AN ELECTROCORTICAL INVESTIGATION OF WORD RECOGNITION IN A BACKWARD MASKING PARADIGM.

by

DANIEL ULRICH BRANDEIS

DIPL. NATW., EIDGENOESSISCHE TECHNISCHE HOCHSCHULE ZUERICH, 1979

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

in
THE FACULTY OF GRADUATE STUDIES
PSYCHOLOGY DEPARTMENT

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
AUGUST 1982

© DANIEL ULRICH BRANDEIS, 1982
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date August 9, 1982
Abstract

Three aspects of stimulus content, i.e. meaningfulness, familiarity and task relevance, were manipulated without the subjects' awareness. A number of subliminal (backward-masked) stimuli were presented to the subject whose task it was to estimate an interval of 1 sec (starting with the presentation-flash) by pressing a button. Supraliminal words were randomly interspersed among these subliminal stimuli, appearing above or below the masked field. Whenever the subject detected a previously assigned target among the supraliminal stimuli, he/she was required to press the button as fast as possible. The meaningfulness of the subliminal material was manipulated using words, nonwords and blanks. Three groups of words were used: the targets, the nontargets and other, 'new' words (which were never presented supraliminally). Task relevance (targets vs. nontargets) and familiarity ('new' words vs. other words) were thus manipulated. Unexpectedly, detection performance was better with words than with nonwords. This suggests that detection is a late process drawing on lexical information. Several components of the event-related potential (ERP) differentiated as early as 140 msec poststimulus between sub- and supraliminal conditions. More importantly, differences within the subliminal conditions were observed: familiarity was discriminated after 260 msec and simple presence of a string after 300 msec. These results are consistent with the conclusions drawn from detection performance, and they support the notion that backward masking does not disrupt processing.
# Table of Contents

Abstract ................................................................. ii
List of Tables ......................................................... iv
List of Figures ........................................................... v
Acknowledgement ....................................................... vi
I. INTRODUCTION ...................................................... 1
   1.1 Detection And Recognition ..................................... 4
   1.2 Pattern Masking .................................................. 6
   1.3 Coding And Recovery Of Information ......................... 9
   1.4 Rationale For The Present Study .............................. 17
II. METHOD ............................................................... 21
   2.1 SUBJECTS .......................................................... 21
   2.2 MATERIALS AND DESIGN ........................................ 21
   2.3 APARATUS .......................................................... 24
   2.4 PROCEDURE ........................................................ 25
   2.5 ANALYSIS .......................................................... 29
III. RESULTS ............................................................ 33
   3.1 Detection Performance .......................................... 33
   3.2 Response Latencies .............................................. 34
   3.3 Event Related Potentials ....................................... 35
IV. DISCUSSION ........................................................ 41
   4.1 Statistical And Methodological Considerations ............ 41
   4.2 Detection Performance .......................................... 42
   4.3 Time Estimates .................................................. 47
   4.4 Event Related Potentials ....................................... 47
   4.5 General Discussion ............................................. 50
BIBLIOGRAPHY .......................................................... 58
APPENDIX A - LIST OF STIMULI ..................................... 64
APPENDIX B - GROUP MEAN ERPS AT P4 FOR ALL CONDITIONS .... 65
APPENDIX C - GROUP MEAN ERPS AT OZ FOR ALL CONDITIONS .... 66
APPENDIX D - GROUP MEAN EOG FOR ALL CONDITIONS .............. 67
APPENDIX E - AVERAGED ERPS FOR ONE SUBJECT (S1) AT P3 .... 68
APPENDIX F - AVERAGED ERPS FOR ONE SUBJECT (S2) AT P3 .... 69
List of Tables

I. Detection performance: mean probability of hits and d's ..34
II. Mean time estimates and invalid trials ......................35
III. F-ratios for contrasts on time estimates ..................35
IV. Mean ERP component latencies and standard deviations .....36
V. Mean ERP component latencies at P3 for all conditions ....37
VI. Mean ERP component amplitudes at P3 for all conditions ...38
List of Figures

1. Sequence of events in a single trial ..................25
2. Set of contrasts ........................................37
3. Group average ERPs at P3 for all conditions ..........39
Acknowledgement

I would like to thank my committee members, Dr. A.M. Treisman and Dr. D. Kahneman for their encouragement and the helpful theoretical guidance throughout the study. I am grateful for the help of Jeff Jutai and John Lind, who introduced me to the Lab and assisted generously whenever needed with practical help and discussions. Special thanks go to my supervisor, Dr. Robert Hare, whose sensible advice, constant support and generosity were invaluable. This research was supported by Grant No. MT 4511 from the Medical Research Council of Canada to Dr. R.D. Hare while the author was on a University of British Columbia Graduate Student Fellowship.
I. INTRODUCTION

While reading in the adult is a highly overlearned skill, remarkable in its speed and efficiency, there is no clear evidence on the extent to which it reflects unique, specialized or universal cognitive functions. The present study is concerned with one aspect of reading, word recognition, which preserves a central feature of reading: the transformation of a sequence of arbitrary symbols into subjective meaning.

Cognitive psychologists have been reasonably successful in treating word recognition within a general framework of information processing, thus relating it to other forms of perception. It is generally assumed that word recognition proceeds in stages, with some amount of preconscious feature extraction preceding identification, recognition and, where required, initiation of a response. The efficiency of word recognition may be partly explained by assuming that it is automatic to a large extent; we seem able to achieve a meaningful representation of printed words quickly and with only little attention and effort. Posner and Snyder (1975) argue that

(1) Automatic processes precede controlled processes involving conscious attention;
(2) they may occur without intention;
(3) they may occur without interfering with other ongoing mental activity; and
(4) they may occur without giving rise to conscious awareness.
A recent report by Marcel (in press) has shed light on the relation between word recognition and the involvement of consciousness in processing. He presented evidence that lexical or semantic access for words under viewing conditions preventing conscious awareness is possible. In brief, Marcel (in press) claims that:

(1) At a stimulus onset asynchrony (SOA, the interval between onset of the target and onset of the masking stimulus) yielding chance performance on both presence/absence and graphic similarity judgements, semantic similarity judgments are still above chance, which means that there is a 'sensitivity' to the meaning of unseen words;

(2) A subliminal word (i.e. masked to yield chance detection performance) can produce significant semantic facilitation or interference in a Stroop (color naming) task;

(3) A subliminal prime word can facilitate a lexical decision about the target, thus producing semantic priming. The effect is obtained with pattern masking only;

(4) Repetition of a subliminal prime increases priming but leaves detection performance unchanged at chance level.

Marcel (in press) argues that different aspects of a word (e.g. form, location, lexical identity, colour) are coded independently and in parallel. For each of these aspects or
domains, processing proceeds automatically up to the highest possible level of representation. The resulting multiple and separate codes must be recovered by elaborating processes that integrate them into a unitary percept. Pattern masking supposedly interferes with the figural record that is essential for conscious perception. Since records decay over time, the highest level codes achieved last are easiest to recover. Partial support for these claims comes from studies by McCauley, Parmelee, Sperber & Carr (1980) and Fowler, Wolford, Slade & Tassinary (1981) but the validity of the findings has also been questioned (Merikle, 1982; Purcell, Stewart & Stanovich, Note 1). If Marcel's argument is tenable it would have consequences that go beyond issues limited to word recognition. More specifically, Marcel questions commonly held beliefs by claiming that:

(1) detection may not be a prerequisite for recognition;
(2) pattern masking may not interrupt processing; and
(3) perceptual effects believed to reflect encoding might actually reflect recovery processes.

The present study approaches these issues by examining both the cortical event related potential (ERP) and the behavioral responses of the subjects to subliminal stimuli. Several components of the ERP have been associated with functional stages in information processing (Donchin, Ritter & McCallum, 1978). Since the temporal ordering of events in the information
processing sequence is crucial to Marcel's position, the ERP, with its high temporal resolution, might be a particularly suitable measure for this purpose.

1.1 Detection And Recognition

Detection and recognition are typically assumed to be hierarchically organized, i.e. detection is viewed as a prerequisite for recognition. This leads to the prediction that recognition must be at chance levels when a stimulus goes undetected (Coren, Porac & Ward, 1979). For signal detection theory and its extension to recognition however (Green and Birdsdall, 1978), recognition without detection is a teneable outcome. Rather than being hierarchically organized, detection and recognition are conceived of as reflecting the same basic process, one in which a set of internal detectors is being monitored. In detection, any of these detectors is compared to a preset criterion. In recognition, the detector with the largest output in the set is chosen. The theory makes the counterintuitive prediction that recognition performance can be superior to detection performance, a prediction that has been confirmed in several experiments (e.g. Diener, 1981). Application of the theory has so far been limited to simple stimuli (e.g. tones of a certain pitch) for which specialized detectors are likely to exist. An important question arises when more complex stimuli, such as words are considered: what are the relevant detectors that are being monitored? Obviously when detectors are being monitored for identification or recognition, the identity of single words must be specified.
Probably only word detectors at a lexical level conform to this requirement. For detection, however, a small set of feature detectors could signal presence of print, but a set of letter or word detectors would serve the same purpose. What signal detection theory predicts for the relation between detection and recognition holds only if the same sets of detectors, at the same functional level, are monitored in both detection and recognition tasks. The counterintuitive prerequisite for an application to word detection thus would be that to detect a word, it is not the early encoding processes decomposing the word into graphic features that one monitors, but rather the later output at the lexical level. If this assumption holds, any experimental manipulation affecting lexical access, regardless of whether it does so by influencing the feature analysis or by affecting a later stage, would have to influence detection.

