditman, Holcomb, & Kuperberg, 2007; Otten & Van Berkum, 2007), or the rapid use of basic world knowledge. In a cartoon-like story about an amorous peanut, for example, the contextually appropriate sentence "the peanut is in love" elicits a smaller N400 at the last word than a much more typical sentence like "the peanut is salted" (Nieuwland & Van Berkum, 2006). What primarily seems to matter is how things fit what is being talked about right now, be it the real world of papers and plagiarism or some imaginary world with unicorns and happy peanuts.

If modern analyses of language use (e.g., Clark, 1996) are correct, the identity and perspective of the speaker are also an indispensable early source of meaning. Recent ERP evidence confirms this (Van Berkum, Van den Brink, Tesink, Kos, &
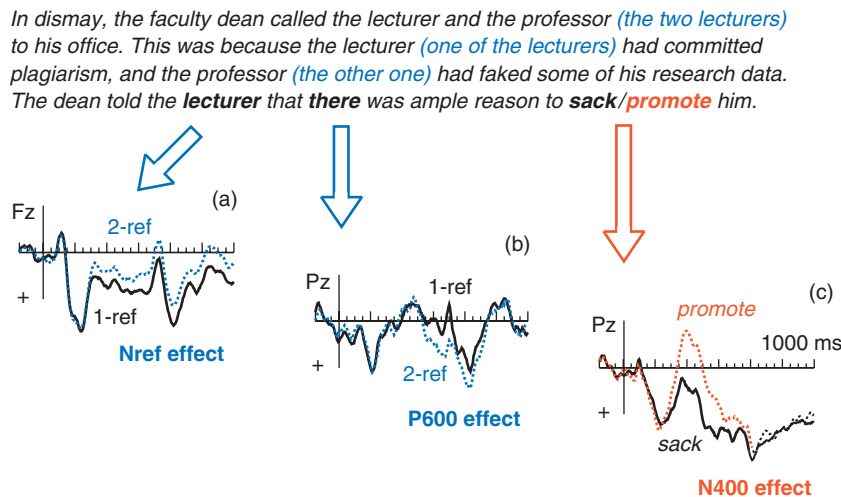
Fig. 1. Rapid consequences of discourse context for the referential (a), syntactic (b), and semantic (c) analysis of an unfolding sentence (top). Graph (a) shows a sustained frontal negative shift or *Nref* effect to a discourse-induced referential problem (in the 2-referent version of the story, "lecturer" is ambiguous). Graph (b) shows a *P600* effect to a discourse-induced syntactic problem (the provisional relative-clause analysis temporarily pursued at "that" in the 2-referent context is subsequently ruled out by "there"). Graph (c) shows an *N400* effect to a discourse-induced semantic problem ("promote" does not fit the wider story context). The example sentence is shown here in several variants (1- and 2-referent contexts, coherent/anomalous ending), but any one participant saw just a single version. ERP waveforms are time-locked to the presentation of the critical word (0 milliseconds) and are shown for 1200 milliseconds each. Negative voltage is plotted upwards. All data were obtained in a single reading experiment, and were replicated with spoken language. See Van Berkum (in press-a) for references.

Hagoort, 2008). When people heard utterances whose content sometimes did not match what could be inferred about the speaker from his or her voice (e.g., "I have a large tattoo on my back" in an upper-class accent or "I like olives" in a young child's voice), words at which the mismatch became apparent ("tattoo," "olives") elicited a standard N400 effect starting about 200 milliseconds after their acoustic onset (relative to the same words uttered by a more appropriate speaker). This shows that listeners immediately take the speaker into account. Because this brain response hinges on voice-based and usually stereotype-dependent inferences about the speaker, it also reveals that listeners rapidly classify speakers on the basis of their voices and bring the associated social stereotypes to bear when interpreting what is said.

This fits well with everyday intuitions, of course. But note that inferences about the speaker kick in well before a word has ended, and do so *along with* more classic semantic constraints such as the meaning of the preceding words. This goes against simple models in which the local linguistic context constrains interpretation well before knowledge about the speaker can do so. Although we as (psycho)linguists might like things to occur in neatly ordered steps, the linguistic brain seems much more "messy" and opportunistic, taking any partial cue that seems to bear on interpretation into account as soon as it can.

