VARIATIONS IN THE LATENCY OF P300 AS A FUNCTION OF VARIATIONS IN SEMANTIC [1] CATEGORIZATIONS

M. KUTAS [2] AND E. DONCHIN

Cognitive Psychophysiology Laboratory, University of Illinois, Champaign, IL, U.S.A.

Several recent studies of P300 (Ritter and Vaughan 1969, Tueting et al. 1970, Squires et al. 1977) utilized the following experimental paradigm: the subject is presented with a stream of stimuli, each of which may belong to one of two categories. The subject counts and reports the number of stimuli that belong to one of the two categories. The probability is low that a stimulus will belong to one of the categories and correspondingly high that it will belong to the other category. Stimuli in the low-probability category were found to elicit an enhanced P300 component. In most experiments reported to date, categorizations were based on physical features of stimuli, such as the frequency of tones, hue of light flashes, or specific pattern of the visual stimulus. It seemed necessary to determine if the same results would be observed if the categorization required of the subject was based on semantic features of the stimuli (cf. Friedman et al. 1975).

For this purpose, subjects were presented with sequences of words, each of which could be categorized, on the basis of a semantic rule, into one of two categories. The two categories appeared with the probability of either 0.20 or 0.80. The intent was to determine the extent to which the appearance of stimuli belonging to the rare category would enhance the P300 component. As the complexity and latency of the categorization response varied with semantic categories, the relationship between the duration of cognitive operations and the latency of the P300 component could also be examined. If P300 reflects specific cognitive processing activity (Donchin 1975, Donchin et al. 1973), then the latency of P300 relative to the physical stimulus would depend upon the latency and duration of the cognitive process and would vary as a function of its complexity. This proposition was tested in the present study.

Methods

The experiment utilized PLATO, a computer-assisted instruction system developed at the University of Illinois. The PLATO terminal uses a plasma panel for display (Smith and Sherwood 1976). The display is achieved by illuminating any of 512 x 512 luminous dots. In this experiment, the PLATO system was programmed to present a sequence of words on the terminal, one at a time every 2000 msec. Each word was preceded by an external trigger, which was led to a PDP-8/E computer. The trigger activated the digitizing process so that EEG data would be acquired in relation to the presentation of the stimuli.

Data from three studies are reported. In each study, subjects were presented with four different sequences of words. Each sequence consisted of about 200 words selected randomly on each trial with the appropriate probability. The following series were used:

1. *Fixed names.* The words were either "Nancy" or "David." "Nancy" appeared 20% of the time.
2. *Variable names.* Words were selected from a list of 20 female or 20 male names. Each name was a two-syllable word consisting of five letters. Twenty percent of the names were selected from the female name list and 80% from the male name list.
3. *Rhymes.* Words were selected either from a list of different words rhyming with "cake" or from a list of 20 four- or five-letter nonrhyming one-syllable words. Rhymes were presented 20% of the time.[3]
4. *Synonyms.* Words were selected from a list of 20 arbitrarily chosen words and 9 synonyms of the word "prod." Synonyms were presented 20% of the time.[4]

---

[1] An extended report of these experiments has been published elsewhere (Kutas et al. 1977).

[2] Present address: Department of Neurosciences, University of California Medical School, La Jolla; CA, U.S.A.

[3] The nine words rhyming were: bake, rake, stake, lake, make, fake, wake, flake, and take. In the second experiment, the words ache and steak were included.

[4] The synonyms presented were goad, poke, shove, nudge, urge, push, prompt, spur, and press.

Brain potentials were monitored with Burden electrodes placed at Fz, Cz, Pz, C3, and C4 referred to linked mastoids in the first experiment and to the chin in others. EOG was recorded between supraorbital and canthal positions. The subject was grounded on the forehead. EEG was recorded on analog tape with a 2-sec time constant and 30-Hz high-frequency cutoff amplitude. Data were digitized off-line by an IBM 1800 computer and stored on digital magnetic tape. Digitizing started 220 msec prior to the stimulus and ended 780 msec after the stimulus. The sampling rate was 10 msec per point. A PDP-8/E computer determined, on line, whether eye movement artifacts were present during each trial by comparing EOG variance to a criterion value.[5] Contaminated trials were not included in the average (the synchronizing pulse on the analog tape was inhibited).

Subjects sat in a comfortable chair in a semi-darkened, shielded room and completed the four conditions in the following order: fixed names, variable names, rhymes, and synonyms. Prior to each block of trials, the subject was instructed to watch the words and count stimuli from the rare category. At the end of each run, the subject was asked to report his count. Each condition consisted of approximately 200 trials (40 rares).

