Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity
Author(s): Marta Kutas and Steven A. Hillyard
Source: *Science*, New Series, Vol. 207, No. 4427 (Jan. 11, 1980), pp. 203-205
Published by: American Association for the Advancement of Science
Stable URL: http://www.jstor.org/stable/1683915
Accessed: 18-09-2016 21:42 UTC

Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity

Abstract. *In a sentence reading task, words that occurred out of context were associated with specific types of event-related brain potentials. Words that were physically aberrant (larger than normal) elicited a late positive series of potentials, whereas semantically inappropriate words elicited a late negative wave (N400). The N400 wave may be an electrophysiological sign of the "reprocessing" of semantically anomalous information.*

Language comprehension has often been characterized as a continual testing and updating of hypotheses about the words that are likely to occur next in a text or conversation (*1*). For example, most people familiar with English would complete the sentence "I take coffee with cream and _____" with the word "sugar," although reasonable alternatives exist. In general, words that are expected in a given context are recognized, remembered, and verbalized more rapidly and accurately than the same stimuli presented alone or in a semantically inappropriate context (*2*).

We have investigated event-related brain potentials (ERP's) that occur when an inappropriate word occurs unexpectedly at the end of a sentence. This approach stemmed from extensive human research showing that certain components recorded from the scalp are sensitive to a person's expectations. In particular, unexpected or surprising stimuli are typically followed after some 300 to 600 msec by a positive ERP component known as the P300 (*3*). We now report, however, that an occasional semantic deviation of a word in a sentence is followed by a negative brain wave (N400) that is quite distinct from the P300 and associated waves (*4*).

In these experiments, each subject read 160 different seven-word sentences presented one word at a time. Subjects were instructed to read silently, "in order to answer questions about the contents of the sentences at the end of the experiment." In the first two experiments, a random 25 percent of the sentences ended in a semantically inappropriate (but syntactically correct) word. The degree of semantic incongruity was "moderate" in experiment 1 (for example, "He took a sip from the *waterfall*") and "strong" in experiment 2 ("He took a sip from the *transmitter*") (*5*). In experiment 3, all sentences ended with semantically appropriate words, but a random 25 percent of these were larger in letter size than the preceding words (*6*). Different subjects were used in each experiment (*7*).

The timing of stimulus presentation is indicated in Fig. 1A. Each sentence was forewarned with a slide containing XXXXX and followed by a 2-second interval. Slides containing the words were flashed for 100 msec at a rate of one per second. After ten practice sentences were exposed, 16 experimental blocks of ten sentences each were administered while the subject reclined in an easy chair.

Scalp electrical activity was recorded using nonpolarizable electrodes from frontal (Fz), central (Cz), and parietal (Pz) locations, all referred to linked mastoids and amplified with a band-pass of 0 to 40 Hz. Eye movements and blinks were monitored by an electrode below the right eye and a bipolar supraorbital to external canthal montage. Trials contaminated with eye movements or blinks were excluded from the average ERP's. Data from all channels were stored on FM tape for off-line signal averaging with a PDP-11/45 computer.

The ERP averaged across the entire seven-word sentence (Fig. 1A) consisted of a sustained negative potential (CNV) upon which were superimposed the phasic ERP's to the individual words. The terminal words were followed by a combination of the positive-going "resolution" of the CNV and the specific ERP's associated with deviant words. The physically deviant, "large" seventh words typically elicited a late positive complex of waves (P560), which was absent following the smaller seventh words. In contrast, the semantically deviant seventh words were followed by a negative component (N400) beginning at about 250 msec and peaking at about 400 msec (*8*) after stimulus onset (dashed waveform).