FINDING REFERENTS

What about determining what or whom speakers are actually referring to? Here, the evidence also points to rapid discourse-dependent processing. For example, in a story with two equally salient lecturers, ambiguous expressions such as "The dean told the lecturer" elicit a negative ERP shift right at the ambiguous word, emerging at about 300 to 400 milliseconds over the front of the head (see Fig. 1a for an example of this *Nref* effect, and Van Berkum, Koornneef, Otten, & Nieuwland, 2007, for review; these effects are relative to the same word used in a single-referent control context). Also, the mechanism involved is relatively smart: In a context in which, say, one of the two lecturers has just fled from the building, "The dean told the lecturer" does not elicit the effect, even though the episodic memory of the discourse does contain two lecturers (Nieuwland, Otten, & Van Berkum, 2007). Thus, based on the unfolding sentence as well as prior text, people rapidly work out which are the conceptually suitable referents. And because the effect of referential ambiguity is largest over the front of the head—clearly different from the N400 effect—we can infer that different aspects of interpretation are handled by at least partly different networks in the brain—an inference that has recently been confirmed with fMRI (Nieuwland, Petersson, & Van Berkum, 2007).

If you do not know who or what is being referred to, this can also immediately affect your word-by-word syntactic analysis of the sentence. For instance, when "the lecturer" in "The dean told the lecturer that . . ." is referentially ambiguous, people are lured into analyzing the subsequent word "that" as starting a relative clause that will tell them which lecturer was meant (e.g., ". . . the lecturer that had committed plagiarism") rather than as starting some other syntactic structure (e.g., ". . . that there was ample reason to sack him"). If the sentence then continues with the latter after all, we see a so-called *P600 effect*, the ERP effect that is always elicited by a word at which the currently pursued syntactic analysis runs into problems (see Fig. 1b, "there"; Van Berkum, Brown, & Hagoort, 1999). Thus, referential factors can lead readers to momentarily perceive a syntactic error when there is none. The same thing can be observed with fragments like "David praised Linda because . . . ," in which people expect to read more about Linda, not David. This referential expectation is so strong that if the sentence continues with a masculine pronoun ("David praised Linda because he . . ."), it elicits a P600 effect, indicating that people momentarily take this perfectly legitimate pronoun to be syntactically incorrect (Van Berkum et al., 2007).

Even more strikingly, the conceptual "pull" of referential factors can sometimes briefly lure people into pursuing a syntactic analysis that is ungrammatical. For example, in a Dutch version of "The dean told the lecturer that . . . ," referential ambiguity can lead people to momentarily read "that" as a relative pronoun even if a relative pronoun of this form does not agree in gender with the gender of the preceding noun (see Van Berkum et al., 1999). Such observations are important, because they suggest that when interpreting language, people don't just slavishly follow the syntax. Other recent ERP findings also suggest that good sense can sometimes outweigh good syntax (e.g., Kuperberg, 2007). Against so-called "syntax-first" or "syntactocentric" models of processing, this indicates that comprehenders are trying to deal with multiple levels of linguistic structure simultaneously, without giving one of them absolute priority.
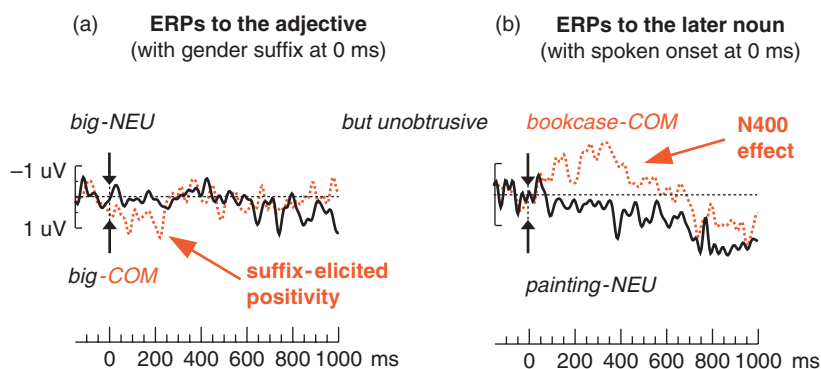
PREDICTIONS AND SHORTCUTS

The rapid context-dependent ERP effects raise an important question: How does the linguistic brain get to be this fast? Part of the story may well be that we are unconsciously yet continuously predicting what might be talked about next. Recent ERP findings suggest, for example, that people use their knowledge of the wider discourse to predict specific upcoming words. As can be seen in Figure 2a, the prediction of an upcoming Dutch noun can be detected by preceding it with an incorrectly gender-marked adjective (Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; see also Otten & Van Berkum, in press). Also, even if a specific word cannot be predicted, as in "He ate the . . . ," people can usually anticipate particular semantic features (Federmeier, 2007).