Recent research involving the ERP is relevant to this discussion of detection and recognition. Two components of the ERP, N100 and P300, have consistently been shown to manifest successive stages of selection in information processing, with N100 (a negative going wave peaking at around 100 msec) being sensitive to early channel selection according to simple stimulus features, and P300 (a positive going wave peaking around 300 msec) reflecting selection of a specific target within the appropriate channel (Picton, Campbell, Baribeau-Braun & Proulx, 1978). Parasuraman, Richer and Beatty (1982) predicted that N100 would be sensitive to detection only, while
P300 should vary with both detection and recognition. In their experiment both detection and recognition judgements of tones differing in pitch were made on each trial. The results provided clear support for their hypothesis: N100 varied with detection only, while P300 was affected both by detection and recognition. In addition, a slow positive shift (SPS), which was identified by principal component analysis (PCA) and peaked after P300, varied with recognition only, regardless of detection confidence. An important finding was that recognition was above chance on trials where (doubtful) 'absent' judgements were made, i.e. when subjects were fairly confident that the target was not present. Moreover both P300 and the SPS yielded discrimination between correctly and incorrectly identified targets. Thus both behavioral judgements and the ERP suggest that detection and recognition are concurrent rather than sequential processes.

1.2 Pattern Masking

Backward masking with patterned masks can profoundly interfere with word recognition. The locus of this form of masking is central, since masking can be obtained dichoptically, with stimulus and mask each being presented to a different eye (Turvey, 1973) This suggests that pattern masking cannot adequately be explained in terms of degradation of the stimulus by the mask. Turvey (1973) argues that the pattern mask interrupts the processing of the target. The findings of Marcel (in press) suggest a different interpretation: pattern masking affects figural codes or records, but leaves lexical or semantic
records of the stimulus largely unaffected.

McClelland & Rumelhart (1981) and Rumelhart & McClelland (1982) have developed an interactive activation model for letter and word recognition. It specifies the effects of pattern masking on print in a way which seems basically compatible with Marcel's (in press) interpretation. Their model is hierarchical, in that it consists of a feature-, a letter- and a word-level. Despite its advanced features, it bears a strong similarity to the influential 'logogen' model (Morton, 1969), in that the word level consists of an array of detectors or nodes, each of which accumulates evidence for a single word. However, contrary to the logogen model, activation and inhibition are possible between and within levels, i.e. feedback from higher to lower levels is implemented and the nodes have no individual thresholds. The effect of pattern masking in this interactive model is to limit amount of time for which a 'legible' representation of the stimulus is available. Information at the word level is less affected by interference from incompatible mask-features than at the feature level, because feedback from the word level can sustain activation of the proper letter nodes, which in turn reactivate the proper word node. Thus, the model explains why higher level, lexical or semantic information is least affected by pattern masking. One could argue that with proper activation at the letter level some figural information should be available. However, representations at the letter level probably must be thought of as abstract rather than feature-preserving; Evett and Humphreys (1981) make a strong
case for access to the lexicon via an abstract graphemic code, where upper- and lower-case letters are indistinguishable. Accordingly, figural information might not be directly available at that level. It may well be that the readouts from letter and word level must be integrated with less abstract, figural information to yield the final percept (Allport, 1977; Evett and Humphreys, 1981).

The effects of backward masking on the ERP have been explored in some detail, and the conclusions drawn are of major importance for the present experiment, although they are solely concerned with the impairment of figural perception. Andreassi, DeSimone & Mellers (1976) and Andreassi, Gallichio and Young (Note 2) reported attenuation of a positive component around 200 msec poststimulus, specific to perceptual masking, under masking conditions similar to metacontrast (i.e. target and mask have adjacent, nonoverlapping contours). The attenuation was specific to the ERP recorded from Oz (central occipital location). Masking was assessed by asking the subjects to draw the sequence of stimuli they saw. Perceptual masking was defined as the failure to draw the targets. Schwartz, Whittier & Schweitzer (1979) and Schwartz & Pritschard (1981) examined ERP under both monoptic and dichoptic viewing conditions. They reported consistent residuals (i.e. waveform differences between target, followed by the mask and mask alone) with discrimination between letters close to chance. However, the residual did not change when a second masked followed, while discrimination significantly improved. Similarly, over a range
of SOAs yielding a U-shaped masking function, the residual monotonically decreased with smaller SOAs. The authors concluded that the ERP does not correlate with performance or subjective experience in backwardmasking. The findings of these two groups (Andreassi et al., 1976, Note 2; Schwartz et al., 1979, 1981) are clearly at odds. While there are procedural differences which make a direct comparison difficult (e.g. the former group used central, the latter peripheral stimulation), it is clear that Schwartz et al. (1979, 1981) used a far narrower range of SOAs to induce changes in perceptual masking (i.e. 0 to 60 msec in six steps as opposed to 10, 40 and 110 msec by Andreassi et al., Note 2), along with a more sensitive detection measure (d' as opposed to the drawings in Andreassi et al., 1976, Note 2). The findings of Andreassi et al. (1976, Note 2) are therefore more difficult to interpret. However, an additional problem characterizes the research of both groups: the conditions were not randomized, but rather presented in blocks, which clearly allows for different state variables (arousal, expectations) across conditions. To summarize, it is unclear whether correlates for perceptual experience under backwardmasking can be found in the ERP. Similarly, it is unclear whether a residual, of whatever functional significance, is present under conditions where target detection is at chance levels.

1.3 Coding And Recovery Of Information

A central aspect of the work of Marcel (1980; in press) is that pattern masked words are shown to produce priming without
becoming conscious. Priming refers to the finding that some responses to a word (e.g., a lexical decision, the classification of a letter string as a regular word or a nonword) are speeded when an appropriate context (e.g., visually similar or semantically associated) precedes this word (e.g., Meyer, Schvaneveldt & Ruddy, 1975). Marcel (1980; in press) has argued that semantic priming, as well as word frequency or familiarity, affect word recognition after a word has gained lexical access rather than in the encoding stage, i.e., that the effects occur at a stage where the recovery of lexical information takes place. In his view then, lexical access is largely invariant and determined by the encoding of physical properties of the word only.

This contrasts with the logogen model (Morton, 1969) where the locus of semantic priming is in the encoding stage. In this model, the subjective lexicon consists of an array of detectors or logogens with individual thresholds. Each logogen accumulates evidence for a single word from feature analysers, from other logogens or from the cognitive system. A single mechanism, criterion bias, accounts for semantic priming and for the effects of other variables affecting word recognition, such as stimulus familiarity and graphic context. Words primed by semantically related context, as well as familiar or graphically primed words, all have the criterion of their corresponding logogen lowered, so that these logogen units need less evidence from feature detectors to exceed their threshold. This criterion bias explanation is parsimonious; however, it can not
account for all of the data (Becker and Killion, 1977). Schvaneveldt and McDonald (1981) recently presented evidence supporting Marcel's (in press) view, namely that semantic priming does not affect encoding. They argue that semantic priming is due to a process testing and verifying hypotheses about the identity of the stimulus after lexical access. This process is characterized as memory driven and 'top-down' operating.

However, the immediate verbal context is only one determinant of semantic priming. There is evidence that instructions to attend to certain words can also result in some form of priming. For example, search for targets among distractors is considerably aided when membership in a single category defines targets and separates them from nontargets (e.g., letters vs. digits, Jonides & Gleitman 1972). Karlin and Bower (1976) have extended this result to words, where a search for exemplars of semantic categories, e.g. the search for 'trees' among 'colours', is required. Again, the search is aided when targets belong to one semantic category and distractors to a different one. These 'category effects' have been interpreted in terms of 'set' for a target category (Gleitman & Jonides, 1978): a top-down, or memory driven form of priming. Thus, instructions to attend to a target category may result in semantic priming.

Not every manipulation affecting word recognition does so at a semantic level, as the discussion of the logogen model has already indicated. Priming effects with graphically similar
word- or nonword- primes have repeatedly been demonstrated (e.g. Meyer, Schwanefeldt & Ruddy, 1975; Evett & Humphreys, 1981). Word frequency or familiarity also affects word recognition. Scarborough, Gerard and Cortese (1979) presented evidence that word frequency is a highly plastic phenomenon. Familiarity effects were readily generated by the repetition of words, they lasted for a long time, and they occurred automatically, without subjects realizing that the words had been repeated. However, no transfer between familiarity in different modalities was found. Murrell and Morton (1974) reported similar results. Warren & Morton (1982) have extended these latter findings to picture recognition, where no transfer from words to pictures or vice versa occurred. Since semantic priming is not modality specific, it seems that familiarity operates at a lower functional level than does semantic priming.

To summarize, word recognition is affected by a variety of experimental manipulations. Effects of unconscious words on the recognition of conscious words have yielded important new evidence about the temporal sequence of events in word recognition. It may well be that different functional loci underly the effects of familiarity and of semantic context on word recognition.

Thus far unconscious reading has been examined by manipulating the immediate semantic context, as in the Stroop or the semantic priming paradigm (Marcel, in press). It may be just as important, however, to examine how other procedures which affect word recognition (e.g. the ones mentioned above,
semantic set, graphic priming or familiarity) in turn affect the unconscious processing of words. This would lead to more conclusive evidence about the temporal ordering of the functional events in word recognition. Equally important, it would also help to characterize unconscious processes and to explore their limits.

There has been very little ERP research on the dynamics of information processing involved in word recognition. This does not mean, however, that research on ERP and linguistic stimuli is sparse, only that this research is not directly concerned with word recognition.

