Phoneme Identification and the Lexicon

ANNE CUTLER

MRC Applied Psychology Unit, Cambridge, United Kingdom

JACQUES MEHLER

Centre de Science Cognitive et Psycholinguistique, CNRS, EHESS, Paris, France

DENNIS NORRIS

MRC Applied Psychology Unit, Cambridge, United Kingdom

AND

JUAN SEGUI

Laboratoire de psychologie experimental, associe au CNRS, Paris, France

In seven experiments reaction time to detect the initial phoneme of words and nonwords was measured. Reaction time advantages for words over nonwords come and go according to the particular characteristics of the experimental situation. One relevant characteristic is degree of task monotony, an effect which is most parsimoniously explained by attention shifting between levels of processing. General classes of models of the relationship between levels of processing in comprehension are discussed in light of the results. Serial models incorporate an attention shift explanation of the monotony effect more elegantly than do interactive models. Alternative serial models are available in the literature in this area. One recent model, which allows only a single outlet point for phoneme detection

responses, and hence requires that apparent reaction time advantages for words are artefactual, can be unambiguously rejected on the basis of the present data. It is argued that a serial model involving competition between target detection based on a prelexical representation and detection based on a lexical representation most satisfactorily accounts for the overall pattern of results. © 1987 Academic Press, Inc.

INTRODUCTION

Language comprehension involves a number of distinguishable levels of processing. The current debate in the study of comprehension centers upon whether these levels may or may not interact. On the one hand there are models of comprehension which allow information flow between processing levels; a crucial aspect of such models is that "top-down" information flow may occur, i.e., decisions at levels close to the input level may be influenced by processing taking place at "higher" levels of comprehension, i.e., levels further removed from the input level. Models belonging to this general class include the interactive activation model of comprehension (Elman & McClelland, 1984; McClelland & Elman, 1986) and the cohort model of auditory word recognition (Marslen-Wilson, 1980). On the other hand are comprehension models which do not allow interaction; in such models, each level of processing is autonomous, and the flow of information is strictly serial and bottom-up, with decisions at "lower" levels being determined only by information coming up from the input. To this class belong inter alia the model of comprehension propounded by Fodor, Bever, and Garrett (1974) and the checking model of word recognition (Norris, 1986).

This debate has prompted experiments which typically examine the relationship between two levels of comprehension by measuring performance on a task which is claimed to reflect processing at the lower level, in the light of experimental manipulations at the higher level. The two levels at issue in this paper, the phonemic processing level and the lexical identification level, provide a case in point. Can lexical information influence phonemic processing? Interactive models claim that it can; serial autonomous models hold that it cannot.

The processes underlying language comprehension cannot, alas, be directly observed. Considerable psycholinguistic ingenuity has therefore been expended upon devising tasks which can be expected to reflect the characteristics of processing at different levels. Response latency is the most common dependent variable—"by default: there simply isn't much else that can be measured" (Pachella, 1974).

Phoneme monitoring is just such a tool for examining comprehension in progress, and it is designed to reflect processing at the phonemic level. Subjects in this task—originally developed by Foss (1969)—listen to

speech and press a button as soon as they hear a word beginning with a specified target sound. For instance, the target might be specified as "/d/ as in dog," in which case the subject would be expected to respond on hearing the word "deep," say, or "doorway," or "drunkenness," or any other word beginning with /d/. The task cannot therefore be performed by storing an acoustic template of the target and searching for a simple acoustic match in the input, since a given phoneme is represented by different acoustic patterns in different phonetic contexts; the subject has to search for a token of the phoneme /d/. Thus response time in this task reflects difficulty of processing at the phonemic level. Of course, whether or not the normal comprehension process includes an explicit phonemic level of representation is an open question; the important point for the present discussion is that if there is a phonemic level of processing in comprehension, it is closer to the input than the lexical level is. This is because there is little reason for a phonemic level of processing to be postulated subsequent to lexical processing, since once the comprehension process has progressed to a representation of meaning there is no necessity to return to a representation of sound. Some researchers have argued that speech is indeed classified prelexically into phonemes (e.g., McNeill & Lindig, 1973; Foss & Gernsbacher, 1983), while others have argued that—in the perception of some languages, at least—a prelexical classification into syllables is undertaken (e.g., Mehler, Dommergues, Frauenfelder, & Segui, 1981; Cutler, Mehler, Norris, & Segui, 1983, 1986). However, because the phoneme-monitoring task actually forces a phonemic representation, it is well suited for examining the relationship between phonemic and lexical processing via the investigation of lexical effects on phoneme detection latency. Indeed, there is a substantial body of research which has directly or indirectly addressed this issue.

Morton and Long (1976) found that phoneme targets on highly predictable words were responded to faster than targets on less predictable words. They argued that phoneme-monitoring response output followed identification of the word bearing the phoneme target. This had also been claimed by Foss and Swinney (1973) on the basis of the finding that response time to a word target was faster than response time to a phoneme target on that same word. Similarly, Rubin, Turvey, and Van Gelder (1976) found that phoneme-monitoring responses were faster when the target began a word rather than a nonword—further indication that lexical access of the target-bearing word preceded phoneme identification. Segui and Frauenfelder (1986) found that targets beginning high-frequency words were detected faster than targets beginning low-frequency words, if subjects were monitoring for targets occurring anywhere in a word, not just word initially. Cutler and Fodor (1979) found that response time was shorter if the word beginning with the target was focused by

context from a preceding sentence. Dell and Newman (1980) replicated Morton and Long's finding and found that the predictability effect interacted with effects of phonetic distraction by phonemes similar to the target. An experiment by Mehler and Segui (reported by Mehler, 1986) also demonstrated such contextual predictability effects.

On the other hand, Foss, Harwood, and Blank (1980) found no evidence of lexical factors such as word-nonword status or frequency of occurrence of the target-bearing word affecting response time. Segui and Frauenfelder (1986) also found no word frequency effects on target detection when they used standard phoneme-monitoring instructions (monitor word-initial phonemes only). Earlier, McNeill and Lindig (1973) and Healy and Cutting (1976) had shown that a match between target level and response item level (e.g., monitoring for a phoneme when the response item was itself a phoneme) facilitated response latency, and had argued from this that the phoneme-monitoring task requires listeners to focus on a phonemic representational level rather than a lexical representation. Further evidence for the match interpretation was adduced by Mills (1980a, 1980b). Newman and Dell (1978) found that low-level phonetic factors could affect monitoring response time; in particular, false alarms were generated by phonemes which were only one phonological feature different from the target phoneme. Newman and Dell argued that a phonological representation constructed subsequent to word identification should allow a simple yes-no response as to whether or not the input matched the target; there should be no gradient of similarity. Since a gradient of similarity defined in terms of phonological features was definitely affecting response time in their experiment, Newman and Dell argued that the phoneme-monitoring response can be executed on the basis of a phonological representation computed *prior* to lexical identification.

As this summary reveals, the question at issue in this research was largely seen to be whether phoneme-monitoring responses can be executed prior to, or must be executed subsequent to, lexical access. Lexical effects on phoneme identification, such as those found by Morton and Long (1976) and by Rubin et al. (1976) were held to argue for the latter position. Nevertheless, there was also strong evidence—from the Foss et al. (1980) study, and that of Newman and Dell (1978)—in favor of prelexical phoneme identification. It is perhaps not surprising that these apparently contradictory results inspired an explanation in terms of an interactive model. Stemberger, Elman, and Haden (1985) interpreted the phonemic/lexical relationship in speech comprehension, as evidenced by the phoneme-monitoring data, in terms of Elman and McClelland's (1984) interactive activation model of comprehension. According to this explanation—which, Stemberger et al. claim, "accounts well for all presently known data" (p. 482)—comprehension does involve an explicit phonemic

level of representation. This level "mediates between" an acoustic feature level and the lexical level. Activation at any one level feeds both to higher and to lower levels, so that lexical effects result from activation at the lexical level feeding back to the phonemic level. This level is the only point at which phoneme detection responses can be made, so that the model can account equally well for the absence as for the presence of lexical effects, by assuming that activation from lower levels can under appropriate circumstances be more rapidly effective than activation from higher levels.