The grand average ERP's (across all subjects) to the seventh words showed

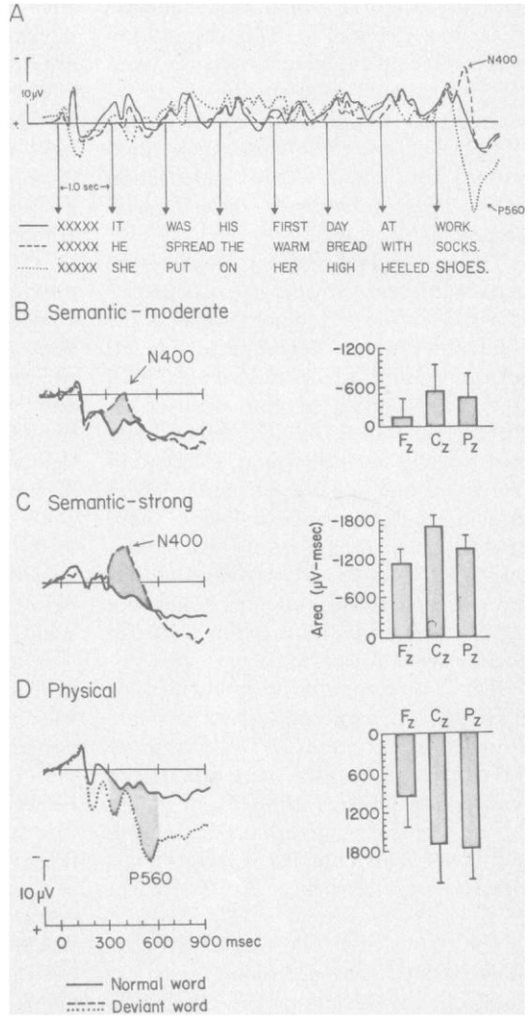

Fig. 1. (A) The timing of word presentations for three sample sentences and typical ERP waveforms recorded over the entire seven-word sentence, averaged over three subjects in experiment 2 (dashed and solid lines) and three subjects in experiment 3 (dotted line). (B–D) Data from experiments 1, 2, and 3, respectively. In each comparison, the grand average ERP's (over all subjects) from Pz for the normal (solid line) and deviant (semantic, dashed line; physical, dotted line) seventh words are superimposed. The 300-msec region used for quantitative analyses of these data is indicated by shading. The means and standard errors of this "difference area," formed by subtracting the normal seventh word ERP's from the deviant seventh word ERP's, are depicted in the bar graphs to the right of the corresponding ERP's.

0036-8075/80/0111-0203$00.50/0

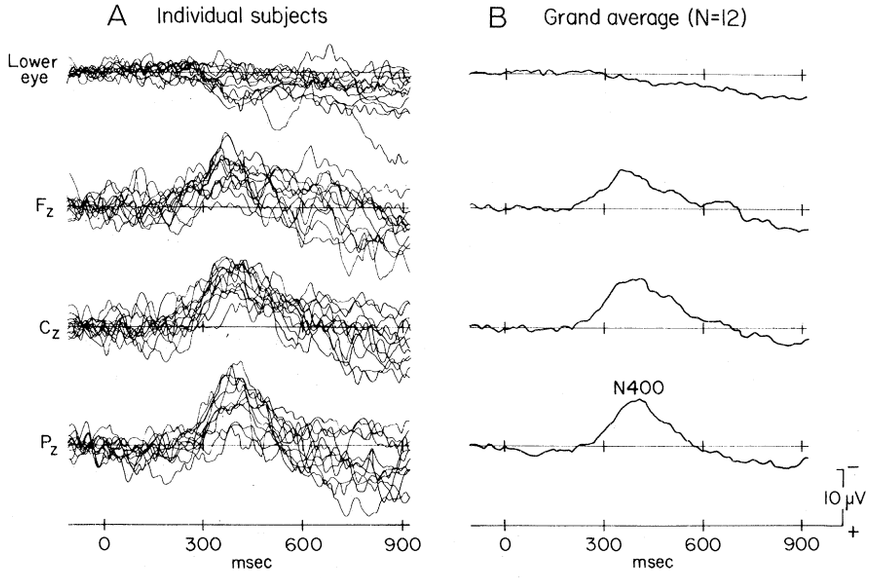

Fig. 2. Difference waves from experiment 2 (strong semantic incongruity). These difference waveforms were obtained by subtracting the averaged ERP's to the semantically congruous words from the ERP's to the semantically incongruous seventh words. Each superimposed tracing (A) represents the difference wave from one subject. The ERP's in (B) are the corresponding grand average waveforms over all 12 subjects.

the N400 component to be substantially larger after the strong semantic mismatches (Fig. 1C) than the moderate mismatches (Fig. 1B). This comparison was significant at all electrode sites, with the N400 quantified as the area in the latency zone between 300 and 600 msec (*9*). Further, it is clear that the N400 constitutes an actual negative deflection above the prestimulus d-c baseline level (Fig. 1C) (*10*).

The consistency of the N400 wave across different subjects in experiment 2 is shown in Fig. 2A. Each tracing is the "difference wave" for a single subject, formed by subtracting the averaged ERP to the appropriate seventh words from the ERP averaged over the semantically incongruous seventh words. Every subject was found to have a negative deflection in the N400 range following these grossly inappropriate words. The grand-average difference waves over all 12 subjects (Fig. 2B) showed that N400 is a monophasic negativity with a broad scalp distribution (*11*).