A few articles have examined ERP differences between words and nonwords. Molfese, Papanicolaou, Hess & Molfese (1979) compared the ERP's to auditorily presented consonant-vowel-consonant (CVC) syllables, which could be meaningful [kaeb] or meaningless [kaek]. Subjects were to press a key if they heard a word and another key if they heard a nonsense syllable. They found differences in the factor-analysed ERP for three factors, peaking at 6-60, 240-280 and 320-440 msec poststimulus, respectively. Shelbourne (1972, 1978) compared visual ERPs to meaningful and meaningless CVC-trigrams in normal subjects and in reading disabled children. The letters of the trigrams were presented one second apart. No differences between meaningful and meaningless trigrams were found, perhaps because the letters of the trigram were presented one at a time rather than simultaneously. Molfese et al.'s (1979) positive finding, on the other hand, may have been the result of lower level
differences between stimuli, since the phonemes used for words and nonsense CVCs were not identical. ERP differences between meaningful [kaeb] and meaningless [kaek] stimuli occurred long before the decisive third phoneme of the CVC was presented. Subjects must have correctly anticipated the critical third letter by utilizing coarticulatory cues in the earlier phonemes. Moreover, Sutton (1979) points out that in a task where the subject has two choices (meaningful vs. meaningless), the linguistic meaningfulness of only one set of stimuli might render them more salient or important. To conclude, there seems to exist no convincing demonstration of an ERP correlate of meaningfulness in general.

A sizeable number of studies has been conducted on more specific aspects of meaning, especially on possible ERP correlates of similarities and differences in meaning. Different ERPs to physically identical stimuli with different meanings have been obtained by Johnston & Chesney (1974), Teyler, Roemer, Harrison & Thompson (1973), and Brown & Lehmann (1979). In general, they used ambiguous stimuli (e.g., a figure which could be read as "B" or as 13; ambiguous sounds in sentences like "a pretty rose" vs. "a boatman rows") which were rendered unambiguous by the context. Different ERPs for number vs. letter interpretation and noun vs. verb meaning were obtained.

Several papers have compared ERPs to different classes of meaning. Usually, subjects had to rate presented words on semantic or affective scales. Chapman (1979) and Chapman,
McCrary, Chapman & Martin (1980) found that ERPs distinguished well between words in different semantic classes, representing the poles of Osgood's semantic differential (Osgood, 1964). Interestingly, they also found similar discrimination when subjects only listened to the words. Begleiter, Porjesz & Garozzo (1979) found different ERPs for word groups rated differently on an affective scale.

Megela & Teyler (1978) have tried to assess the relative importance of semantic vs. physical or task specific manipulations for ERP changes. They compared the correlations between ERPs to synonyms ("Throw the small ROCK/STONE") and homonyms ("Shuffle /walk on the DECK"). They found very similar ERPs to the homonyms but virtually random correlations between the ERPs to synonyms. Obviously, in this experiment low order effects were much more powerful than semantic effects.

Thatcher (1977) compared ERPs to synonyms, antonyms and neutral words in a semantic matching task. The subject's task was to indicate which of the three conditions had occurred. No differences between synonyms and antonyms were found, but both these conditions differed from the neutral condition. The factor-analytically identified component yielding the best discrimination peaked around 400 msec poststimulus, discrimination being limited to parietal and temporal leads (P3, T5) on the left hemisphere. More recently, however, Vaughan, Sherif, O'Sullivan, Herrmann & Weldon (1982) have reported ERP differences for synonyms and antonyms in a similar task in a latency range of 250-650 msec. Boddy (1981) analysed ERPs in a
semantic categorisation task. Differences between positive and	negative instances of a category were reflected at a latency
earlier than 212 msec in the N1-P2 component. He interprets the
findings in terms of category priming.

Some of the studies mentioned are open to a number of
interpretations. Thatcher (1978) points out the difficulty of
discriminating between contributions of semantic content and of
invariant processes operating on this semantic content (e.g.
effects of syntactic context), especially when using ambiguous
stimuli in different contexts. In the studies concerned with
different semantic or affective classes, it is difficult to
assess the relative contributions of the specific semantic
content and the task specific context. Interpretations in terms
of differences in emotional involvement, differences in
subjective salience of stimuli, or differences in task
difficulty can not be ruled out. The work of Chapman (1979) and
Chapman et al. (1980) seems to escape this criticism. Their use
of various combinations of rating scales and word classes shows
that the differences for the ERPs to different word classes are
independent of the semantic scale used to judge them, and
persist when the words are simply read. Similarly, Thatcher's
(1977) use of extensive controls leads one to believe that his
conclusions about correlates of semantic matching are warranted.
On the other hand, these latter studies illustrate to what
extent the results depend on the use of sophisticated
multivariate techniques and manipulations of the raw ERP.
Simple visual inspection of the untransformed averaged ERPs to
the different word classes only shows them to be extremely similar. Several researchers have suggested that in order to obtain reliable language related ERP-effects, tasks should be designed to engage subjects in active linguistic processing (e.g. Friedmann, 1978; Megela & Teyler, 1978). However the work of Chapman et al. (1980) indicates that ERPs discriminate equally well between semantic classes independent of whether the subject simply repeats the words or rates them on a semantic scale. It may be that ERP correlates of early linguistic processing would be more apparent in the ERP when the subject is not actively engaged in decision making, since such higher order cognitive activity might obscure effects due to earlier processes such as lexical access. It may even be useful to prevent the subject from engaging in interpretative processes and from building up interpretations of the experimental stimuli, e.g. by interpreting the words and nonwords as being differently important. Backward masking with a pattern mask may be appropriate for this purpose, since it seems to preserve the lexical code of the stimulus word while preventing conscious experience (Allport, 1977; Marcel, in press).

1.4 Rationale For The Present Study

A primary purpose of the present study was to obtain a single threshold for detection of physically similar stimuli (e.g. words and nonwords differing only in the sequence of their constituent letters) in order to insure subliminal viewing conditions. As suggested earlier in discussing the relation between word-detection and word-recognition, it seems
counterintuitive that Ss would rely on other than figural information in order to detect words. The study controlled for this possibility, however, by determining the threshold for words and nonwords with matched physical features.

The second reason for this study is related to the current state of research on ERPs and backward masking. Most experimenters have been concerned with the ERP correlates of perceptual experience. However, perceptual experience is probably based on a complex sequence of information processing, with some processes reflected more and others less directly in ERP components. The supposedly complex structures underlying conscious experience do not lead one to expect a simple relationship with the ERP as a whole. The reinterpretation of pattern masking mentioned above (e.g. Marcel, in press) allows for a more specific question: how is information processing under conditions of pattern masking manifested in the ERP structure? By manipulating instructional set and stimulus familiarity but not conscious experience, the present experiment attempts to examine specific information processing structures and to establish a functional and temporal relationship between ERP structure and processes in word recognition.

The third reason for the present experiment was to explore unconscious processes themselves. Marcel (in press) claims that unseen words access their lexical entries and produce semantic priming, and he accounts for these findings by distinguishing encoding and recovery of information. Word familiarity may affect unconscious processing of words, since it probably
operates at a lower functional level than semantic priming, possibly on the encoding stage. In addition, the present study also examined whether processing advances to the point where contact is made with semantic set, i.e. whether differential processing of targets and their controls can be observed. For this purpose the procedure includes an instructional manipulation which is thought to result in semantic set, a 'top-down' form of priming: the instruction to attend to targets in one semantic category. Semantic set is more likely to aid recovery of lexical information than encoding of the stimulus and thus is expected to affect a later stage of word processing than familiarity. Marcel (in press) has demonstrated that a subliminal prime affects processing of a conscious target. The manipulation of set in the present study can be viewed as a test of whether conscious priming can affect processing of a subliminal target.

Words (and the appropriate controls) were chosen because they supposedly invoke a significant amount of mandatory, automatic processing under pattern masking. The time estimation task was chosen mainly because the delayed response allows recording of uncontaminated ERPs, and because it has been shown to be sensitive to a number of cognitive manipulations, e.g. difficulty of information processing (Fraisse, 1979), stimulus familiarity (Avant, Lyman & Antes, 1975; Avant and Woods, Note 3), stimulus processing time (Thomas & Weaver, 1975). More specifically, Avant and Woods (Note 3) found relative duration judgements to be affected by lexical status of pattern-masked
letter strings yielding chance-discrimination of letter case. In addition, the task required no overt response to the 'unseen' aspect of the stimulus, thereby avoiding the frustrating situation where subjects have to judge something they do not see. It was hypothesized that with decreasing familiarity, i.e. blanks > targets > nontargets > new words > nonwords, processing would be more demanding and take longer, resulting in overestimation of real time, for which the subject compensates by producing shorter 1 sec estimates (Fraise, 1979).
II. METHOD

2.1 SUBJECTS

The subjects (Ss) were six female and six male students. All were right-handed, native English speakers ranging in age from 18 to 23 years (mean = 20.3, standard deviation = 1.5 years). Payment was $5.00 per hour, amounting to $15.00 for the complete session of approximately three hours. All subjects reported normal or corrected-to-normal vision and could easily read tachistoscopically presented words.

2.2 MATERIALS AND DESIGN

The experimental procedure included both sub- and supraliminal conditions. All experimental stimuli were followed by a pattern mask; for the supraliminal presentations, however, the mask was ineffective since the words were typed above or below the masked central field. The complete set of practice and experimental stimuli is given in Appendix A.