McClelland and Elman (1986) also specifically consider phoneme-monitoring results in the light of their TRACE model; their account is essentially identical to that of Stemberger et al. (1985). However, they do suggest (p. 30) that when a phoneme is unambiguously articulated, and in word-initial position, a monitoring response should always be possible on the basis of activation of the phoneme detector independently of top-down activation from the word level. That is, they suggest that the findings of Foss et al. (1980), and of Newman and Dell (1978), are more in line with the predictions of their model than the lexical effects found by Morton and Long (1976) and by Rubin et al. (1976).

In summary, though, any interactive activation account of the phoneme identification process is one in which there is only a *single outlet point* for the phoneme detection response, and in which the phonemic and lexical levels of representation can *interact.*

The same body of phoneme-monitoring results, however, has also been accounted for within a model in which phonemic and lexical processing are *serial* and *autonomous.* Two versions of this account exist in the literature; they have in common that comprehension is seen as a series of mandatory autonomous processes, with information flow only in a bottom-up direction. In order to account for the comings and goings of lexical effects on phoneme-monitoring latency, as summarized in the brief review above, these accounts postulate *more than one outlet point* for the phoneme detection response.

The first such account was produced by Cutler and Norris (1979), who proposed that the monitoring task can be effectively performed either as a result of a target detection procedure carried out on a prelexical representation or on the basis of phoneme information associated with a lexical representation. The monitoring task triggers a target detection procedure which operates on the earliest accessible representation of the input. This procedure goes on in parallel with the normal mandatory process of locating a lexical entry from a prelexical representation. The two procedures race. If the target detection procedure produces an output prior to the completion of lexical access, the response will be based solely on prelexical information; if lexical access is achieved before the target de-

tection procedure produces an output, the response will be based on lexical information.

Two important characteristics of this account should be noted: (i) the target detection procedure is *not* a part of the normal comprehension process, but is a task-specific operation; (ii) the prelexical representation on which the target detection procedure operates may, but is not logically required to, itself contain phonemic information. That is, although target detection requires a phonemic representation, this may have to be computed, rather than simply retrieved, from the normal prelexical representations. It is assumed that phonemic information is represented in some way in lexical entries, but it is not necessarily assumed that it is required for a prelexical access code.

Cutler and Norris outlined the conditions under which each of the two alternative detection procedures should win the race, (i) All conditions being equal, the prelexical procedure wins. (ii) If word identification is speeded by contextual focus, or by high predictability of the word in context, then the lexical procedure wins. (iii) If lexical access is completed very quickly simply because the target-bearing word is very short, the lexical procedure wins.

A further characteristic of this account is that lexical factors can only be facilitative; if word identification is slowed for some reason this should have no effect whatsoever on the target detection procedure.

A very similar account was proposed by Foss and Blank (1980), whose formulation provided for the monitoring response to be effected on the basis of either of two codes, which are both normally computed as a part of auditory comprehension. The first, a prelexical code, they called phonetic. Its important characteristic is that it may be incomplete, since acoustic information in the signal may be too distorted or too densely coded to allow for a complete phonetic transcription. The second code is postlexical, and called the phonological code. The phonological representation of the word in the lexicon is necessarily a complete one.

An important feature of this "dual-code model" is that it claims that both the phonetic and the phonological representations are normally computed during language comprehension. The target detection process is a simple matching procedure which compares an internal representation of the target with the segments present in the internal representation of the speech input. It is essentially the same procedure irrespective of what level of speech representation it is operating on. Foss and Blank remain neutral as to whether the target specification can give rise to separate phonetic and phonological internal representations, or whether it produces a single representation so that one or other of the representations of the speech input will require conversion.

Foss and Blank hypothesized that task difficulty would be one of the

primary factors affecting which code forms the basis of target detection. If task difficulty is high (if, for instance, there are two simultaneous targets, or phonemes very similar to the targets are present in the input), then the likelihood of the phonetic code decaying before it can be fully analyzed is heightened, and responses are more likely to be based on the postlexical phonological code. Similarly, contextual predictability is likely to speed lexical access and hence make postlexical target detection more likely. Finally, assigning a higher status to the task of comprehending the input as opposed to the task of detecting the target should lead to processing resources being diverted from target detection, with the consequence that once again this is unlikely to be completed before the phonetic representation has decayed, so that the postlexical phonological code will be used instead.

Dell and Newman (1980) adopted the Foss and Blank (1980) model to account for their finding of apparent interaction between phonetic and lexical effects, although in an earlier publication (Newman & Dell, 1978) they had preferred an interactive model.

In some minor respects these two accounts, despite their essential similarity, appear to make slightly different predictions. Cutler and Norris' (1979) "race model" specifically claims that lexical information can only be facilitative, whereas the dual code model allows for situations in which the prelexical phonetic code has disappeared, so that a response based on it is impossible; under such conditions lexical information will be free to exercise facilitative or inhibitory effects. The dual code model claims that a prelexical phonetic code is normally computed during comprehension, whereas the race model is neutral with respect to the nature of the prelexical representation upon which target detection can be performed. The race model claims that the race occurs for every detection response, so that in a given experiment the result of the race is specific for every individual response, whereas the dual-code model at least suggests that experimental situations can be devised in which one or other route must be employed throughout the experiment. In this respect the Cutler and Norris account appears to propose *multiple outlet points* for the phoneme detection response in every comprehension situation, whereas the Foss and Blank account proposes *alternative outlet points* according to the experimental characteristics of the particular situation. Nevertheless, the similarities between the two accounts are far more striking than these differences: both are models in which the phonemic and lexical levels of representation are serially ordered and noninteractive, and both allow more than one outlet point for the monitoring response. Recently, however, some of the results upon which these accounts were founded have been disputed.

Foss and Gernsbacher (1983) attempted to test predictions of the dual

code model, in particular the prediction that task difficulty would in-
crease the likelihood of postlexical responding. They found no support
for the model. Instead, in several experiments, they found evidence that
monitoring responses were strongly influenced by phonetic factors. The
major phonetic factor which they identified was the nature of the vowel
following the consonant target—reaction times to [d], for instance, were
slow when the target was followed by the vowel *[ae],* as in *dash,* or the
nonword *dat,* fast when the vowel was [e], as in *debt* or *desh.* Foss and
Gernsbacher referred to this effect as one of vowel length, because they
found that mean response times in one of their experiments correlated
with a measure of vowel length. But the measure of vowel length which
they used was that computed by Peterson and Lehiste (1976) for a
northern United States' dialect in 1960. Twenty-five years later, Foss and
Gernsbacher's (1983) subjects and materials used a southern United
States' dialect. Vowel length is one of the major differences between
northern and southern United States' dialects. It is thus at least question-
able whether Peterson and Lehiste's measure is valid for Foss and
Gernsbacher's experiment. Foss and Gernsbacher did not measure the
actual length of their materials. Therefore, while not disputing that the
ordering of their RTs was affected by the nature of the vowel, we would
prefer to refer to this effect by the more neutral term of vowel identity.

Foss and Gernsbacher found no effect on response times, under any
conditions, of lexical status of the target-bearing word. Moreover, they
found that the target-bearing items from Morton and Long's (1976) study
of contextual predictability produced the contextual predictability effect
even out of context. That is, those target items which had been predict-
able in context (e.g., *door* in "Slowly he opened the door") produced
faster response times in isolation than did those items which had been
less predictable in context (e.g., *dance* in "Slowly he opened the
dance"). Foss and Gernsbacher concluded that Morton and Long (and
other previous researchers who had shown apparently postlexical effects
in phoneme monitoring) had inadvertently confounded their independent
variables with the vowel identity factor.

Foss and Gernsbacher therefore rejected the dual-code model and pro-
posed instead that phoneme-monitoring responses are always based on a
prelexical representation of the input. (It should be noted that Foss and
Gernsbacher's account, in contrast to the dual-code model but in agree-
ment with the race model, is neutral with respect to the exact nature of
this prelexical representation and in particular with respect to whether it
contains segments which are isomorphous with phonemes.) In compar-
ison to the earlier models, then, Foss and Gernsbacher's account can be
seen as a *serial autonomous* model of comprehension in which there is
only *a single outlet point* for phoneme detection responses. Thus there

are essentially three contenders for an account of the work described above: interactive processing levels with a single response outlet; serial autonomous processing levels with more than one response outlet; and (based on a rejection of some of the above work) serial autonomous processing levels with a single response outlet. Thus the theoretical position is depressingly unresolved.