Figure 1D compares the grand average ERP's to the large and small terminal words in experiment 3. The physically deviant, large words were associated with a late positive complex of waves having three principal peaks (P210, P360, and P560). This complex waveform was characteristic of nearly every subject (*12*).

The results of a follow-up experiment demonstrated that occasional semantic and physical deviations presented randomly within the same series of sentences each elicited their characteristic ERP's—that is, an N400 after semantic deviations and a late positive complex after physical deviations (*13*).

The finding that words out of context elicited an N400 wave was unexpected in light of the evidence that stimuli which disconfirm a subject's expectancies generally elicit an enhanced late positive component (P300) and associated waves (*3*; *4*). It is unlikely that N400 could be a delayed manifestation of the N2 or N200 component that frequently precedes the P300 wave to surprising stimuli, as there was no sign of any additional positivity after N400 (*14*). Nor does the central-parietal scalp distribution of N400 resemble that reported for N2 to unexpected visual stimuli (*15*). It also seems unlikely that N400 is equivalent to the negative phase of the "slow wave" that sometimes follows the P300, because the slow wave is reportedly positive over the posterior scalp (*16*).

A number of purely negative ERP's have been reported in association with various perceptual and cognitive processes (*17*). Of these, the N400 seems most plausibly related to the CNV, which builds up in anticipation of significant stimuli and may be prolonged when those stimuli require more detailed processing (*5*, *18*). If the N400 is a phasic augmenting of the CNV, however, this would be the first time such an effect has been identified as a marker of complex linguistic processing. Whatever the relation of N400 to previously described ERP's, it is clear that semantic deviations activate a different constellation of brain activity than do physical deviations. This physiological dichotomy between P300 and N400 suggests that different modes of cognitive reactions to surprising events need to be delineated.

We cannot yet be certain which aspect of a person's reaction to semantic incongruity is indexed by the N400 component. It remains to be seen whether N400 is specific to semantically inappropriate word or whether it accompanies the violation of other linguistic or nonlinguistic expectancies as well. However, N400 is not a general response to all linguistic or meaningful stimuli because judgments about such stimuli have been specifically associated with the P300 wave (*19*). Rather, the N400 seems to reflect the interruption of ongoing sentence processing by a semantically inappropriate word and the "reprocessing" or "second look" (*20*) that occurs when people seek to extract meaning from senseless sentences. The extent to which this further processing is automatic or requires the conscious switching of attention to the offending word remains to be determined (*21*).

The N400 wave seems likely to provide useful information about the timing, classification, and interactions of cognitive processes involved in natural language comprehension, and possibly in wider contexts as well. Measurement of the N400 may reveal where in a sentence (or paragraph) semantic information of different types is assimilated and may thus contribute to theoretical analyses of reading. Finally, since the presence of N400 during reading depends on semantic processing of a high order, this wave has considerable promise as a clinical tool in the evaluation of reading impairment and language disorders.

MARTA KUTAS
STEVEN A. HILLYARD
Department of Neurosciences, University of California, San Diego, La Jolla 92093

References and Notes

1. K. S. Goodman, in *Theoretical Models and Processes of Reading*, H. Singer and R. B. Ruddell, Eds. (International Reading Association, Newark, Del., 1970); J. Hochberg, in *Basic Studies in Reading*, H. Levin and J. P. Williams, Eds. (Basic Books, New York, 1970).
2. R. E. Schuberth and P. D. Eimas, *J. Exp. Psychol. Hum. Percept. Perform.* **3**, 27 (1977); I. Fischler and P. A. Bloom, *J. Verbal Learn. Verbal Behav.* **18**, 1 (1979).
3. Reviewed by P. Tueting, in *Multidisciplinary Perspective in Event-related Brain Potential Research*, D. Otto, Ed. (Government Printing Office, Washington, D.C., 1978).
4. The P300 component often occurs in conjunction with a preceding N2 or N200 component and a subsequent slow wave [E. Don-