The stimuli for the subliminal presentations consisted of blank cards with a central, grey patch of 2 x 0.8 cm, corresponding to a visual angle of 2.4 x 1.0 degrees at a viewing distance of 48 cm. The use of the grey patch was necessary to reduce contrast of the stimuli. With other stimuli masking was not severe enough to prevent detection at reasonably high SOAs (above 20 msec). This probably reflects characteristics of the tachistoscope used. An IBM Selectric was used to type either words or nonwords on the grey patch. Typeface was letter gothic, 12 letters per inch. Words were three to seven letters long, with a four letter word.
corresponding to a central visual angle of 1.0 x 0.4 degrees. Eight words were chosen from each of three semantic categories, i.e. furniture, vehicles and body parts, listed in the Battig & Montague (1969) category norms. Only words with at least ten occurrences per million (Kucera & Francis, 1967) were used. They were matched across categories in terms of mean word frequency (Kucera & Francis, 1967), mean log frequency and the respective standard deviations, as well as in terms of word length and, to the extent still possible, in terms of typicality, i.e. production frequency and position rank. Nonwords were made by scrambling letters of 8 experimental words, at least two from each category, so as to yield nonpronounceable strings bearing little resemblance to the original words. String lengths in this group were matched with the word categories. Each word was used for two identical stimuli to yield 16 stimuli per category.

The pattern mask consisted of a central white field slightly larger than the grey patches containing fragments of letters of the same typeface as used for the stimuli. It was produced by treating a densely typed field of letters in random orientations with liquid paper to fragment the letters. A high quality Xerox copy was made and a field of 2 x 1 cm cut out and glued on the center of a white card.

For supraliminal presentation, the same words were used again. They were typed above or below the grey patch on otherwise identical stimulus cards. Each word occurred in both possible positions, yielding 16 stimuli per category. These words appeared outside the masked field, centered 0.8 cm (1.0
degree) above or below the center of the card.

For the threshold determination, a separate 'detection deck' was used. Stimuli were of the type used for subliminal presentations. Half of the 48 cards had a letter string on the grey patch. Of these, half contained nouns, matched again in frequency and wordlength to the ones in the experimental conditions. No exemplars of the categories used in the experimental conditions were used. The remaining strings were nonwords derived from the corresponding words, as in the experimental condition (with the exception that here every word had a corresponding nonword). The deck was divided into two 'subdecks', each with 50% strings and balanced as well as possible.

Two of the semantic categories were assigned to each subject as targets and nontargets, respectively. Only target and nontarget words were presented supraliminally, i.e. outside the masked field. This led to the following seven experimental conditions: targets and nontargets presented supraliminally, i.e. above or below the center; nonwords, targets, nontargets, and 'new' words (the remaining category, i.e. the words they never saw in supraliminary presentation) presented subliminally, i.e. centrally; and a blank condition (no print) for control. Combinations of target, nontarget and new word categories were counterbalanced across sex, with each possible combination administered once to a female and once to a male subject.
2.3 APARATUS

Stimuli were presented in a Cambridge two-field tachistoscope (t-scope). Luminance of stimulus and masking field were equalized at 3.2 footlamberts (maximal setting).

EEG (Electroencephalogram) was recorded from Oz, C3 and C4 referred to the nose. EOG (Electrooculogram) was recorded between electrodes supraorbital for the left and at the outer canthus for the right eye. Grass electrodes in combination with Grass EC-2 paste were used at all sites. Grounding was via a Beckmann electrode on top of the spine. Resistance was kept below 7kOhms for all leads. A Beckmann R711 Polygraph was used for recording. A time constant of 1 sec and an amplification of 50 microvolts/cm were used for all channels. The signals of the four physiological channels were multiplexed via a Vetter model 4 FM recording adapter and stored on one channel of a cassette tape deck. A 1000 Hz tone indicating warning tone, stimulus onset and buttonpress response was recorded on the other channel. Tachistoscopic exposure and tone signals to loudspeaker and cassette tape were under control of an MCS-85 microcomputer. Trials were initiated by the experimenter. The EOG was monitored throughout the experiment and the S was told about the occurrence of excessive artifacts in the recording, where present.

Pilot studies had indicated that this procedure yielded comparable sensitivity for vertical and horizontal eyemovements.
2.4 PROCEDURE

For all but three Ss the two sessions were separated by a 15 min. break. For the remaining three Ss, sessions were 2 days, 8 days and 15 days apart. The first session was used to determine the S's threshold for detection of backward-masked words. EOG electrodes were attached and the S was screened on the first few trials for excessive eye movement artifacts. These artifacts were present in three Ss, and they were replaced. The S was seated in a dimly illuminated, shielded room. The general instructions about the sequence of events in a single trial were read to the S. Throughout the experiment important instructions were accompanied by a summary presented in the stimulus field of the tachistoscope. Every trial consisted of a warning tone of 750 msec, a waiting period of 1 sec, a stimulus presentation followed at a variable SOA by the pattern mask, a short tone of 10 msec occurring 1 sec after the stimulus presentation, and an end tone of 200 msec, 1.5 sec after stimulus presentation (see Figure 1).

Figure 1 - Sequence of events in a single trial
The S was told not to move and to keep her \(^2\) eyes centered in the middle of the mask-field between warning tone and end tone. A head rest was used to ensure a constant viewing distance (48 cm). Instructions for the detection task were then presented. In general, an attempt was made to prepare Ss optimally for the detection task, and they were given considerable advance information, strategic hints and feedback. The S had to respond vocally with either 'string' or 'blank' and was first given 5 trials at an SOA of 200 msec. Then the detection cards were given to her to look at. She was asked to look carefully at some of the stimuli, since decisions would become very difficult, and was told to verify, that half of the cards contained letter strings. The S was given the assurance, that only this deck would be used for the detection task. An instruction card provided a summary of stimuli and their relative frequency. The following statements were emphasized to optimize performance:

Decide only on presence or absence, use any available cue; don't be afraid of guessing, rely on your feelings; don't give up when it gets frustrating; strings and blanks are half and half, so always use this information to adjust your sensitivity, and avoid using just one response.

A minimum of 3 trials, depending on how confident the S was, were given at SOAs of 150, 120 and 100 msec each. Then SOA was

\(^2\) For the sake of simplicity, 'she' and 'her' will be used in the sense of 'she' or 'he' and 'her' or 'his', respectively.
reduced slowly until the S made a mistake. Feedback was then given on approximately 30 trials in an attempt (generally successful) to improve the S's performance. At the next lower SOA where the S made two mistakes in a row, a 'staircase' method was used for a subdeck (24 trials). For two consecutive correct responses, SOA was reduced by 10 msec, and for two incorrect responses SOA was increased by 10 msec. To obtain convergence towards the highest SOA yielding responses at a chance level an incorrect trial was automatically recorded after 12 trials (thus if a S would have given alternating correct and incorrect responses throughout the 24 trials, the SOA would nevertheless have been increased by one step, after 13 or 14 trials, due to the preset incorrect response between trial 12 and 13). In addition, the SOA was not reduced below 20 msec. The SOA at which the maximal number of trials had been given was determined and 0, 5 or 10 msec were subtracted from it (up to the experimenter). A subdeck was presented at the resulting SOA. Whenever performance exceeded 60% correct responses, the SOA was further reduced, usually in steps of 5 msec, but at least by 2 msec. The first SOA at which less than 60% (i.e. less than 14 items) correct responses were obtained in two consecutive subdecks (i.e. 48 trials) was used for the experimental session.

The S filled out a questionnaire on the threshold determination to indicate strategies, confidence and estimated performance. Ten targets and nontargets of the supraliminal conditions were then presented at this SOA while the S was asked
Subjects were generally able to identify most of them.

The S was then instructed for the time estimation task. She was to press a button 1 sec after the stimulus presentation, i.e. concurrent with the short tone, with the right thumb. No supraliminal trials were given for that purpose. Up to 60 practice trials were given in order to train the S to press reliably within an interval of 900 to 1100 msec poststimulus. The first session lasted between 40 and 80 min.

EEG and EOG electrodes were attached at the beginning of the second session. For Ss entering the second session on a different day, a subdeck of detection trials was administered in order to check the stability of the previously determined SOA. As a result of this procedure, the SOA of one S was reduced by 5 msec. The time estimation task was practised further until performance became reliable again.

Then the S was instructed for the semantic task. She was shown her targets and nontargets, and was told that nontargets were irrelevant, that it was sufficient to decide whether a word was a target, and that, only for targets, she had to press a button as fast and accurately as possible with the middle finger of her right hand. She was instructed not to respond to words other than targets. Practice was given with the supraliminal targets and nontargets. The two tasks were then combined. When the S did not see a word she estimated one sec with the right thumb. When she saw a word she decided whether or not it was a target, in the manner just described for practice trials.
Accuracy was stressed. When practice resulted in reliable performance, physiological recording was started and the experimental trials were given in two blocks of 116 trials (7 x 16, plus 4 practice trials repeated at the end). Presentation sequence was randomized for each block separately, with the constraint that no more than two exemplars of a single supraliminal condition, and no more than 3 supraliminal trials regardless of condition, followed each other. The S was encouraged to relax and to ask for short breaks when her eyes felt tired. Immediately after the recording session, the S filled out another short questionnaire on her strategies and on what she saw in the time estimation trials, and was then shown the instruction card for detection trials and given 48 detection trials (the two subdecks). She then filled out a final questionnaire (similar to the first one on detection trials), was debriefed about the nature of the experiment, and asked for her comment.