Such anticipation can be based on a highly sophisticated, precise analysis of what has been said before (Otten & Van Berkum, 2007; Otten & Van Berkum, in press). At the same time,

*The burglar had no trouble whatsoever
to locate the secret family safe.
Of course,it was situated behind a...*

(discourse biases towards:*painting-NEU*)

—— consistent with prediction

·········· inconsistent with prediction

(a) **ERPs to the adjective**
(with gender suffix at 0 ms)

big-NEU   *but unobtrusive*

−1 uV

1 uV

big-COM   **suffix-elicited positivity**

0 200 400 600 800 1000 ms

(b) **ERPs to the later noun**
(with spoken onset at 0 ms)

bookcase-COM   **N400 effect**

painting-NEU

0 200 400 600 800 1000 ms

Fig. 2. Immediate use of knowledge of the wider discourse to predict specific upcoming words. Participants listened to (Dutch) mini-stories (left) that, when truncated after the article ''a'' in an earlier completion pretest, were predominantly completed with one specific critical noun—in this case, ''painting,'' a word that has neuter (NEU) grammatical gender in Dutch. In the ERP study, the article ''a'' was first followed by a semantically coherent adjective whose gender suffix did or did not agree with the anticipated noun. Graph (a) shows that ''incorrectly'' inflected adjectives (such as ''big-COM'' when ''painting-NEU'' might be expected; COM denotes common gender) elicited a positive ERP effect right at the suffix, showing that listeners were at this point indeed anticipating that particular noun, and showing that strongly anticipated ''ghost'' words are involved in syntactic analysis before they actually appear in the input. Graph (b) shows that a coherent but much less expected later noun (e.g., ''bookcase'') elicits a large N400 effect. Negative voltage is plotted upwards. See Van Berkum et al. (2005).

however, we know that the brain also takes shortcuts, using its statistical knowledge of what we tend to talk about (and how) to make a quick-and-dirty heuristic analysis of the input without waiting for all the details, in parallel with a more sophisticated analysis (e.g., Sanford & Sturt, 2002). There is good evidence that the N400 can also reflect this more shallow heuristic processing. For example, in stories where a highly anomalous word is scenario-relevant, related to the appropriate word, and presented in a "business as usual" way (as in offhandedly stating that they buried the survivors of a plane crash), the severe anomaly in meaning does *not* generate an N400 effect (Nieuwland & Van Berkum, 2005).

That the N400 can index shallow, associative processing is also evident from the fact that simply presenting one associatively related prior word can already reduce a word's N400 component (e.g., "arms"–"legs"; Camblin et al., 2007; see Kutas et al., 2006, for review). The rapid N400 effects observed in language comprehension thus seem to reflect a system that uses both a sophisticated (message-level) and a more heuristic (e.g., scenario- and association-based) analysis of the available context to "get ready" for the next bit of information. As argued elsewhere, much of this probably comes for free as a result of how our long-term semantic memory works (Kutas & Federmeier, 2000; Van Berkum, in press-b). That is, the context-dependent N400 effects might well be indicating the processing consequences of rapid memory-mediated anticipation.

NEW CHALLENGES

The ERP findings converge with behavioral research (see Recommended Reading) in showing that to make sense of an unfolding utterance, readers and listeners rapidly draw upon a wide variety of information sources. Also, people don't just stick to

what is given: They anticipate what might be said, about whom; they use heuristics to arrive at the earliest possible interpretation; and if it makes sense, they sometimes ignore the syntactic rules of their language. What we see is an opportunistic, proactive brain at work.

So what's next? One clear limitation of current neurocognitive research is that, although we see the brain doing its job, it is difficult to pin down what exact mechanisms we're looking at. I actually suspect that, without detailed computational models of how utterance meaning is constructed as language unfolds, we won't really get much further than this. It is tempting to think that, now that we can "look under the hood of the car" with neuroimaging tools, we just need more experiments. I think this is a mistake, and that playing 20 Questions with nature (asking simple yes-or-no questions to work something out) is as limited as it was several decades ago, when Alan Newell (1973) made his plea for computational models. The big challenge, then, is to construct precise models of interpretation that not only fit the behavioral facts and linguistic analyses, but whose machinery can also be mapped onto the neuroimaging data.