chin, W. Ritter, C. McCallum, in *Event-related Brain Potentials in Man*, E. Callaway *et al.*, Eds. (Academic Press, New York, 1978)].
5. Behavioral assessments of the degree of semantic incongruity were based on subjects' ratings of the last words on a seven-point scale, where 1 was "expected" and 7 was "totally unexpected." The mean rating for congruous words was 1.39 ± 0.13, for moderately incongruous words 6.28 ± 0.16, and for strongly incongruous words 6.96 ± 0.02. The difference between the strong and moderate scores was significant (Mann-Whitney U test, $P < .002$).
6. The small and large letters were 0.89° and 2.64° high.
7. All subjects were normal young adults between 18 and 35 years old who did not know the aims of these studies. The number of subjects run in experiments 1 to 3 were 11, 12, and 9, respectively.
8. The mean peak latencies of N400 at the Pz site were 397 ± 12 msec for the moderate semantic condition and 391 ± 10 msec for the strong semantic condition.
9. As assessed by Mann-Whitney U tests, the N400 area was greater for strong than for moderate semantic deviations at frontal ($P < .01$), central ($P < .025$) and parietal ($P < .05$) electrode sites. Baseline-to-peak measures showed similar levels of significance.
10. Measured as the peak negativity between 300 and 600 msec after the stimulus relative to the baseline voltage averaged over 100 msec before the seventh word, the mean N400 amplitudes were: −6.79 ± 0.74 μV at Fz, −7.14 ± 1.00 μV at Cz, and −7.81 ± 1.49 μV at Pz.
11. The peak N400 amplitudes in the difference waves relative to the baseline before the seventh word were −8.05 ± 0.98 μV at Fz, −9.62 ± 0.72 μV at Cz, and −9.24 ± 1.09 μV at Pz. All statistical evaluations of N400 amplitude to incongruous versus congruous words in experiment 2 were significant beyond the .001 level.
12. The mean latencies of the three peaks at Cz were as follows: P210, 209 ± 4; P350, 351 ± 9; and P560, 558 ± 13 msec. One subject showed a late negative wave instead of these three peaks. The P350 and P560 peaks were greatly attenuated or absent in a control condition where a single, large word was flashed repeatedly [M. Kutas and S. Hillyard, *Brain Lang.*, in press.].
13. The mean N400 areas at Cz were −1213 ± 253 μV-msec for incongruous versus congruous small words and −955 ± 320 μV-msec for incongruous versus congruous large words. The late positive area was 753 ± 492 μV-msec for large versus small congruous words and 794 ± 262 μV-msec for large versus small incongruous words.
14. Late positivity after N400 was assessed in experiment 2 by calculating the difference between the normal and deviant words in area across the 600 to 900 msec range. This difference (for example, 227 ± 270 μV-msec at Cz) was not significant.
15. E. Courchesne, S. A. Hillyard, R. Galambos, *Electroencephalogr. Clin. Neurophysiol.* **39**, 131 (1975); R. Simson, H. G. Vaughan, Jr., W. Ritter, *ibid.* **42**, 528 (1977); M. Haider, E. Groll, G. Studynka, *Exp. Brain Res.* **5**, 45 (1968). All of these late negative waves were followed by a large positive deflection.
16. J. W. Rohrbaugh, K. Syndulko, D. B. Lindsley, *Electroencephalogr. Clin. Neurophysiol.* **46**, 416 (1979).
17. For a review see R. Näätänen and P. T. Michie, in *Human Evoked Potentials Applications and Problems*, D. Lehmann and E. Callaway, Eds. (Plenum, New York, 1979), p. 251; J. Cohen and W. Grey Walter, *Psychophysiology* **2**, 187 (1966); D. Symmes and M. A. Eisengart, *ibid.* **8**, 769 (1971).
18. M. W. Donald, Jr. *Nature (London)* **227**, 1057 (1970).
19. S. A. Shelburne, *Electroencephalogr. Clin. Neurophysiol.* **32** (1972); D. Friedman, R. Simson, W. Ritter, I. Rapin, *ibid.* **38**, 255 (1975); M. Kutas, G. McCarthy, E. Donchin, *Science* **197**, 792 (1977).
20. D. LaBerge and J. A. Lawry, in *Attention: Theory, Brain Function and Clinical Applications*, D. Sheer, Ed. (Erlbaum, Hillsdale, N.J., in press).
21. K. E. Stanovich and R. F. West, *Mem. Cognit.* **7**, 77 (1979).
22. We thank N. Cohen, R. Galambos, J. Hansen, H. Neville, J. Phillips, and S. Van Voorhis for their help. Supported by NSF grant BNS 77-14923 and PHS grant NIMH 1 R01 MH25594.