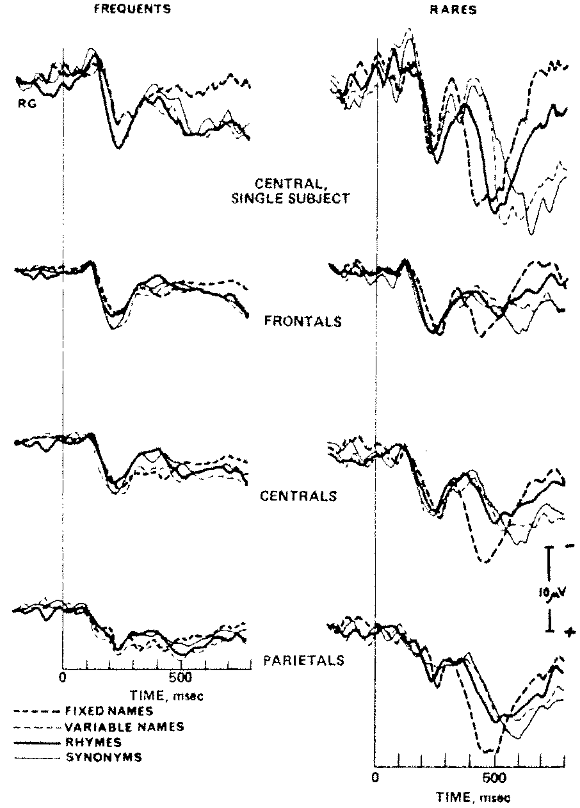

Fig. 1. Superimpositions of average ERPs obtained during experimental conditions requiring different semantic categorizations. At the top are sample Cz waveforms to the rare and requent stimuli from an individual subject. The remaining waveforms are superaverages across six subjects for Fz, Cz, and Pz positions. Only the rare stimuli (prob. = 0.20) were counted. Each rare waveform consists of approximately 18 to 30 single trials.

Results and discussion

Experiment 1

Six subjects participated in the first experiment, in which the task was to count the number of stimuli belonging to the rare category. Fig. 1 presents the data from one subject and superaverages computed over the entire subject group.

Evoked responses elicited by stimuli belonging to the frequent category lack a P300 component, while ERPs elicited by stimuli belonging to the rare category show a marked P300 component. Latency of the P300 component varied widely—shortest for the fixed name categorization and the longest for synonyms. Variable names and rhyming words showed intermediate latencies. The same order of latencies characterized data from all subjects. P300 amplitude elicited during the fixed-name condition was larger than in the other three conditions. There were no systematic differences among the amplitudes of the average P300s.

These data demonstrated that the P300 response was associated with the categorization of stimuli even when categorization depended upon semantic rather than physical characteristics of stimuli. Although averaging was over a diverse array of physical stimuli, such as different female names, a clear P300 response was elicited. The data were consistent with the suggestion that P300 latency varies systematically with the complexity of information processing required.

Experiment 2

To validate the extent to which differences between series were related to different subject decision times, the experiment was repeated with five other subjects, who were asked to respond rapidly by pressing one of two buttons upon the appearance of any stimulus from the two categories. The results, shown in Fig. 2, differ from those obtained in Experiment 1 in two important respects. First, it appeared that the execution of a motor response changed the appearance of the "frequent" evoked response in that

[5] Trials were rejected if the sum of squared digitized values (220 msec prestimulus and 780 msec poststimulus) exceeded a criterion value determined by visual inspection of a large sample of EOG traces and their corresponding digital values.

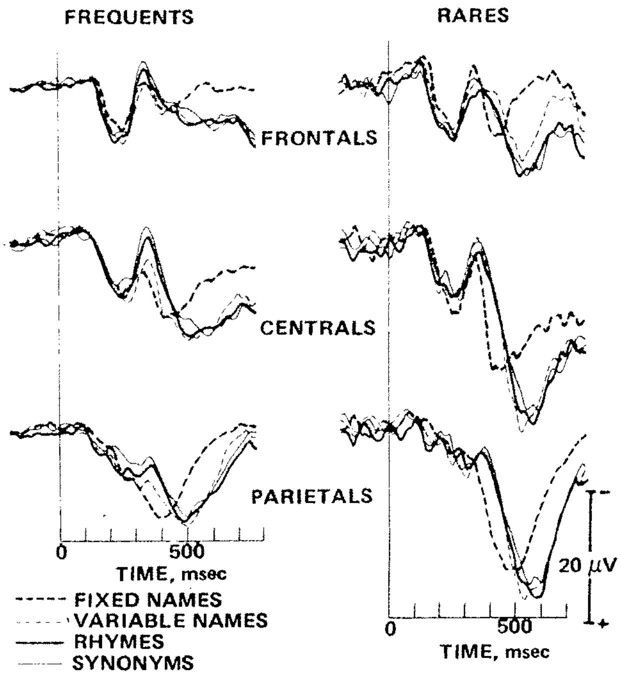

Fig. 2. Superimpositions of ERPs averaged across five subjects from four different experimental conditions. Subjects were required to perform a choice reaction time response, responding to frequent stimuli with one hand and rare stimuli with the other.

a positive component, presumably P2 of the motor potential, could be observed. This positivity could not fully account for the marked enhancement of positivity with a latency of about 400 msec associated with rare stimuli. Latencies obtained in this second experiment, however, were somewhat less differentiated than those observed in the first experiment. The fixed-name latency was still considerably shorter than that associated with the other three conditions; however, these three conditions were no longer as differentiated as they were during the count condition. Table 1 presents means and standard deviations of reaction times (RT) averaged over the five subjects for each of the experimental conditions. These means are based on RT scores obtained from all trials on which subjects responded correctly. The large variance of RT is noteworthy. The fixed-name RT was substantially shorter than the RT of the other conditions, which were essentially equal. The variance of P300 latency (Fig. 2) was equally large, suggesting a substantial degree of trial-by-trial variation in P300 latency and RT. An analysis of the relationship between P300 latency and RT is presented by Kutas et al. (1977).