2.5 ANALYSIS

Percent correct responses and the signal detection parameter d', indicating sensitivity to a stimulus, were calculated for the pre- and postexperimental threshold sessions. Both response latency data and ERPs were retrieved from magnetic tape. During the experiment response latencies were only available in hexadecimal form and the only records written were on the sequence of trials. In order to avoid possible experimenter bias, data analysis was deferred until the last S was run. Data from both physiological channels and marker
channel were digitized and averaged off line. Sampling rate was 100 Hz for the physiological channels and 3000 Hz for the marker channel. Response latencies were averaged for each condition while the ERPs were averaged for each combination of site and condition. Only buttonpress responses between 200 and 1500 msec poststimulus were included. Time estimation responses outside a latency window of 700 to 1300 msec were discarded as invalid. Targets trials with responses slower than 1000 msec were excluded, as were nontarget trials where a button was pressed.

Planned orthogonal contrasts between the condition means were performed using the BMDP4V program (Dixon, 1981). The time estimates, i.e. the response latencies for the subliminal conditions, were subjected to the following contrasts: Blanks vs. strings ('presence'), words vs. nonwords ('lexicality'), familiar vs. unfamiliar words ('familiarity') and targets vs. nontargets ('task relevance').

Sampling and averaging of the physiological signals was for a period of 3.25 sec, starting with the warning tone (-1750 msec, i.e. 1750 msec prestimulus) and extending to the endtone (1500 msec poststimulus). However, sampling with 100 Hz only applied to a 1600 msec period, starting at -100 msec. Artifact rejection of EOG contaminated trials was based on three criteria:

1. Absolute amplitude: amplitude of EEG or EOG

The supraliminal conditions were not considered, because for nontargets, a no-go situation (i.e. no responses) resulted, and the relatively short latencies for the targets were instructed, obvious and therefore trivial.
exceeded preset upper or lower limits.

(2) Relative amplitude: the absolute difference between amplitude and average amplitude in the EEG or EOG of a trial exceeded preset criteria.

(3) Correlated amplitude: fast amplitude changes exceeded criteria for EEG and EOG simultaneously or with a short lag (± 10 msec).

The P3 derivation was used as the critical EEG channel. A conjunction rule was applied to the relative and the correlated amplitude criterion, i.e. both types of artifacts had to occur for a trial to be rejected. Criteria were adjusted individually when a first run failed to yield flat EOG recordings with less than 20% rejection. A critical period of at least 700 msec poststimulus was scanned. The ERPs of individual Ss were plotted for all combinations of conditions and sites. ERP amplitude was scaled according to recorded calibration signals and referred to a baseline, defined as the average of the 50 msec period preceding stimulus presentation. Components occurring consistently in most of the Ss were identified and scored. Six components, labelled here N1, P1, N2, P2, N3, P3, were identified. Amplitude and latency measures were analysed in contrasts similar to the ones for the time estimate. However, an additional contrast between subliminal and supraliminal conditions was performed, and each comparison was done for a specific site, i.e. at Oz for the blanks vs. strings contrast, and at P3 for the remaining comparisons, presumably involving
linguistic information processing (see also Fig. 2, p.37) ⁴. Contrasts were run with BMDP4V (Dixon, 1981). Analyses were performed on all component latencies, all peak-to-trough amplitude measures and, in addition, N1 and P3 amplitude (referred to the baseline) ⁵. For the physiological data, a significance level of .01 rather than the usual .05 level was adopted in order to reduce the familywise type-I error rate. The .05 level was used when interpreting trends. Contrasts were performed using the error terms corresponding to the individual contrasts rather than the conventional pooled estimate common to all contrasts. Mitzel & Games (1981) have shown that this robust approach sacrifices relatively little power, and that the conventional solution is often inaccurate.

⁴ As mentioned in the introduction, Andreassi (Note 2) found changes in the ERP with backwardmasking specific to Oz; and Thatcher (1977) found effects of semantic relatedness (synonyms or antonyms) at left hemisphere leads (P3 and T5) only. These findings were thus incorporated in the design of the contrasts. ⁵ Peak-to-trough measures were felt to be more immune to baseline fluctuations. N1 and P3 amplitudes were analysed because these bordering components are underrepresented in the five difference-amplitudes, i.e. because they enter only one difference as opposed to two for the other components, and because a major theoretical framework has been built around N1 and P3.
III. RESULTS

3.1 Detection Performance

Of the 14 Ss run originally, two were reruns for Ss whose data were discarded. One S was excluded because in the first threshold session criterion was not reached at an SOA as short as 15 msec. Since it was felt that neither the controlling switches nor the tachistoscope were accurate at such short exposures, SOA was not reduced further. This resulted in 71% correct responses in the postexperimental detection trials. The second S was excluded because there was not enough time to give detection trials at the beginning of the second session, two days after the first session. This S gave 75% correct responses in the postexperimental detection trials and reported having seen one of the target words in the time estimation task.

The following analyses are based on the counterbalanced design with 12 Ss. Experimental SOAs ranged from 23 to 60 msec, with a mean of 36 msec and a standard deviation of 9.7 msec. For all Ss included in the analysis, correct responses in the postexperimental detection trials were less than 60%. Table I summarizes detection performance at the experimental SOA for both pre- and postexperimental threshold.
Table I - Detection performance: mean probability of hits and d's

<table>
<thead>
<tr>
<th></th>
<th>EXPERIMENTAL</th>
<th></th>
<th>POST-EXPERIMENTAL</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ALL</td>
<td>WORD</td>
<td>NONWORD</td>
<td>ALL</td>
</tr>
<tr>
<td>HIT d'</td>
<td>.45 .13</td>
<td>.49 .05</td>
<td>.42 .25*</td>
<td>.48 .11</td>
</tr>
<tr>
<td>M</td>
<td>.07 .26</td>
<td>.13 .47</td>
<td>.12 .30</td>
<td>.10 .22</td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note - M: mean; SD: standard deviation; * p < .05.

The same false alarm rate was used in determining the d' for overall sensitivity as well as for sensitivity to words or nonwords only. Sensitivity measures (d') were analyzed for deviations from chance performance. In the preexperimental session, sensitivity to nonwords was below chance, t(11) = -2.88, p < .05. In the postexperimental session, sensitivity to words was above chance, t(11) = 3.051, p < .05. False alarm rates were .50 for the pre- and .43 for the postexperimental session. For the preexperimental session, Ss responded correctly on 47.4% of the trials; 49.1% and 45.6% were the respective values for words and nonwords. For the postexperimental session, Ss responded correctly on 52.3% of the trials, 54.9% for words and 49.7% for nonwords.

3.2 Response Latencies

Table II shows the mean response latencies for the time estimations and for the target identification, together with the corresponding standard deviations and percentage of invalid trials.
Table II - Mean time estimates and invalid trials

<table>
<thead>
<tr>
<th></th>
<th>BLANK</th>
<th>NONWOR</th>
<th>TARGET</th>
<th>NONTAR</th>
<th>NEWORD</th>
<th>TARGET</th>
<th>NONTAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>939.0</td>
<td>932.1</td>
<td>937.8</td>
<td>944.9</td>
<td>931.7</td>
<td>659.8</td>
<td>--</td>
</tr>
<tr>
<td>SD</td>
<td>38.8</td>
<td>37.4</td>
<td>45.9</td>
<td>32.5</td>
<td>35.1</td>
<td>40.8</td>
<td>--</td>
</tr>
<tr>
<td>-%</td>
<td>2.1</td>
<td>3.4</td>
<td>2.9</td>
<td>2.6</td>
<td>3.4</td>
<td>10.4</td>
<td>8.3</td>
</tr>
</tbody>
</table>

Note - NONWOR=nonword, NONTAR=nontarget, NEWORD=new word.
M=mean estimated interval in msec; SD=standard deviation; -%=percentage invalid trials.

None of the differences between the subliminal conditions was significant; the F values are given in Table III. The associated type-I error probabilities were all above 0.1.

Table III - F-ratios for contrasts on time estimates

<table>
<thead>
<tr>
<th></th>
<th>BLANK vs. STRING</th>
<th>WORD vs. NONWORD</th>
<th>FAMILIAR vs. NEW WORD</th>
<th>TARGET vs. NONTARGET</th>
</tr>
</thead>
<tbody>
<tr>
<td>F(1,11)</td>
<td>0.17</td>
<td>1.00</td>
<td>3.01</td>
<td>1.16</td>
</tr>
</tbody>
</table>

3.3 Event Related Potentials

Six components were identified and scored. They are labeled N1, P1, N2, P2, N3, P3; with 'N' and 'P' corresponding to the polarity (i.e. positive or negative) of the wave. Not all components were equally prominent across subjects, with P2 and N3 often characterized by changes in slope rather than by clear polarity reversals. The windows used for scoring the components were 90-130 msec for N1, 110-180 msec for P1, 160-220 msec for N2, 200-300 msec for P2, 250-340 msec for N3 and 360-
440 msec for P3. An attempt was made to achieve high consistency within a single S's data. In doubtful cases the wave closest to the one scored in the clearest trace of the S's data was chosen. One subject did not show a P3 component. In this case, latencies were defined as the center of the window (i.e. 400 msec) for all sites and conditions. Amplitudes were defined as the average of the three central time points (i.e. 390, 400, 410 msec). Table IV depicts the mean latencies of the scored components, collapsed across Ss, sites and conditions, with the corresponding standard deviations.