Moreover, the empirical situation is little better. Foss and Gernsbacher's conclusion that vowel identity is the strongest influence on phoneme-monitoring response time still does not account for all the available results. Although they have cast doubt on Morton and Long's experiment, other findings cannot be similarly dismissed. For instance, Mehler and Segui (reported in Mehler, 1986) demonstrated contextual predictability effects with target-bearing words in which vowel identity was fully controlled. In particular, the lexicality effect—faster response to targets on words than to targets on nonwords—still has some support. Although one study which controlled vowel identity across words and nonwords (Segui, Frauenfelder, & Mehler, 1981) failed to find an effect of word-nonword status, another study in which vowel identity was also controlled across words and nonwords (Rubin et al., 1976) found that words were responded to significantly faster than nonwords.

Inspection of the materials in these two studies suggested one further difference: the Rubin et al. (1976) study used monosyllabic target-bearing items, while Segui et al. (1981) used bisyllabic. Recall that at least one of the models discussed above, namely Cutler and Norris' (1979) race model, predicts greater lexical involvement in phoneme detection for shorter than for longer words. Foss and Gernsbacher apparently used both monosyllabic and bisyllabic target-bearing items in all but one of their experiments, therefore their results do not address this issue. Nor did they, unfortunately, conduct a study investigating lexicality effects with their vowel identity factor controlled.

The series of experiments reported in the present paper was designed in the hope of resolving both the empirical and theoretical uncertainties with respect to the relationship between phonemic and lexical processing in models of comprehension. All our experiments deal with the lexicality effect in the phoneme-monitoring task. It was hoped that if we could establish under precisely what conditions lexicality effects appeared and disappeared, we would not only clear up the empirical picture, but might also shed some light on the general characteristics of a model of comprehension with respect to these two processing levels. For reasons described immediately above, we decided to begin by investigating the lexicality effect using monosyllabic targets with, or course, vowel identity strictly controlled. The experiment was carried out, in French, in the same laboratory as the Segui et al. (1981) study.

EXPERIMENT 1[1]

Materials

Twenty-five word-nonword pairs were chosen. In each pair both word and nonword had the same vowel. Five pairs had CV structure, with a change in the initial C effecting the change from word to nonword (e.g., *pis-ti*). A further 10 pairs had CCV structure, with the same variation (e.g., *glas-clas*). A further 10 pairs had CVC structure, and in these pairs it was a change in the final C which distinguished word from nonword (e.g., *date-dac; belle-berre*). The 25 pairs are listed in the Appendix. Three hundred twenty-four further items were chosen, of which 60% were words and 40% were nonwords, and 75% had two syllables and 25% one syllable.

The total of 374 items were distributed into 100 sequences ranging in length from 1 to 6 items. Ten sequences, each 6 items in length, did not contain an occurrence of the specified target. A further 40 filler sequences had targets occurring in first, second, or sixth position. The remaining 50 sequences each had one of the 50 experimental items occurring in third, fourth, or fifth position. Each word was matched with its paired nonword on position in list. Each list terminated (after the occurrence of the target, or after the sixth item) with the word *fini*.

The lists were divided into two blocks of 50, with list length, filler type, and word-nonword targets counterbalanced across blocks. Each block was preceded by 10 practice lists similar in structure to the lists in the main blocks. All lists were recorded by a female native speaker of French. The lists were spoken at a rate of approximately one item per second.

Subjects

Subjects were 30 students of the University of Paris V. Half the subjects heard the two sets of lists in AB, half in BA order.

Procedure

Subjects were tested individually and were presented with a set of printed target specifications, each phoneme target being specified by a single letter. Subjects were instructed to proceed to the next target on hearing the word *fini* at the end of each sequence.

A timing mark was placed on the tape coincident with the onset of the target-bearing word. This timing mark, inaudible to the subjects, started the clock of a minicomputer, which was stopped by the subject's keypress response. Actual positions of timing mark and word onset were measured and the collected response times adjusted to reflect response time exactly from word onset.

Results

Response times shorter than 100 ms or longer than 1000 ms were discarded. Data discarded for this reason, plus missed targets, accounted for 3% of total responses.

Means for each condition of phonetic structure (CV, CVC, CCV) by

---

[1] This experiment was carried out in collaboration with Jean-Yves Dommergues.

TABLE 1
Mean RT (in milliseconds) per Condition for Experiment 1

	Words	Nonwords	X
CV	435	464	450
CVC	440	482	461
CCV	451	491	471
X	443	482	

lexical structure (word, nonword) are shown in Table 1. Words were re-sponded to significantly faster than nonwords ($F_1(1,28) = 57.3, p < .001$, $F_2(1,44) = 13.08, p < .001$). The difference between mean RTs to items of different phonetic structure was significant in the analysis by subjects ($F_1(2,56) = 4.8, p < .02$) but failed to reach significance in the items analysis ($F_2(2,44) = 1.43$). The interaction between these two factors was not significant, and separate analyses on the word-nonword effect showed it was significant for each phonetic structure separately at at least the .005 level.

Discussion

This experiment showed a lexicality effect, in direct contradiction both to McClelland and Elman's (1986) prediction that lexicality effects should never occur when target phonemes are clearly articulated and in word-initial position, and to Foss and Gernsbacher's (1983) prediction that controlling vowel identity should remove any lexicality effects; in-stead, the result is as predicted by both the serial multiple-outlet models described above. In conjunction with the Segui et al. (1981) failure to find a lexicality effect with bisyllabic items, however, the result is exactly as predicted by Cutler and Norris' (1979) race model: lexicality effects ap-pear with short target-bearing items but not with longer ones.[2]

Foss and Gernsbacher (1983) did not publish a full list of their mate-rials. It is possible that their failure to find an effect of lexicality resulted more from a majority of polysyllabic target-bearing items than from vowel identity effects in most of their experiments. Their Experiment 4, however, used all CVC items, and failed to find a word-nonword differ-ence; thus there is one result which is still in contrast with the present

[2] It might seem probable that an interactive activation model would also make this pre-diction, on the grounds that activation from the lexical representation of the target-bearing word would appear earlier for short words than for longer words. However, a worked-out example by McClelland and Elman (1986, p. 32) clearly shows that in their model, at least, activation of the initial [t] in the word *target* has reached response criterion before lexical facilitation is available even for the first syllable (e.g., from *tar).*

experiment's results. Could it be that the present experiment introduced some new low-level confound, even though the particular low-level confound which Foss and Gernsbacher identified was controlled for?

It is, of course, difficult to design an experiment to test for the presence of an unidentified acoustic/phonetic factor. However, there is a possible route for testing whether an observed apparent lexicality effect is in fact due to a low-level acoustic/phonetic artefact. All human listeners use the same auditory apparatus, and should be subject to the same low-level acoustic/phonetic effects such as vowel length. Lexicality, however, is a language-specific property. It is possible to determine whether an apparent lexicality effect is due to a low-level confound, therefore, by replicating the original experiment exactly except in that the subjects are not speakers of the language in question. The acoustic/phonetic factor hypothesis predicts that the results will replicate exactly with the new subject population.

It was decided, therefore, to replicate Experiment 1 with subjects who were native speakers not of French, the language in which Experiment 1 was conducted, but of a language with similar phonetic stock but very different lexical stock, namely English. Foss and Gernsbacher's (1983) hypothesis that phoneme-monitoring response latencies are solely reflective of low-level prelexical factors predicts the same results as we found in Experiment 1, including the apparent lexicality effect.

EXPERIMENT 2

Subjects and Materials

Subjects were 18 members of the Applied Psychology Unit subject panel. Each was paid a small sum for participating. Half of the subjects heard the sets in each of the two possible orders. All of the subjects were native speakers of British English; none was fluent in French, and none had studied it as far as or beyond the end of secondary schooling. The materials were those used in Experiment 1.