Failure to observe differences in P300 latency with three of four experimental conditions could be attributed to the fact that in the second experiment speed of response was emphasized without requiring accuracy in categorization. Thus, subjects tended to execute erroneous categorizations as they attempted to maximize response speed. The error rate varied across experimental conditions, with fewest errors occurring during the fixed-name condition and most occurring during the synonym condition. Clearly, subjects could have traded accuracy for speed. A third experiment was therefore run to assess this possibility.

Experiment 3

In the final experiment, the rhyme series was not used. Five subjects participated under three experimental conditions with each of the remaining series. The "count" condition replicated Experiment 1 and the "RT-accuracy" condition subjects made a choice reaction to the stimulus, but were instructed to be very accurate. The results are shown in Fig. 3. When accuracy was emphasized, P300 latencies varied in the same manner as in the count condition. When speed was a prime consideration, subjects seemed to maximize speed by reducing processing time invested

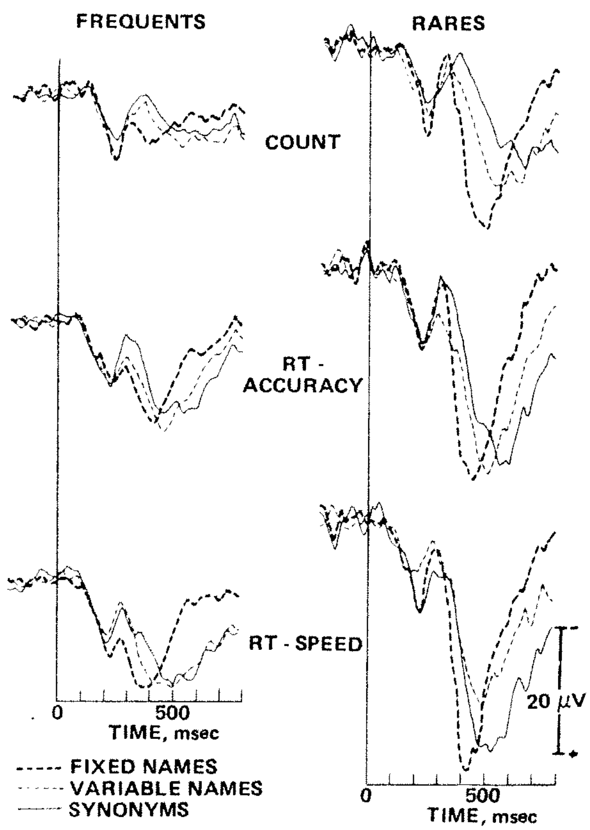

Fig. 3. Superimposition of central (Cz) ERPs averaged across five subjects for three semantic categorizations obtained during three different response regimes.

Table 1. Means and Standard Deviations of Reaction Times, Experiment 2 (RT-speed, five subjects)

Condition	Rare stimuli		Frequent stimuli	
	Mean RT	Standard deviation	Mean RT	Standard deviation
Fixed names	514.41	83.31	419.43	96.88
Variable names	613.25	89.82	522.22	107.24
Rhymes	633.48	123.79	524.22	125.39
Synonyms	666.55	109.43	513.48	105.78

Table 2. Grand Means and Standard Deviations of Reaction Times, Experiment 3 (Count, RT-accuracy, RT-speed, five subjects)

Condition	Rare stimuli		Frequent stimuli	
	Mean RT	Standard deviation	Mean RT	Standard deviation
Fixed names-accuracy	543.56	111.19	453.27	110.09
Variable names-accuracy	573.03	164.67	498.79	99.38
Synonyms-accuracy	619.73	100.60	520.25	119.56
Fixed names-speed	455.47	77.23	353.15	102.89
Variable names-speed	506.92	100.65	409.29	92.33
Synonyms-speed	531.43	94.70	413.75	96.99

in the categorization, thereby reducing the variability in P300 latency. Grand means and standard deviations of RT for the six conditions are presented in Table 2.

Conclusion

The latency of the late positive component (P300) associated with rare occurrences of relevant stimuli varies with stimulus evaluation time. The differences in P300 latency cannot be attributed to the effects of the motor response on P2. Latency differences observed during the count condition were quite similar to those observed in the RT-accuracy condition, yet no manual responses were required during count conditions. The data are consistent with the view that the variable-latency parietal-maximum, late positive waves are manifestations of the activity of the same intracranial processor. Alternate views that tend to differentiate between late positive components by their latency (e.g., Thatcher 1977) seem less parsimonious.