<table>
<thead>
<tr>
<th>LATENCY</th>
<th>N1</th>
<th>P1</th>
<th>N2</th>
<th>P2</th>
<th>N3</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>111.1</td>
<td>141.5</td>
<td>183.3</td>
<td>259.1</td>
<td>301.3</td>
<td>400.7</td>
</tr>
<tr>
<td>SD</td>
<td>11.8</td>
<td>12.5</td>
<td>15.6</td>
<td>25.9</td>
<td>23.8</td>
<td>18.4</td>
</tr>
</tbody>
</table>

Sets of contrasts were performed on the component latencies, on the peak-to-trough amplitude measures (e.g. N1-P1, P1-N2...) and on N1 and P3 amplitudes (referred to baseline). Contrasts were essentially identical with the ones done for the time estimates; in addition, subliminal and supraliminal conditions were compared. The following figure depicts the set of contrasts used, along with the site(s) at which they were tested. The first four contrasts involve only means of subliminal conditions.
Figure 2 - Set of contrasts

(1) Blank vs. string: at Oz
(2) Words vs. nonword: at P3
(3) Familiar vs. new words: at P3
(4) Targets vs. nontargets: at P3
(5) Subliminal vs. supraliminal trials: at all sites

Table V gives the mean latency-measures at the P3 derivation, and Table VI does so for the amplitude measures. Figure 3 depicts the group average ERPs at P3. Further ERP averages are shown in Appendices B-D (group averages) and E-F (individual subject averages).

Table V - Mean ERP component latencies at P3 for all conditions

<table>
<thead>
<tr>
<th></th>
<th>SUBLIMINAL</th>
<th>SUPRALIMINAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BLANK</td>
<td>NONWOR</td>
</tr>
<tr>
<td>N1</td>
<td>113.8</td>
<td>117.5</td>
</tr>
<tr>
<td>P1</td>
<td>142.1</td>
<td>145.4</td>
</tr>
<tr>
<td>N2</td>
<td>184.2</td>
<td>185.8</td>
</tr>
<tr>
<td>P2</td>
<td>264.2</td>
<td>259.6</td>
</tr>
<tr>
<td>N3</td>
<td>311.3</td>
<td>310.0</td>
</tr>
<tr>
<td>P3</td>
<td>406.7</td>
<td>399.2</td>
</tr>
</tbody>
</table>

Note - NONWOR=nonword, NONTAR=nontarget, NEWORD=new word. Latencies in msec poststimulus.
For the latencies, the only difference occurred between sub- and supraliminal presentations, with N2 being earlier for subliminal presentations (F=20.02, p < .01).

Table VI - Mean ERP component amplitudes at P3 for all conditions

<table>
<thead>
<tr>
<th>Component</th>
<th>BLANK</th>
<th>NONWOR</th>
<th>TARGET</th>
<th>NONTAR</th>
<th>NEWORD</th>
<th>TARGET</th>
<th>NONTAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASELINE</td>
<td>-8.96</td>
<td>-8.03</td>
<td>-8.99</td>
<td>-8.64</td>
<td>-8.93</td>
<td>-10.20</td>
<td>-9.52</td>
</tr>
<tr>
<td>N1</td>
<td>-1.18</td>
<td>-1.58</td>
<td>-1.28</td>
<td>-0.73</td>
<td>-0.61</td>
<td>-1.08</td>
<td>-0.26</td>
</tr>
<tr>
<td>P1</td>
<td>1.32</td>
<td>0.50</td>
<td>1.07</td>
<td>1.63</td>
<td>1.07</td>
<td>2.29</td>
<td>2.45</td>
</tr>
<tr>
<td>N2</td>
<td>-3.28</td>
<td>-4.64</td>
<td>-4.06</td>
<td>-3.33</td>
<td>-4.03</td>
<td>-2.62</td>
<td>-3.88</td>
</tr>
<tr>
<td>P2</td>
<td>7.67</td>
<td>6.54</td>
<td>5.58</td>
<td>7.06</td>
<td>7.30</td>
<td>4.94</td>
<td>3.60</td>
</tr>
<tr>
<td>N3</td>
<td>5.49</td>
<td>4.33</td>
<td>5.57</td>
<td>6.43</td>
<td>4.99</td>
<td>4.48</td>
<td>2.98</td>
</tr>
</tbody>
</table>

Note - NONWOR=nonword, NONTAR=nontarget, NEWORD=new word. Component amplitudes in microvolts referred to baseline. Signs reflect polarity of voltage.
Figure 3 - Group average ERPs at P3 for all conditions

Note - the upper two tracings depict the supraliminal, and the lower five tracings the subliminal conditions, respectively. Epoch 500 msec poststimulus; negativity upwards; dashed lines (baselines) 10 microvolts apart.
For the amplitude measures, the baseline was more negative (by 1.08 microvolts) for supraliminal as opposed to subliminal presentations (F=11.09, p < .01). This contrast also differed for N1-P1 amplitude (F=11.83, p < .01), N1-P1 being 1.19 microvolts less in the subliminal conditions; for P1-N2 amplitude (F=16.35, p < .01), with P1-N2 by 1.8 microvolts smaller in the subliminal conditions; and for N3-P3 amplitude (F=13.93, p < .01), N3-P3 being smaller again, by 3.49 microvolts, in the subliminal conditions.

Comparisons within the subliminal conditions yielded an effect on P2-N3 amplitude for familiar vs. new words (F=10.42, p < .01, P2-N3 being 1.99 microvolts smaller for familiar words) and on N3-P3 for blank vs. strings (F=11.64, p < .01); N3-P3 amplitude was .98 microvolts smaller for trials presenting a letter-string. A trend was observed for discrimination between words and nonwords in terms of P3 amplitude (F=4.87, p < .05), with P3 amplitude 1.79 microvolts larger (i.e. more positive) for words.
IV. DISCUSSION

4.1 Statistical And Methodological Considerations

The number of statistical tests performed in this study is considerable. On the ERP data alone, 14 sets of 5 orthogonal a priori hypotheses were tested. The adjustment of the significance levels made here is not nearly sufficient to keep the experimentwise error rate at a nominal 0.05 level (it does so for one of the 14 sets only). Of these 70 contrasts, 14 compared the subliminal and the supraliminal conditions. Five significant effects were found there. The remaining 56 contrasts involved only the subliminal conditions. Two significant effects and a trend (at the .01 and the .05 level, respectively) resulted. Considering this, it is of some concern that the significance of none of the effects encountered within the subliminal conditions is striking, in fact, none of them reaches significance at a next lower level (e.g. at the 0.01 level for effects tested at the 0.05 level). Since planned orthogonal comparisons reflect a powerful procedure, we are clearly dealing with weak effects. To exclude the possibility that they are the result of random processes replication of the present findings seems essential. The possibility of methodological artifacts must also be considered. Whether the threshold obtained was too liberal is a point to be discussed, since the use of the sensitivity parameter d' and the separate assessment of detection performance for words and nonwords yielded a distinct pattern of detection accuracies, different from what would be expected by chance.
However, there is some evidence indicating that the results obtained are at least internally consistent. First, it is felt that they lend themselves to a consistent interpretation, for which the ERP measures and the detection performance measures provide converging evidence. Secondly, while the results are arguably weak, they are at least specific, in that they do not seem to generalize from the set of a priori hypotheses chosen to a number of other possible hypotheses. Each contrast was hypothesized to yield maximal discrimination at specific recording sites. To assess the value of these specific hypotheses, each contrast on the peak-to-trough amplitudes was performed for the remaining sites. This set of contrasts, including over twice the number of the specific contrasts performed, yielded two trends ($p < .05$) as opposed to two significances ($p < .01$) and one trend for the specific hypotheses. While these arguments clearly do not question the need for a replication of the present findings, neither do they favor the view that the results should be dismissed as being due to chance.

4.2 Detection Performance

While the overall results indicate that subjects judged presence or absence of a string no better than chance, the detailed analysis of the d's yielded different pictures for words and nonwords. It seems likely that in the preexperimental session the criterion (two blocks in sequence with less than 60% correct responses) artificially depressed the estimate of the Ss detection performance below the actual performance level, i.e.
that it introduced a sampling bias towards the blocks of trials with poor overall performance. It seems that Ss were actually slightly above chance for detection of words, as reflected in the (unbiased) second session, while for nonwords their sensitivity was at chance level. It is assumed that this relation held for both sessions; in the first one, however, the bias forced actual random performance (for nonwords) below chance level, and forced the actual above chance performance (for words) to a level no more different from chance.

An obvious question then is whether 'subliminal' or 'unconscious' perception is an adequate term to describe the word trials in the time estimation task. Several findings bear on this issue. First, maximum sensitivity (for words in the postexperimental session) was still very low, yielding a $d'$ of 0.23 and 55% correct responses. Of all Ss (other than the excluded ones) questioned, only one reported with some confidence having seen letters or words, after the first detection session. An additional 48 detection trials were run for that S, and he was to report all he could from trials in which he felt he had actually seen something. The S responded to these instructions on 5 trials. He reported letters or words on three 'string' and on two 'blank' presentations. For the string presentations, he reported 'tie' as 'hat' (both words were among the 12 words in the detection deck), 'AHT' as starting with 'S', and 'PEOECPF' as a nonword starting with 'MB'. For one of the blanks he reported 'ROOF', for the other blank 'MBL' in the middle. These data are clearly at odds with his
confident report of having seen letters and words. All but one subject debriefed after the experiment were clearly amazed when told that strings had been presented on 80% of the trials, although the trials looked identical in detection and time estimation trials.

The response latency data make a similar point: There was no sign that 'subliminal' presentation of targets or nontargets led to responses similar to those obtained under normal viewing conditions, nor was there any evidence for an induced response conflict.