Procedure

Subjects were instructed in French grapheme-phoneme correspondences. They were permitted to hear the initial practice list repeatedly if they wished (but in fact none asked for a repeat). In all other respects the procedure was as in Experiment 1.

Results

RTs longer than 1000 ms or shorter than 100 ms were discarded; these with missed targets constituted 6.3% of the total data. Mean RTs per condition are shown in Table 2. Analyses of variance showed that both main effects and interaction were insignificant. The means for the phonetic structure conditions, however, were ordered exactly as in Experiment 1.

TABLE 2
Mean RTs (in milliseconds) per Condition for Experiment 2

	Words	Nonwords	*X*
CV	496	488	492
cvc	511	501	506
ccv	524	521	522
X	513	506	

Discussion

These results suggest that the results of Experiment 1 were not due to an unidentified low-level confounding—the lexicality effect appeared with speakers of the language in question but not with speakers of another language, whereas the weak effects of phonetic structure of the target-bearing item were precisely comparable across the two experiments.

We therefore decided to attempt to establish the lexicality effect in English also. Moreover, we decided to replicate the characteristics by which Foss and Gernsbacher's (1983) Experiment 4 differed from our Experiment 1. As in Foss and Gernsbacher's Experiment 4, subjects in our next experiment listened for only a single target, and all target-bearing items and all filler items had consonant-vowel-consonant (CVC) structure.

Furthermore, although vowel identity was again controlled across words and nonwords, we decided to undertake a direct test of Foss and Gernsbacher's hypothesis that vowel length affects phoneme-monitoring response latency, by measuring the actual length of our items and correlating item lengths with response times. We took two measures on each item: total item length (from the burst of the initial stop consonant to the offset of the final consonant) and length of the initial CV sequence. (Vowel onsets, especially after stops, are hard to identify, consonant onsets easier. The total length measure allowed us to assess the contribution of the final consonant; the CV length measure was essentially a vowel length measure, since the initial C was always either [b] or [d], which have very similar durations.)

EXPERIMENT 3

Materials

Twenty-two word-nonword CVC pairs were constructed, 11 pairs with initial [b], 11 with initial [d]. Within each pair the medial vowel was held constant, and only the final consonant changed. The materials are listed in the Appendix.

Four hundred sixty-five further CVC strings were chosen, of which half were words and half nonwords. Two sets of lists, a [b] set and a [d] set, were constructed. Each set contained

50 sequences, with roughly equal numbers of sequences four, five, and six items in length. Twenty-two sequences in each set contained one of the experimental items, always in penultimate (i.e., third, fourth, or fifth) position in its list. Each word was matched with its paired nonword on position in list. Of the remaining sequences, 10 contained no occurrence of the target, while the remainder contained a target-bearing item in first, second, or final position.

The filler items preceding target items were also controlled. An equal number of target-bearing words and nonwords was preceded by words and by nonwords. The items preceding a target item never began with a stop consonant or an affricate. These controls were necessary since under certain conditions monitoring responses may be affected by lexical status of the preceding item (Foss & Blank, 1980) or by phonetic similarity of the preceding item's initial phoneme to the specified target (Newman & Dell, 1978).

Each set was preceded by a practice list of similar structure. The sets were recorded by a female native speaker of British English. The lists were produced at a rate of approximately one item per second.

Subjects

Twelve members of the Applied Psychology Unit subject panel participated in the experiment, for which they were each paid a small sum. All were native speakers of British English. Six subjects heard the [b] set first, 6 subjects the [d] set.

Procedure

Subjects were instructed to listen for any item beginning with the specified target and to press the single response key as soon as they detected one. A timing mark, inaudible to the subjects, was aligned with the onset of each target-bearing word and started the clock of a minicomputer, which was stopped by the subject's keypress response. At the end of the first set subjects were instructed of the change of target.

The items were digitized and the exact interval between timing mark and word onset determined; the measured response times were then corrected for these values. Total word length, and length of the initial CV sequence, was also measured for each item.

Results

RTs longer than 1000 ms or shorter than 100 ms were discarded. These plus missed targets accounted for 3% of the total data. The mean RTs for each condition are shown in Table 3. No main effect or interactions reached significance (all $Fs < 1$).

Mean item RTs were correlated with the obtained length measurements. RT did not, overall, correlate significantly either with total item length $(r = .02)$ or with CV length $(r = .14)$. However, the lack of signifi-

TABLE 3
Mean RTs (in milliseconds) per Condition, Experiment 3

	Words	Nonwords	X
[b]	521	527	524
[d]	521	522	522
X	521	524	

cance was due only to the [d] items (r with total length -.21, with CV length = -.07); mean item RT for the [b] items did correlate significantly with length (r with total length = .52, $p < .02$; r with CV length = .50, $p < .02$).

Discussion

The results of this experiment were not as predicted by the hypothesis that lexicality effects will always appear with short target-bearing items. Instead, the results essentially replicate Foss and Gernsbacher's finding that monitoring for a single phoneme target on a CVC syllable in a list of other CVC syllables shows no response time difference to words versus nonwords. Moreover, we found support for their hypothesis that vowel length influences response time—albeit only for half our materials, and notably *not* for those items which more directly replicated Foss and Gernsbacher's study, namely the [d] items.

We argued that the results of Experiment 1 showed a true lexicality effect, as predicted for monosyllabic items by Cutler and Norris' (1979) race model. But the present experiment (and Foss & Gernsbacher's (1983) Experiment 4) also used monosyllabic items, and showed no lexicality effect. If the results of Experiment 1 are—as we argued—not artefactual, what differences between the experiments could possibly account for the differences between the results?

We decided to begin with the most glaringly obvious difference: that Experiment 1 was conducted in French, but Experiment 3, and Foss and Gernsbacher's experiments, in English. Although we find it quite incredible that there could be cross-linguistic processing differences which would result in a lexicality effect in phoneme monitoring in one language but not in another, it could conceivably be the case that, say, monosyllabic nonwords were of necessity less wordlike in French than in English, making a word-nonword distinction more salient. Our next experiment therefore consisted of a replication of Experiment 1 in English.

EXPERIMENT 4

Materials

Twenty-five word-nonword pairs were chosen, each pair with the same vowel. Two pairs had CV structure, 14 pairs CVC structure, and 9 pairs CCV structure. All except one CCV pair had a complementary pair with the identical phonetic alteration effecting a change in the reverse direction (e.g., *car-dar, dye-kye; dash-tash, tack-dack; blue-ploo. plough-Mough*). Limitations of the English language constrained the number of CV and CCV items chosen. The materials are listed in the Appendix.

Filler items were chosen and arranged as in Experiment 1. Each list terminated with the word *end*. The lists were recorded by a native speaker of British English, at a speaking rate of approximately one item per second.

Subjects and Procedure

Subjects were 20 members of the Applied Psychology Unit subject panel, who were paid a small sum for participating. All were native speakers of British English. Half the subjects heard the two sets of lists in AB, half in BA order. The procedure was as in Experiment 1.

Results

Response times longer than 1000 ms or shorter than 100 ms were discarded. These, with missed targets, totaled 4% of all the data.

Mean response times for each condition are shown in Table 4. The main effect of word-non word status was highly significant in the analysis by subjects ($F(1,18) = 9.03$, $p < .008$), but missed significance in the analysis by items ($F(1,44) = 2.38$, $p = .13$). The effect of phonetic structure was also significant in the subjects analysis ($F(2,36) = 3.41$, $p < .05$) but not in the items analysis ($F(2,44) = 1.04$). The interaction of the two main effects was not significant. Post hoc analyses showed that targets on CVC items were responded to significantly faster than on either CV or CCV items. Further post hoc comparisons showed that the word-non-word difference was significant in the 11 V-final pairs ($F_1(1.18) = 7.8$, $p < .02$; $F_2(1,20) = 4.35$, $p < .05$), but not in the 14 C-final pairs.

The V-final (CV and CCV) and CVC sets were imbalanced with respect to target phoneme. However, inspection of the item means showed that this was not an important factor. Within the V-final set, lexicality effects were substantial for bilabial (65 ms advantage of words over nonwords), alveolar (27 ms), and velar (53 ms) stops. Within the CVC set, neither bilabial nor alveolar stops showed a lexicality effect, though the single velar pair did.