11 October 1979

Prenatal Exposure to Diazepam Alters Behavioral Development in Rats

Abstract. *Characteristic potentiation of rat locomotion responses and acoustic startle reflexes that normally appear in the third postnatal week was absent in rats exposed to diazepam during the third week of gestation. Loss of these behaviors suggests a long-term effect that may result from changes in cellular development. Tissue undergoing neuronal differentiation may be especially sensitive to drugs that act on the central nervous system, and the period in which differentiation occurs is perhaps critical for the induction of changes that are later expressed as altered behavior.*

Behavioral alterations in offspring that were exposed prenatally to drugs (*1*) and environmental chemicals (*2*) are noted with increasing frequency. Traditionally, studies of the sequelae of prenatal exposure to drugs and chemicals have concentrated on exposure during organogenesis and on resultant malformation. Exposure during other developmental periods, however, may produce functional disturbances. Given the widespread use of ataractic drugs, including diazepam and chlordiazepoxide (*3*), and the clinical reports (*4*) that prenatal exposure to these agents may yield a variety of regulatory dysfunctions in the neonate, we designed experiments to obtain measures of behavioral activity and reactivity in developing rats exposed to diazepam during the last week of gestation (*5*).

Diazepam was administered subcutaneously once daily between 9:00 and 10:00 a.m. to female rats (Long Evans) on days 13 to 20 of gestation; the females were pregnant for the first time, having been bred at 100 days of age with male rats of the same strain. Day 0 of gestation was the day on which a vaginal smear positive for sperm was first obtained (*6*). On day 13, the pregnant females were weighed and assigned to one of five groups: uninjected controls (N = 15); vehicle-injected controls (N = 3); and three experimental groups given diazepam (2.5 mg/kg, N = 2; 5.0 mg/kg, N = 3; and 10.0 mg/kg, N = 2) (*7*). On day 20, all females were placed in plastic cages for parturition and, beginning on day 21, were checked every 3 hours from 9:00 a.m. to 10:00 p.m. The day of birth was designated as postnatal day 0.

Within 24 hours after birth, the litters of both diazepam-injected and control rats were reduced to seven or eight pups and given to uninjected dams for fostering (each litter had one dam). Behavioral testing was begun on postnatal day 12.

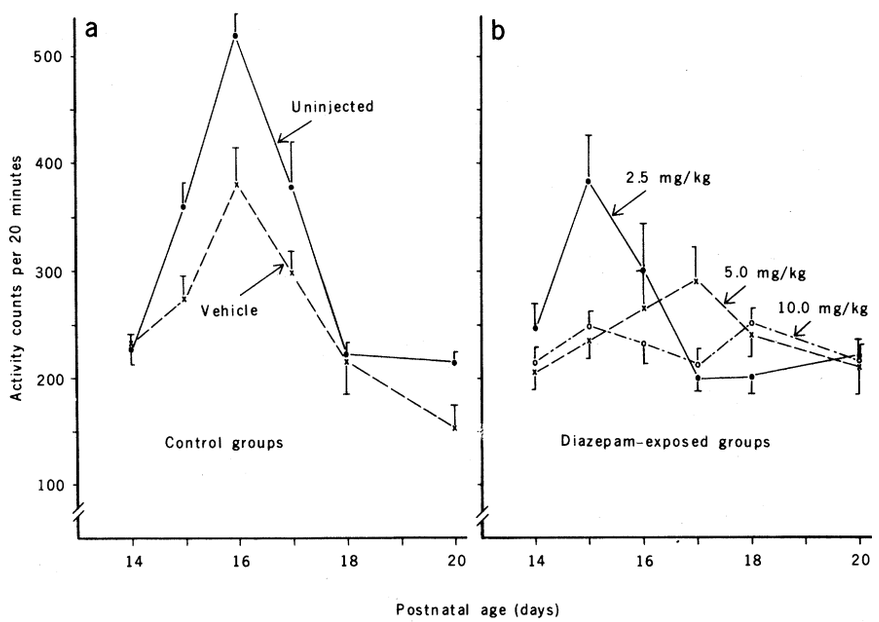

Fig. 1. Development of spontaneous locomotor activity in control and diazepam-exposed rats. Each animal was tested for a 20-minute period on each testing day. Numbers of rats tested were as follows: uninjected, 15; vehicle-exposed, 12; diazepam-exposed, 7, 8, and 7 to 2.5, 5, or 10 mg/kg, respectively. Animals tested represented a minimum of two litters per group. Vertical bars indicate standard error.

0036-8075/80/0111-0205$00.50/0