Furthermore, the results indicate that the term to question is 'detection' rather than 'subliminal' or 'unconscious' perception. Overall, Ss proved unable to judge, at better than chance accuracy, whether any stimulus had preceded the pattern mask. This is how detection was originally defined in the experiment. While the finding that presence/absence of some strings (i.e. words) was judged with slightly better than chance accuracy is in need of an explanation, it does not reflect what is commonly understood as 'detection'. Similarly, the early ERP components associated with detection and stimulus set (Parasuraman et al., 1982) discriminated between subliminal and supraliminal conditions but not between blanks and 'subliminal' strings. Parasuraman et al. (1982) reported that N100 reflected detection performance. The results obtained here for the components in the same latency range indicate that detection accuracy, where above chance, was not accompanied by the expected early ERP effects. This might again indicate that
detection, when confounded with the content of the stimulus, cannot be equated with 'simple' detection. To assess the possible contributions of 'detection' performance to the late ERP effects of familiarity and presence of a string, correlations between the critical ERP differences and the sensitivity to words (in the second, unconstrained session) were calculated. The correlations were not significant, $r(10) = -0.25$ for familiarity and $r(10) = -0.26$ for presence ($p > 0.10$ in each case). Together with the negative values of the correlations, this result most probably precludes an interpretation of the ERP effects in terms of residual awareness of some aspect of the stimuli.

Finally it should be noted that the sensitivity measures were obtained under optimal conditions. Great care was taken to assure that Ss had a clear idea of the stimuli they were confronted with, and that reliance on intuitive feelings and vague impressions rather than confidence was emphasized. In addition, considerable amounts of practice and feedback were given.

To summarize, several lines of evidence favor the conclusion that the overall level of performance, the lack of confidence accompanying the judgements, and the peculiar pattern of accuracy do not support the view that words were 'detected'. However, the results also indicate that discriminating (rather than detecting) verbal behavior did occur. A complete dissociation between verbal report and indirect measures of stimulus discrimination, which is the focus of a rather hot debate around the issue of subliminal perception (e.g. Merikle,
1982; Purcell et al., Note 1), was not found in this study.

The differential sensitivity to words and nonwords is interpreted in line with recent results reported by Avant & Woods (Note 3). They found that lexical judgements on two successively presented, pattern-masked letter strings (i.e. to judge the strings as 'same' or 'different' with respect to lexical status) were made with better than chance accuracy, while for the same stimuli figural judgements (i.e. 'same' or 'different' with respect to letter case) were at chance.

In the present study Ss performed as if making a lexical decision; i.e., they discriminated words, but not nonwords, from blanks. Consistent with this is the finding (for the unconstrained postexperimental session) that only 42% hits were scored for nonwords. This yields 58% 'blank' responses to nonwords, a figure comparing favorably with the 57% 'blank' responses to actual blanks. Subjects might actually have responded with 'blank' in the sense of 'no word', thus successfully performing a lexical decision task.

Unfortunately, reinterpreting the 'blank' response to a nonword as a hit in a hypothetical lexical decision task confounds sensitivity and response bias: thus the figures might simply indicate a general tendency to respond 'blank' along with an increased sensitivity to words.

Whatever task Ss actually performed it seems that they monitored internal evidence confounded with lexical processing. This extends the results of Avant & Woods (Note 3) to the domain of 'simple' detection. Detection under conditions of pattern
47

masking, i.e. when figural properties of the target are interfered with, thus might rely on information at a higher level than previously assumed. In this respect, it is an interesting possibility that the instructions given to optimize performance, with their emphasis on 'simple' physical features, were perhaps detrimental rather than useful - detecting 'simple' physical properties might be different from, and more demanding than, detecting some internal events confounding physical and lexical properties of stimuli.

4.3 **Time Estimates**

The negative findings in terms of response latencies are difficult to discuss. Unfortunately, the experiment does not tell us whether the time estimation procedure chosen would be sensitive to similar stimulus manipulation under normal viewing conditions. The finding of Avant and Woods (Note 3), that relative durations were affected by lexical status of pattern masked material, presented in the compared intervals, might be limited to relative duration judgements. However, there are no positive results to support such speculations at this point.

4.4 **Event Related Potentials**

The difference in baseline values between the subliminal and the supraliminal conditions is puzzling. There are several possibilities that could produce such an effect. The possibility that the effect was artifactual was considered. Since prestimulus negativity does correlate with performance in some situations (Donchin et al., 1978), it is possible that the trials excluded as invalid in the subliminal and in the
supraliminal conditions (as reflecting poor performance) were characterised by different anticipatory negativity, since different types of preparation might be useful or detrimental in different types of tasks. The data were therefore digitized again, with all criteria for trial exclusion (except for experimenter errors) removed. The contrast between subliminal and supraliminal conditions performed on this new set of baselines still yielded $F(1,11) = 8.82$ with $p = .013$. Thus baseline differences were present in the 'raw' averages obtained from all trials. Since there is no evidence that the effect is spurious, it seems very likely that Ss had some sort of advance information as to what type of trial would occur. The possibility that the experimenter communicated to the S nonverbal cues based on his bias for a specific outcome seems unlikely, partly because the experimenter was fully occupied with changing cards, recording the codes, and monitoring the eye movements of the S, and partly because the main interest of the study was in differences within the subliminal conditions, where no baseline differences were observed. However, the possibility cannot be completely refuted. Wastell and Kleinman (1980) suggested that in experiments with random permutations of stimuli, predictability of a stimulus type is still above chance, due to the constraint of equal (or just fixed) numbers of stimuli per type. The additional constraint, that no more than three supraliminal stimuli and no more than two supraliminal stimuli of one type were allowed in sequence might have increased predictability to the point where it resulted in
differential preparation. It is argued here that the validity of the main conclusions is not affected by this finding of baseline differences, be it spurious or not. If the differences reflect availability of advance information, then the lack of difference within the more interesting subliminal conditions would seem to imply that within these conditions no cues were available. Nevertheless, care must be taken interpreting differences between subliminal and supraliminal conditions.

The difference in N1-P1, starting before approximately 140 msec poststimulus, most probably reflects early effects of selective attention. No difference was obtained for N1, which is tentively assumed to be homologous with N100. Wastell and Kleinman (1980) similarly found no effect of selective attention on N1 in an experiment where targets and nontargets differed in simple physical cues, such as shape and intensity, but appeared at the same location. They argue that N1 reflects mainly attention to a single spatial source. It is unclear to what extent the two possible spatial positions in the present experiment can be considered as such a single source. In the present experiment, N1 occurred somewhat earlier than is usual for the visual modality. However, the finding that differences still do occur before 140 msec poststimulus is generally in agreement with other data on early selection (Picton et al., 1978). It is also possible that an even earlier effect is obscured by the different baselines, e.g. the high prestimulus negativity for the supraliminal conditions might have led to a ceiling effect for the early N1 negativity. The other
differences will not be discussed in detail, since differences in stimuli, task and baseline cannot be separated. It should be noted, however, that the differences are major and obvious on simple visual inspection.

The differences within the subliminal conditions were small and not obvious on simple visual inspection. Components of the ERP discriminated between familiar and unfamiliar words at 250-300 msec, while presence vs. absence of a string was discriminated by later components, at 350-400 msec. This result is consistent with the view that detection under conditions of backward-masking is a late process which draws on evidence from specialized lexical (and maybe other) processors, an interpretation supported by the trend for words vs. nonwords: discrimination of lexical status also occurred after the effect of familiarity, at a latency similar to the detection effect. However, it also seems possible that the effect of lexical status is spurious.

4.5 General Discussion

The results provide converging evidence from sensitivity and ERP measures that detection under the present conditions is a late process. They also argue against the view that pattern-masking disrupts processing: ERP-effects did not occur before 260 msec poststimulus (latency of P2), and were found prominently in comparisons involving even later components (N3 at 300 msec, P3 at 400 msec). A reinterpretation of the effect of familiarity in terms of task relevance is not supported because of the negative finding for the target vs. nontarget
comparisons, and it seems as if the words were identified by the perceptual system, but not recognized as being tied to different interpretations within the task.

The present study indicates that specific ERP correlates of stimulus processing can be obtained under conditions of pattern masking. However, little can be concluded about the correlates for the simple stimuli commonly used in backward masking paradigms (e.g. Andreassi et al., 1976; Schwartz et al., 1981). The effects reported here are content specific, potentially indicating the usefulness of complex stimuli in the study of the ERP under backward masking.

An interesting finding in the present study was the apparent convergence between ERP effects and interpretations derived from behavioral measures. That familiarity is coded earlier than presence/absence of a letter string has been interpreted as reflecting the dependence of 'detection' on lexical information under conditions of pattern masking. Even more intriguing is the fact that the trend for discrimination between words and nonwords in the ERP appears temporally confounded (i.e. occurs in the same latency range of 300-400 msec) with the discrimination between strings and blanks. This finding is paralleled in the behavioral domain - detection performance was confounded with content of the presented stimulus. Taken together, the findings extend the usefulness of the ERP in the study of human information processing to situations where automatic (in the sense of unconscious) processing prevails.
Familiarity effects may occur automatically, i.e. without the S being aware of differential familiarity (Scarborough et al., 1979). Their locus seems to be lexical rather than semantic (Warren & Morton, 1982) since they do not show transfer across modalities. Most recently, Tulving, Stark & Schacter (1982) have succeeded in separating effects of familiarity on priming and recognition. In their study words showed effects of priming (in a fragment completion task) from previous exposure without being recognized; and the effects were found to have different decay functions. Priming thus seems to be unrelated to episodic memory and might be indicative of a lower level memory system. This seems to parallel the present study in which the ERP indicates that items were identified as familiar, but not recognized as being of different task relevance.