Discussion

This experiment showed a lexicality effect at least for the V-final items. In Experiment 1 all three item types had shown a clear lexicality effect. We felt justified in rejecting the hypothesis that the previous pattern of results reflected intrinsic differences between languages. Accordingly we set out to investigate the many methodological differences between experiments which might cause lexicality effects to appear and disappear.

TABLE 4
Mean RTs (in milliseconds) per Condition, Experiment 4

	Words	Nonwords	X
CV	409	435	422
CVC	397	400	399
CCV	402	455	428
X	400	423	

We began with number of targets. The two similar experiments which had failed to show lexicality effects—our Experiment 3 and Foss and Gernsbacher's Experiment 4—had involved monitoring for a single target throughout the experiment. Experiments 1 and 4 had used multiple targets, with target specification varying from trial to trial. We therefore decided that our next study would replicate Experiment 3 as closely as possible except that the target specification would vary.

EXPERIMENTS

Materials

The materials set from Experiment 3 was expanded so that every word-nonword pair had, as in the preceding experiment, a complementary pair in which the same change resulted in a reverse effect e.g., *boat-boke, poke-pole.* Thus not only vowel identity and initial consonant identity but also final consonant identity were controlled across words and nonwords. [t] and [p] were added to [b] and [d] as phoneme targets. There were a total of 40 word-nonword pairs. They are listed in the Appendix.

Five hundred twenty further CVC words and nonwords (260 of each) were chosen, and arranged in 120 lists, 40 each of four, five, and six items in length. Eighty lists contained an experimental item in penultimate (i.e., third, fourth, or fifth) position. Of the 40 filler lists, 20 contained a target-bearing item in first, second, or final position, while the remaining 20 contained no occurrence of the specified target. Each of the four targets served 30 lists. Two sets of 60 lists were drawn up, with all factors counterbalanced across lists. Each member of a given pair and its complement occurred in the same position in their respective lists.

A set of 12 practice lists of similar structure to the experimental lists was also compiled. The lists were recorded by a female native speaker of British English at a rate of approximately one item per second. The target for each list was specified immediately prior to the list, using a CVC name as example, e.g., *b* as in *Ben* or *p* as in *Pain.*

Subjects and Procedure

Twelve members of the Applied Psychology Unit subject panel took part in the experiment, and were paid a small sum. All were native speakers of British English. Half the subjects heard the two sets of lists in AB, half in BA order. Except that there were separate target specifications for each list throughout the experiment, the procedure was as in Experiment 3.

Results

RTs over 1000 ms or under 100 ms were again discarded; with missed responses these accounted for 2.4% of the total data.

Mean RTs for each condition are given in Table 5. The word-nonword effect narrowly missed significance ($F_1(1,11) = 4.36$, $p < .06$; $F_2(1,72) = 3.29$, $p < .075$). The main effect of target phoneme was not significant in either analysis and did not interact with the word-nonword effect. However, it can be seen from Table 5 that, as in Experiment 3, the word-nonword effect is stronger with bilabial than with alveolar targets. Post hoc t tests showed that the word-nonword difference was significant for items

TABLE 5
Mean RTs (in milliseconds) per Condition, Experiment 5

	Words	Nonwords	X
[t]	398	395	397
[p]	385	425	405
[b]	397	422	409
[d]	394	404	399
X	394	411	

beginning with [p] (t(ll) = 3.38, $p < .01$), but not for [b] items $(p > .15)$ or for [d] or [t] items.

A correlation analysis was again carried out on the mean item RTs and the length measurements. Collapsed across target phonemes, correlations with both total item length and CV sequence length were insignificant. Correlations for each target phoneme separately were also all insignificant. The correlation which came nearest to significance was that for [b] items with total length $(r = .41, p < .07)$.

Discussion

The results of this experiment were disappointingly inconclusive. Experiment 3 had shown no trace of a lexicality effect, but some evidence of vowel length correlations. The present experiment produced no significant vowel length effects but some evidence of a lexicality effect.

We reasoned that varying the number of targets was perhaps a step in the right direction toward identifying the methodological variable or variables involved in the comings and goings of the lexicality effect, but it could not be the whole story; what was needed was to identify the more general methodological characteristic of which variation in number of targets was merely an instance. There are two ways in which a phoneme-monitoring experiment with multiple targets can be considered to differ from a phoneme-monitoring experiment with a single target: multiple targets can be seen as making the task more difficult, since the subject has to adjust to a new target specification for each list; or they can be seen as making the stimuli more variable, and hence breaking the monotony of single-target monitoring.

Both task difficulty and stimulus monotony are viable candidates for a general methodological characteristic which might be expressed in several different ways and might well affect the way phoneme monitoring is performed and hence the likelihood of lexicality effects appearing. We decided to investigate the two variables separately. We began with task difficulty, since, as discussed in the introduction, this was the variable which Foss and Blank had hypothesized was chiefly involved in deter-

mining which of the dual codes was to be involved in a given monitoring response.

One manipulation of task difficulty in phoneme monitoring is well established. As discussed above, it has been repeatedly observed that monitoring for more than one target simultaneously is more difficult than monitoring for a single target (Foss & Dowell, 1971; Steinheiser & Burrows, 1973; Treisman & Squire, 1974). Foss and Gernsbacher (1983) had also attempted this manipulation, but without controlling for vowel identity. Most importantly, the Rubin et al. (1976) study, which showed a lexicality effect with CVC targets, had used simultaneous monitoring for two targets.

Accordingly, our next experiment required subjects to monitor simultaneously for the occurrence of either of two targets; as Rubin et al. had done, we compared response with a single key to either target (in which there is only one source of difficulty, namely memory load involved in two target specifications) with response to separate keys for the two targets (in which to memory load for two targets is added a second source of difficulty, namely that of making the response decision). In other respects the experiment amounted to a replication of Experiment 3.

EXPERIMENT 6

Materials

The materials were those of Experiment 3, rearranged into two sets with equal numbers of [b] and [d] targets in each set. The Experiment 3 tape was cross-recorded into the new order.

Subjects

Twenty-four members of the Applied Psychology Unit subject panel, all native speakers of British English, participated. Each was paid a small sum. Half the subjects responded with one key for either [b] or [d] targets, half with separate keys for [b] and [d]; within these groups, half heard the two sets in AB, half in BA order.

Procedure

The procedure was as in Experiment 3, except that subjects were monitoring for both [b] and [d] throughout the experiment. For the subjects who had separate keys for [b] and [d], half had the [b] key on the left, half the [d] key on the left.

Results

RTs over 1000 or under 100 ms were discarded; with missed targets 6.8% of the total data was lost in this way. Mean RTs for each condition

are shown in Table 6. The main effect of response type was almost significant in the analysis by subjects ($F_1(1,22) = 3.82$, $p < .065$) and highly significant in the items analysis, where it was of course a within-items factor ($F_2(1,40) = 60.61$, $p < .001$). The word-non word difference, however, was not significant, nor was the main effect of target phoneme type (all $Fs < 1$). The only interaction to approach significance was that between word-nonword status and target phoneme type: across both conditions, [b] items showed more evidence of a lexicality effect than [d] items ($F_1(1,22) = 9.00$, $p < .01$; $F_2(1,40) = 3.42$, $p < .075$). Correlations with measured item length were again performed. Across the complete set of items, the correlation with total length was insignificant, but that with CV length significant ($r = .33$, $p < .03$). This was again entirely due to the [b] items, which correlated significantly on both measures (total length: $r = .47$, $p < .03$; CV length: $r = .70$, $p < .001$), whereas the [d] items were again not correlated with length ($r = .03$ and .06, respectively).

Discussion

These results give little support to the suggestion that task difficulty is a crucial factor in the appearance of lexical status effects on phoneme-monitoring response time. Although the disjunctive response condition was clearly more difficult than the single-key response condition, since the response times were considerably longer, there was no suggestion that the more difficult condition produced larger lexical status effects. Nor did the addition of another target (the effect of which can be assessed by comparing the present results with those of Experiment 3) result in a significant increase in lexicality effects.