A second challenge ahead of us is to take neuroimaging a bit further into the arena of real language use. Many interesting questions await us there. One set of issues revolves around the speaker, the type of discourse, and "layers" in communication: For example, how does your brain respond to "the clouds are whispering" if you know that the speaker is a poet, a patient suffering from word-finding problems, or a protagonist in a story you are reading? Furthermore, to what extent do the processes that construct an interpretation depend on whether the message conveyed is relevant to the listener socially, emotionally, or for some task at hand? And what about the intentions of the speaker? Returning to the café example, why was I telling you about my deadline in the first place? Maybe it was just to get it off my chest

or make conversation. But if you have recently burdened me with a 200% teaching load, well, maybe I was communicating something else. Such intentional acts are at the heart of everyday communication, but right now, we haven't got a clue as to how the brain works out all of this as words come in. It is time to find out.

Recommended Reading

Clark, H.H. (1996). (See References). A thorough linguistic and psycholinguistic analysis of comprehension and production in the context of real language use, with strong emphasis on collaborative, social aspects of linguistic communication.

Jackendoff, R. (2007). A parallel architecture perspective on language processing. *Brain Research*, *1146*, 2–22. A linguistically grounded, "non-syntactocentric" theory of how comprehenders analyze multiple levels of language structure (e.g., sound, syntax, meaning) simultaneously.

Kutas, M., Van Petten, C., & Kluender, R., (2006). (See References). A thorough review of psycholinguistically relevant ERP effects, primarily aimed at sentence comprehension.

Trueswell, J., & Tanenhaus, M. (Eds.). (2005). *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions*. Cambridge, MA: MIT Press. A recent overview of behavioral (eye tracking) research on language comprehension in context.

Van Berkum, J.J.A. (2004). Sentence comprehension in a wider discourse: Can we use ERPs to keep track of things? In M. Carreiras & C. Clifton Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 229–270). New York: Psychology Press. A review of the pros and cons of using ERPs to study discourse-level comprehension, illustrated with some of the experiments mentioned above.

REFERENCES

Camblin, C.C., Gordon, P.C., & Swaab, T.Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory & Language*, *56*, 103–128.

Clark, H.H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.

Ditman, T., Holcomb, P.J., & Kuperberg, G.R. (2007). The contributions of lexico-semantic and discourse information to the resolution of ambiguous categorical anaphors. *Language and Cognitive Processes*, *22*, 793–827.

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

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

Kutas, M., & Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*, 463–470.

Kutas, M., Van Petten, C., & Kluender, R. (2006). *Psycholinguistics electrified II: 1994–2005*. In M. Traxler & M.A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 659–724). New York: Elsevier.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W.G. Chase (Ed.), *Visual information processing* (pp. 283–308). New York: Academic Press.

Nieuwland, M.S., Otten, M., & Van Berkum, J.J.A. (2007). Who are you talking about? Tracking discourse-level referential processing with ERPs. *Journal of Cognitive Neuroscience*, *19*, 228–236.

Nieuwland, M.S, Petersson, K.M., & Van Berkum, J.J.A. (2007). On sense and reference: Examining the functional neuroanatomy of referential processing. *NeuroImage*, *37*, 993–1004.

Nieuwland, M.S., & Van Berkum, J.J.A. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, *24*, 691–701.

Nieuwland, M.S., & Van Berkum, J.J.A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, *18*, 1098–1111.

Otten, M., & Van Berkum, J.J.A. (2007). What makes a discourse constraining? Comparing the effects of discourse message and scenario fit on the discourse-dependent N400 effect. *Brain Research*, *1153*, 166–177.

Otten, M., & Van Berkum, J.J.A. (in press) Discourse-based anticipation during language processing: Prediction or priming? *Discourse Processes*.

Sanford, A.J., & Sturt, P. (2002). Depth of processing in language comprehension: not noticing the evidence. *Trends in Cognitive Sciences*, *6*, 382–386.

Van Berkum, J.J.A. (in press-a). The electrophysiology of discourse and conversation. In M. Spivey, K. McRae, & M. Joanisse (Eds.), *The Cambridge handbook of psycholinguistics*. Cambridge, UK: Cambridge University Press.

Van Berkum, J.J.A. (in press-b). The neuropragmatics of 'simple' utterance comprehension: An ERP review. In U. Sauerland & K. Yatsushiro (Eds.), *Semantic and pragmatics: From experiment to theory*. New York: Palgrave.

Van Berkum, J.J.A., Brown, C.M., & Hagoort, P. (1999). When does gender constrain parsing? Evidence from ERPs. *Journal of Psycholinguistic Research*, *28*, 555–571.

Van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *31*, 443–467.

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

Van Berkum, J.J.A., Van den Brink, D., Tesink, C., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*, 580–591.