That familiarity is a powerful variable with long retention, and that it is readily affected by experimental manipulations has repeatedly been shown (e.g. Murrell and Morton, 1974; Scarborough et al., 1979). However, the main interest in word recognition has focussed on semantic effects over short intervals (e.g. semantic priming). Research explicitly concerned with the effects of verbal information over longer periods, as well as the findings of the present study, suggest a somewhat different emphasis: The view of word recognition as the temporary activation of a complex structure might be slightly misleading, since it neglects the important role of input in updating perceptual and cognitive structures with respect to future events.
The failure to obtain differences between the subliminal targets and nontargets might indicate that subliminal stimuli do not interact with task relevance, here conceived of as semantic set. However it may well be that for semantic effects to become apparent, some response to a meaningful stimulus (not necessarily the subliminal one, though) is required. In cases where subliminal stimuli prime a conscious target (Marcel, in press) this requirement is met. It is not met, however, in the present study where effects of (conscious) semantic set on subliminal targets were examined. Possibly a conscious target is needed to 'carry' semantic effects through the cognitive system up to a point where they affect ongoing processes. More specifically, recovery processes acting on a lexical record may not occur automatically to pattern masked stimuli, but when initiated by a normally presented target they may act on any lexical record available, including those from sources that will never become conscious. However, more sensitive experiments must be designed to answer this question.

How do the findings fit with the models of recognition and, more specifically, word recognition, as introduced at the beginning? The result that detection accuracy is affected by lexical status supports the view adopted by signal detection theorists (e.g. Green and Birdsdall, 1978). Detection and recognition thus may reflect a unitary process, with the same set of internal detectors being monitored in both situations. The finding might also imply that the signal detection model applies not only to simple stimuli. However the results do not
seem to favor a single model of word recognition. The strength of the logogen model of Morton (1969) with respect to the present findings is that it explicitly accounts for familiarity effects and the updating of the logogen units with respect to future events. The interactive activation model (Rumelhart & McClelland, 1982) is helpful in explaining the selective effects of backward masking on figural properties of a word, and it might give a slightly more direct account for the higher detection accuracy for words than for nonwords. In the framework provided by the logogen model (Morton, 1969), this result could indicate that Ss monitored the lexical system for the detection task. It seems unlikely, however, that single logogen units exceeded threshold, i.e. that words were actually recognized; rather Ss must have had some information about activation below the threshold level. In the framework provided by the interactive activation model, evidence available at the lexical or word level influences lower levels of representation via feedback. Thus in this model judgements based on evidence from lower levels are expected to show influences of higher level information such as lexical status. However, as mentioned earlier, the present experiment can not decide at which level Ss actually monitored evidence for the detection task. Monitoring either the lexical level or a figural level receiving feedback from the lexical level could both yield the pattern of results actually obtained.

A last general point addresses the issue of the proper threshold for subliminal perception. Chances are that the
outcome of the debate whether or not results like the ones obtained by Marcel (in press) are demonstrations of unconscious perception will be far less interesting than the results themselves, unless a broader view of the 'proper' threshold is adopted.

Merikle (1982) has argued for the need to use signal detection parameters along with a sufficient number of observations for determining a detection threshold, and he obviously suspects the thresholds in the papers by Marcel (in press), Fowler et al. (1981) and McCauley et al. (1980), to be too liberal to warrant an interpretation in terms of unconscious perception. Purcell et al. (Note 1) have pointed out that conditions for the threshold determination and for the actual experiment, while differing in seemingly trivial details, may result in drastically different estimates of the identification or detection accuracy. They demonstrated that adaptation differences could be responsible for the results obtained by McCauley et al. (1980).

It seems indeed unfortunate that the methods used to determine detection thresholds leave so much space for differing interpretations of the level of performance and experience they reflect. In the experiment described here an attempt was made to avoid some of the more obvious criticisms, e.g. by giving rather explicit instructions, by familiarizing Ss with the stimuli and the task, and by assessing detection performance in some detail with signal detection parameters. However, the results indicate that cognitive manipulations affect detection
to such an extent that threshold determination should probably control for them, too. This might actually mean that two experiments with identical setup (including all manipulations of interest) but different tasks should be run, one to assess detection performance and one as the actual experiment. The suggestion obviously is simplistic and limits the range of possible experiments. It may, however, serve the purpose of illustrating what could at least partly underly the much discussed 'complete dissociation' between verbal report and indirect measures of behavior: differences in instructions and tasks.

A more viable alternative is the procedure used by Avant & Woods (Note 3) who compared figural and lexical, semantic or temporal judgements to pairs of stimuli. This approach examines what information is available for a considerable range of tasks. An important feature within this approach is the use of figural forced choice judgements, i.e. a case discrimination. While this approach might arguably address rather abstract figural qualities, it has the advantage of being less confounded with lexical processing than 'simple' detection. Alternatively, the use of concurrent detection and recognition judgements (e.g. Parasuraman et al., 1982) might serve better to characterize detection under difficult conditions. Their additional use of confidence ratings results in a more detailed picture of the relationship between performance and experience in complex detection tasks than simple presence/absence judgements.

Finally, a more radical approach would be to abandon the
concept of detection as a basic, simple process. This approach has some appeal because the preoccupation with 'simple' detection might be founded on a misconception, namely that an early, content-independent process close to the sensory end mediates detection and thus must be accessible for the purpose of verbal report about the stimulus. The argument against subliminal perception then seems to be that if such an early, simple function is precluded, later complex functions are even less likely to occur. The present study, however, indicates that detection under conditions of pattern-masking might be a rather late, content-specific process. Thus, occurrence of an (incomplete) dissociation between detection accuracy and indirect measures of semantic processing may not indicate a counterintuitive dissociation between early, simple and complex, late processes, but between equally late and complex processes of different specialization. If the latter interpretation is correct, then detection of complex and meaningful stimuli under difficult conditions might be interesting in itself and thus be worth further investigation. Allport (1980, p.51) in discussing 'New Directions' for cognitive psychology states that

"... A different and contrasting orientation focusses directly on information content, and on content-specific (hence distributed) mechanisms for its representation and use..."

This contemporary characterisation of cognitive research seems to leave much more room for results of the kind obtained in this exploratory study than does the notion of 'simple' detection.
BIBLIOGRAPHY

i. REFERENCE NOTES


ii. REFERENCES


Gleitman, H. & Jonides, J. The effect of set on categorization


Mitzel, H.C. & Games, P.A. Circularity and multiple comparisons in repeated measure designs. *British Journal*


129-140.
APPENDIX A - LIST OF STIMULI

i. Stimuli for the detection task

<table>
<thead>
<tr>
<th>WORDS</th>
<th>NONWORDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIG ....72</td>
<td>TIE ....23</td>
</tr>
<tr>
<td>HAT ....56</td>
<td>WORD ....274</td>
</tr>
<tr>
<td>BALL .....110</td>
<td>WINE ....72</td>
</tr>
<tr>
<td>TENT .....20</td>
<td>RIFLE ....63</td>
</tr>
<tr>
<td>BIBLE .....59</td>
<td>MOUSE .....10</td>
</tr>
<tr>
<td>COFFEE ....78</td>
<td>TENNIS ....15</td>
</tr>
</tbody>
</table>

ii. Stimuli for the experimental tasks

<table>
<thead>
<tr>
<th>WORDS</th>
<th>NONWORDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAR ....274</td>
<td>BED ....127</td>
</tr>
<tr>
<td>BUS .....34</td>
<td>RUG .....13</td>
</tr>
<tr>
<td>TAXI .....16</td>
<td>DESK ....65</td>
</tr>
<tr>
<td>JEEP .....16</td>
<td>LAMP ....18</td>
</tr>
<tr>
<td>BOAT .....72</td>
<td>TABLE ..198</td>
</tr>
<tr>
<td>TRAIN ...82</td>
<td>CHAIR ...66</td>
</tr>
<tr>
<td>TRUCK ...57</td>
<td>BENCH ...35</td>
</tr>
<tr>
<td>TRACTOR .24</td>
<td>MIRROR ..27</td>
</tr>
</tbody>
</table>

Note - numbers reflect frequency counts from Kucera & Francis (1967). In a few cases, counts given do not reflect the numerosity of a given item (e.g. count given for 'ear' reflects actual count for 'ears').
APPENDIX B - GROUP MEAN ERPS AT P4 FOR ALL CONDITIONS

Note - the upper two tracings depict the supraliminal, and the lower five tracings the subliminal conditions, respectively. Epoch 500 msec poststimulus; negativity upwards; dashed lines (baselines) 10 microvolts apart.
APPENDIX C - GROUP MEAN ERPS AT OZ FOR ALL CONDITIONS

Note - the upper two tracings depict the supraliminal, and the lower five tracings the subliminal conditions, respectively. Epoch 500 msec poststimulus; negativity upwards; dashed lines (baselines) 10 microvolts apart.
APPENDIX D - GROUP MEAN EOG FOR ALL CONDITIONS

Note - the upper two tracings depict the supraliminal, and the lower five tracings the subliminal conditions, respectively. Epoch 500 msec poststimulus; negativity upwards; dashed lines (baselines) 10 microvolts apart.
APPENDIX E - AVERAGED ERPS FOR ONE SUBJECT (S1) AT P3

Note - the upper two tracings depict the supraliminal, and the lower five tracings the subliminal conditions, respectively. Epoch 500 msec poststimulus; negativity upwards; dashed lines (baselines) 10 microvolts apart.
APPENDIX F - AVERAGED ERPS FOR ONE SUBJECT (S2) AT P3

Note - the upper two tracings depict the supraliminal, and the lower five tracings the subliminal conditions, respectively. Epoch 500 msec poststimulus; negativity upwards; dashed lines (baselines) 10 microvolts apart.