Manipulating the second suggested factor, task monotony, was less straightforward. In a sense we had already manipulated it, in the comparison between Experiments 4 and 5. Both used monosyllabic target-

TABLE 6
Mean RTs (in milliseconds) per Condition, Experiment 6

	Words	Nonwords	X
2 targets,			
1 response			
[b]	519	543	531
[d]	536	495	515
2 targets,			
2 responses			
[b]	586	601	594
[d]	601	599	600
X	560	560	

bearing items and multiple-phoneme targets, and were procedurally similar; but Experiment 4, in which lexical status effects appeared, was notably less monotonous than Experiment 5 in that the lists of stimuli mixed monosyllabic and bisyllabic items, whereas the Experiment 5 lists were uniformly monosyllabic. We decided to attempt to replicate this comparison between the two experiments by forcing the same comparison with Experiment 1, which also showed lexical status effects. It was also possible to make the between-experiments comparison more cleanly a second time around; there were a few other procedural differences between the English Experiments 4 and 5, but in setting up the next experiment as an analogous comparison to the French Experiment 1, it was possible to keep all procedural details the same *except* that this time we used no bisyllabic fillers.

EXPERIMENT 7

Materials and Procedure

Twenty CVC French word-nonword pairs were chosen, arranged in complementary pairs as in Experiments 4 and 5, so that initial C, medial V, and final C were controlled across words and nonwords. Examples are *bonne-bomme, pomme-ponne.* The full set is listed in the Appendix. As in Experiment 5, [t], [d], [b], and [p] were used as target phonemes.

Two hundred twenty-two further filler items were chosen, of which half were words. All filler items had CVC structure. The experimental items with filler items formed two blocks of 40 lists, compiled and counterbalanced as in Experiment 1. The lists were recorded by a female native speaker of French at a rate of approximately one item per second. The procedure was as in Experiment 1.

Subjects

Subjects were 20 students of the University of Paris V. Half the subjects heard the two sets in AB, half in BA order.

Results

RTs over 1000 or under 100 ms were discarded; with missed responses 6.5% of the total data was thus lost. Mean RTs for each condition are shown in Table 7. The word-nonword effect was not significant *(F < 1)*.

TABLE 7
Mean RTs (in milliseconds) per Condition, Experiment 7

	Words	Nonwords	X
[t]	498	506	502
[p]	516	498	507
[b]	478	459	469
[d]	496	520	508
X	497	496	

Discussion

Experiment 7 closely mimicked Experiment 1. However, while Experiment 1 showed a significant effect of lexical status, Experiment 7 did not.

The English language Experiments 4 and 5 (although not quite as closely parallel as Experiments 1 and 7) show a similar pattern: an effect of lexical status in Experiment 4, none in Experiment 5.

All four experiments used monosyllabic target-bearing words, multiple phoneme targets, and a single target per list. In each pair of experiments, subjects were drawn from the same population, and the procedure was either highly similar or identical. In all experiments vowel identity was strictly controlled. It was not the case that response times in Experiments 5 and 7 were simply too fast for a significant lexicality effect to show up: mean response time in Experiment 5 was 402 ms, as against 411 ms for Experiment 4, while mean response time in Experiment 7 (496 ms) was actually slower than mean response time in Experiment 1 (463 ms). The one undeniable difference between the two experiments was that the filler items used to construct the lists differed sharply: Experiments 5 and 7 used entirely monosyllabic items, Experiments 1 and 4 had a majority of bisyllabic fillers.

Our results, therefore, suggest that stimulus monotony can effectively determine whether lexical status exercises an effect on the phoneme-monitoring response. However, we do not propose that stimulus monotony is in any sense an independent factor to be accounted for in processing models. As we describe below, we consider it to be merely one methodological characteristic which is capable of precipitating a very general effect, namely a shift of attention between processing levels. The next section also considers the implications of our results for the description of phoneme-monitoring task performance, and for the modeling of comprehension processes.

GENERAL DISCUSSION

Levels of Processing and Attention

Our results fall into the coherent pattern into which fit also the previous findings reviewed in the introduction. Lexicality effects on phoneme-monitoring responses can indeed be made to come and go. Thus our results directly contradict the predictions of models as diverse as those of Foss and Gernsbacher (1983) and of McClelland and Elman (1986) that phoneme monitoring should never produce lexicality effects. We showed that it is possible to make lexicality effects disappear if the stimuli are monotonous, and reappear if the stimuli are more varied. In

our experiments, varying the filler items was sufficient to remove a monotonous effect; a steady stream of CVC items appeared to produce mainly prelexically based responses, but a mixture of monosyllabic and bisyllabic fillers produced a higher proportion of lexically sensitive responses. Varying the number of targets appeared to have some small effect of relieving the monotony also.

Rubin et al. (1976), however, also used all-CVC lists, in two experiments, and in both cases found significant effects of lexical status. There are several respects in which the Rubin et al. experiments may be considered to be less monotonous than ours. First, the ear of presentation of the stimuli was varied from one block of trials to the next. All subjects had some stimuli presented to the left ear, some to the right, some binaurally. Second, in one of their experiments Rubin et al. randomly varied the interstimulus interval, so that subjects could not predict when the next item would be presented. Third, in their other experiment Rubin et al. blocked items by lexical status so that each sequence consisted of all meaningful words or all meaningless nonwords. (Subjects in the Rubin et al. (1976) experiments were also monitoring for two targets simultaneously, but this manipulation, as we have shown, does not have a strong effect on the appearance of lexicality effects.)

Thus, although in our experiments all-CVC stimuli were apparently monotonous, we would argue that this is not necessarily the case. Stimulus monotony is not a variable intrinsically connected with particular stimulus structures. For instance, a sequence of bisyllables with a repeating stress pattern should also be monotonous (and we would predict no lexicality effects with such stimuli, though, as pointed out above, we claim that lexicality effects are less likely with bisyllabic stimuli in any case). Any highly homogeneous stimuli run the risk of being monotonous, but it is possible for other characteristics of the experimental situation to keep subjects interested in the speech input and hence processing it in a normal way. Some such characteristics may involve, for instance, the payoff weightings of the target detection and comprehension tasks, as described immediately below. Alas, in our all-CVC experiments no such counteracting factors were present; the monotony was apparently unrelieved.

Phoneme-monitoring experiments require subjects to perform two tasks—listening to speech and detecting a target. We suggest that the effect of varying stimulus monotony is to shift attention between these two tasks. If the stimuli are highly monotonous, it is boring to listen to them as speech. That is, subjects will not be concentrating attention on the meaning of the incoming speech, and lexical access will be at a low priority. Instead, attention will be focused on the target detection process, increasing the likelihood that detection responses will be based

purely on prelexical representations. Conversely, if the stimuli are inter-
esting, attention is likely to be shifted to the comprehension task, and
hence away from the target detection task; under these conditions the
likelihood of lexically sensitive responding rises.

A very simple unpublished study by Dommergues and Segui supports
this suggestion that monotonous stimuli encourage a shift of attention
away from lexical processing. Listeners were presented with lists of 30
items, of which, in each list, 15 were words and 15 nonwords. They were
then asked to estimate the proportion of the items which had been real
words. The estimates produced by subjects who had heard a list of mixed
bisyllabic and monosyllabic items (as in Experiments 1 and 4) were not
significantly different from the correct answer of 50%; but subjects who
had heard a list of all monosyllabic items (as in Experiments 5 and 7)
produced a mean estimate significantly lower than the actual value, and
significantly lower than the mean estimate of the first group. That is,
subjects presented with the more monotonous list seemed to be paying
less attention to processing the stimuli as words.

If monotony is indeed merely one instance of an effect which results in
a shift of attention, then we should be able to find other manipulations
which will exercise the same attention-shifting effect. For instance, it is
reasonable to expect that varying the payoff of the two tasks may simi-
larly affect which of them is attended to.

The previous factor and this are no doubt closely related. The effect of
stimulus monotony is to make language processing less interesting, hence
to lower its payoff in terms of message yield, and hence to shift attention
away from it. What other aspects of the experimental situation could in-
fluence relative task priority?

If subjects are encouraged to regard the language comprehension task
as primary, they will be more likely to attend to it. Dell and Newman
(1980) explicitly manipulated this factor by comparing lexical effects on
phoneme monitoring (in sentences) in a condition in which subjects were
given a recall test, without preparation, at the end of the experiment, and
in a condition in which they were required to paraphrase each sentence.
The latter condition was held to concentrate attention on the comprehen-
sion task, and indeed, the effects of the lexical characteristics of the
target item were significant in that condition but not in the recall condi-
tion.

On the other hand, there are situations under which subjects may re-
gard the target detection task as primary. For instance, if target detection
is frequently required (i.e., there are relatively few filler items which do
not require a response), the detection task may be the focus of attention.
Most phoneme-monitoring experiments using lists of isolated items, for
instance, use a ratio of fillers to target-bearing items of between 5:1 and

10:1 (the mean for the present experiments was approximately 7:1). Foss and Gernsbacher's (1983) Experiment 4, however (which, as mentioned above, failed to find any trace of lexicality effects), used a ratio of less than 2:1. It is likely that in these conditions subjects assigned higher priority to the detection task, thus increasing the probability of prelexically based responding.

Stimulus monotony therefore appears to be just one of a number of factors which can affect whether attention is focused on lexical processing or on target detection in a phoneme-monitoring task. Samuel and Ressler (1986) have proposed that attention may similarly be shifted between word and phoneme perception in an experimental situation involving the phonemic restoration illusion. This illusion was first studied by Warren (1970), who showed that when a phoneme is entirely excised from a word but replaced by noise, listeners usually report hearing the word intact, but with some extraneous noise. Samuel (1981) found that listeners could not reliably discriminate between versions of real words in which noise had actually replaced a phoneme, on the one hand, and in which it had merely been added to the phoneme, on the other. However, by telling subjects which phoneme in a word to attend to, Samuel and Ressler (1986) were able to elicit significantly more accurate discriminations. That is, subjects were able to switch attention from the lexical level, which normally mediates the phonemic restoration illusion, to the phonetic level, exactly as we have proposed that they do under different conditions of the phoneme-monitoring task.

We believe that this account of phoneme detection performance in terms of attentional factors has a strong claim to ecological validity. The ability to switch attention between different levels of linguistic analysis has an obvious role in everyday language processing. On some occasions a listener may wish to attend specifically to the sounds of language—in order to acquire the pronunciation of a new word, perhaps, or to identify a speaker's dialect. Puns depend for their effect on listeners' ability to attend to sounds at the expense of meaning, for instance. On other occasions, though, the sounds of language will be irrelevant, and paying particular attention to them could detract from the more important process of extracting meaning from speakers' utterances. Under these conditions the listener will attend to the higher levels of processing. Thus the ability to switch attention between levels such as the phonemic and the lexical is independently motivated; our explanation of the empirical data in these terms has no ad hoc character.

Moreover, we believe that this account provides a basis for distinguishing between classes of models of the relationship between phonemic and lexical processing. The oppositions which we identified in the introduction are addressed in the next sections.

Serial Versus Interactive Levels of Processing

Both general classes of model, the serial and the interactive, are very powerful, and most patterns of results can be easily modeled within both general frameworks. As we saw in the introduction, accounts of the phoneme-monitoring data have been advanced within both serial and interactive models. It may prove ultimately impossible to reject either class of model solely on empirical grounds. However, it seems to us that certain types of explanation are more naturally modeled in one way than another. Specifically, we believe that an account of lexicality effects in phoneme monitoring involving the shifting of attention between processing levels fits very naturally into a serial model of comprehension incorporating multiple-outlet points for phoneme detection responses. In any model which postulates autonomous processes it should obviously be possible to alter the outcome of separate processes independently. By shifting attention between lexical and phonemic processing it should be possible, therefore, to control the effective outcome of a race between these processes to output a response; and since each of these processing levels has an outlet point for the detection response, the response can issue from the outlet at whichever level is being attended to. This is the framework within which our explanation of the effects of stimulus monotony was presented in the preceding section.

Nevertheless, it is certainly possible to fit our attention-shifting explanation into an interactive activation framework. We believe, however, that the result would be a less natural combination. In an interactive activation model both phonemically based and lexically influenced responses are the product of a single integrated system. In TRACE (McClelland & Elman, 1986), for instance, or in the Stemberger et al. (1985) interactive activation account of phoneme-monitoring performance, all detection responses in monitoring experiments result from activation of a "node" at the phonemic level of the network. These models make no provision for recovery of phonemic information directly from the lexical nodes. Thus attentional factors cannot alter the degree of lexical involvement in the phoneme detection response simply by shifting attention between phonemic and lexical nodes; attention must to some extent at least remain fixed on the phonemic nodes, because in them resides the only information on which a response can be based. When there is only one outlet point, attention during responding must be directed to that point.

Within the interactive activation framework one can alter the relative contribution of phonemic and lexical information by changing the weightings on the connections between lexical and phonemic nodes. For instance it would be possible to reduce the strength of the bottom-up connections so that phonemic information became less likely to activate

word level nodes. Similarly, reducing the top-down weightings would make partial activation of a word node rather less likely to activate the word's constituent phonemes. Simultaneous reductions in the weightings of both directional connections would, of course, also be possible. Any of these alterations would allow the interactive activation model to mimic the effects of attention shift between response outlets. What the model requires to achieve this, in other words, amounts to a gain control operating over all the connections between phonemic and lexical nodes in the network. This could quite easily be implemented in TRACE, for example, by attaching additional excitatory and inhibitory connections to all lexical nodes and connecting them to a single node which could act as a master gain control.

However, we see two ways in which such a modification is undesirable on general principles. First, the fact that it adds an additional free parameter to the model is undesirable. The predictions made by any interactive activation model result from the connective structure of the network in the model, and the setting of the connection parameters within the network. Whenever such a model is extended to allow greater freedom in parameter setting, the predictive power of the model becomes accordingly less constrained. A model with free-roaming parameters can predict anything. Thus an interactive activation model extended to incorporate an attention-shifting account of the variability of the lexicality effects in phoneme detection loses discriminatory power.

Second, such a change in the model would in any case be unmotivated apart from the need to explain laboratory results such as those of the present experiments.[3] Shift of attention between processing *levels*— which, we have argued, is necessary in any model in order to account for our ability to concentrate to a greater or lesser degree on perceiving specific pronunciations on certain occasions—can be incorporated equally well into any model. But only when the model postulates response outlets at each separate level does the shift of attention between levels automatically bring about a shift of attention between response outlets as well. Thus there is a very natural account of our results within a model incorporating independent processes with independent response outlets. There is no knockdown argument for or against either model in these results—as there indeed may never be, given the power of each class of model; but the interactive activation framework encompasses an attentional shift explanation rather less elegantly than does a serial account.

---

[3] Rumelhart and McClelland (1982) also suggested that variable control over the letter-to-word inhibition parameter in their model of visual word recognition would allow their model to account for the data of Carr, Davidson, and Hawkins (1978). However, they did not suggest how the control variation should be determined.

In the next sections we discuss the various versions of serial model which we identified in the introduction, and suggest ways in which our results may help to distinguish between these as well.

Single Versus Multiple Outlets

Within a serial framework, the question of single- versus multiple-outlet points is easily resolved by our results. Foss and Gernsbacher's (1983) single-outlet serial model must be rejected, since it depends for its rationale on the nonexistence of lexicality effects on phoneme-monitoring response latencies; all previous results showing apparent lexicality effects could only be phonetic artefacts, according to their account. However, we have demonstrated that lexicality effects can be found even when Foss and Gernsbacher's postulated phonetic effects have been strictly controlled. We believe it is no longer necessary to try to explain previous results away. Lexicality effects do exist in phoneme monitoring; varying task monotony is one way of making them come and go. If lexicality effects can come and go, serial single-outlet models are inadequate, since when the direction of information flow is strictly bottom-up, responses from any outlet point can show effects only of information available at and below that point. Thus a single phonemic level outlet, such as Foss and Gernsbacher's model proposed, cannot show lexical effects (and a single lexical level outlet could not fail to show them). Therefore this class of models can be definitely rejected.

Race Model Versus the Dual Code

The most preferable way of accounting for the overall pattern of results from experiments on lexicality effects in phoneme monitoring is therefore within a serial model incorporating more than one possible response outlet point. In the introduction we discussed two versions of such a model, Cutler and Norris' (1979) account involving a race between mandatory processes with separate response outlet points, in every case, and Foss and Blank's (1980) account involving mandatory successive phonetic and lexical representations from which target information can be extracted, with alternative response outlet points of which only one alternative may perhaps be available in a given case. As we showed in the introduction, there is little to choose between these two versions of the general serial multiple-outlet account. However, some finer details of our results do lead us to express a preference for the race model, with its mandatory sequence of processes which is identical in every case.

Recall that the race model lays great stress on the prediction that lexical effects can only be facilitative. This is because target detection can always be accomplished on the basis of a prelexical representation. If lexical access is particularly fast, however, the lexical representation may become available before target detection from the prelexical representation is completed. If lexical access is slow, no effects will be observed,

because the prelexical representation will be used as in the default case. Thus lexical factors can never inhibit phoneme-monitoring responses.

This feature of the model has the odd entailment that lexicality effects on phoneme-monitoring response time will always be weak and fragile. The baseline condition is the time taken for the prelexically based response; responses can be facilitated with respect to the baseline, but never inhibited. Thus observed effects will be small. All nonwords and many words will be responded to prelexically; some words will be responded to with lexical mediation.

Of course, the target detection process does not necessarily take an equal time for each phoneme. Some phonemes are perhaps more encoded and harder to identify, and for these the detection process will take fractionally longer. Under these circumstances the likelihood of lexical access being completed first will also rise. Thus lexical effects will be somewhat more likely with harder-to-identify phonemes.

Note that this means that there may be no significant effect of intrinsic target identifiability on response time, since a slowing in the target detection procedure will be compensated for by an increase in the number of times the lexical outlet response wins the race. However, on the basis of the present results, we would expect [b] to be harder to perceive than [d], since we found consistently larger lexicality effects for [b] than for [d] items. There is very little difference in intrinsic perceptibility between any of the stop consonants. But studies of perceptual confusions among consonants have shown that in fact [b] is somewhat more confusable than [d]. Wang and Bilger (1973) showed that misidentifications were higher for [b] than for [d]. Miller and Nicely (1955) found the same result for consonant identification in quiet conditions. Goldstein (1980) separated out the relative contributions of intrinsic distinctiveness and response bias to confusion matrix patterns; consonants with higher distinctiveness than response bias rankings he labeled relatively unambiguous, consonants with higher response bias than distinctiveness rankings were labeled relatively ambiguous. In prevocalic position [d] was ranked higher on distinctiveness than on response bias, [b] higher on response bias than on distinctiveness. Thus [b] is apparently a more ambiguous sound than [d]. As a more ambiguous sound, it is harder to perceive, and hence slows the target detection process and renders lexical outlet responses more likely. This is one possible reason for the consistent finding in our experiment of stronger lexical effects for [b] than for [d] items. Exactly this asymmetry was also found by Foss and Gernsbacher (1983), it should be noted: in both their Experiments 1 and 3 they found greater lexicality effects for [b] than for [d] items. Foss and Gernsbacher dismissed this finding as probably due to vowel identity, but we would argue that it represents a true asymmetry of lexicality effects, precipitated by intrinsic differences of perceptibility between [b] and [d].

Target identifiability may also be affected by Foss and Gernsbacher's preferred factor, namely identity of the following vowel—some vowels may make all or some consonants easier or harder to identify. Identifiability may also be affected by whether the sound following the target is a consonant or a vowel. We do not believe there is as yet enough evidence to make firm statements on these matters. However, we do reject Foss and Gernsbacher's claim that vowel *length* affects target detection time from a prelexical representation. On the contrary, we believe that the effect of vowel length (with the monosyllabic items under consideration in the present, and preceding, experiments, this is equivalent to item length) is precisely as predicted by the race model: increasing length increases lexical access time and hence decreases the likelihood of the lexical outlet response winning the race.

A consequence of this is that any set of items which shows length effects on response time should also show lexical effects. We claim that length effects—shorter RTs to shorter items—arise because for the shorter items the lexical outlet detection response has been facilitated to an extent that it has beaten the prelexical detection process. Thus these items should also show lexical effects.

Since Foss and Gernsbacher (1983) did not actually measure their items and hence do not actually have evidence of length effects, we cannot test this prediction against their results. However, we can test it on the results from the present experiments.

Recall that we had measured item length in 3 out of 7 experiments (Experiments 3, 5 and 6), and had found no consistent pattern of effects —sometimes length seemed to be correlated with RT, sometimes not. In particular, items beginning with [b] seemed to show much more evidence of length effects than items beginning with other phonemes. We predict that where there are length effects, there will be lexical effects.

Since (as predicted by the race model) word-nonword status effects are weak, we decided to undertake a more stringent test of the presence of lexical effects. Any effects present should show themselves only in the words, of course, as only word RTs can be facilitated by lexical effects. One very well-established effect upon response time to words in a wide variety of tasks is that of frequency of occurrence. Accordingly we determined the frequency of occurrence (Kucera & Francis, 1967)[4] of all of

---

[4] For those few items which were homophones (e.g., *dam-damn*) we summed the frequencies across the alternative readings. In addition to the surface frequency of each word, we computed the combined frequency of base plus inflected forms (e.g., *book, books, booking, booked)*. Separate correlations were computed for the combined and surface frequency measures. However, since for every single set of items the two correlations were either both significant or both insignificant, and indeed, for every correlation which was significant beyond the .05 level the correlation coefficient itself was identical for the two measures to the second decimal place, we are reporting only one coefficient in each case.

the words used in these experiments, and measured the correlation of mean RTs to these words with frequency.

In Experiment 3, the correlation of both length measures with RT was significant for the [b] items, but not for the [d] items and not for the complete set of items. The reaction times showed no significant lexical status effect overall, though a fractionally greater difference between words and nonwords beginning with [b] than between words and non-words beginning with [d]. We therefore predict significant correlations with frequency only for the [b] words. This is exactly what we find: [b] word RT shows a significant negative correlation with frequency (r = - .66, p < .03). The correlation for [d] word RT, however, is insignificant (p > .25), and across the complete set of items the correlation is likewise insignificant.

In Experiment 5, there was no significant overall lexical status effect, though [p] words were responded to significantly faster than [p] non-words. There were no significant correlations of item RT with length, either across the whole experiment or for any target phoneme separately. We therefore predict no significant overall correlation with frequency, and no effect for any phoneme but [p]. There is indeed no significant overall correlation with frequency. Correlations for no target phoneme set reached the set level of significance, but the highest correlation was achieved by the eight [p] words (r = - .64, p < .09).

In Experiment 6, we found that one of the length measures correlated significantly with RT across the whole set, but that this effect was again due solely to [b]-item RTs, which correlated significantly with both length measures, whereas [d]-item RTs did not. The reaction times showed no overall effect of word-nonword status, but a marginally significant interaction between target phoneme type and word-nonword status such that [b] words were on average responded to 20 ms faster than [b] nonwords; [d] words, however, were on average responded to 21 ms slower than [d] nonwords. We therefore predict frequency effects for the [b] words and not for the [d] words. This is exactly what the correlations show: the [b] word RTs are significantly negatively correlated with frequency (r = - .72, p < .02), the [d] RTs show no significant correlation (p > .2). Collapsed over the complete set of items, the frequency correlation, just as did the length correlation, this time reaches significance (r = - .61, p < .01).

Thus the pattern of frequency effects across these experiments exactly mimics the pattern of length effects, which is precisely the result predicted by the race model. The dual-code model, on the other hand, as set out by Foss and Blank (1980) and Foss and Gernsbacher (1983), predicts precisely the reverse: when there are length effects (or any kind of acoustically based effects), there should be no lexical effects. Thus we have managed to distinguish between the alternative serial multiple-outlet

models, and to provide evidence which seems to favor the race model over its rival.

We can adduce yet one further piece of evidence in favor of the race model. The model predicts that the shorter a word, the more likely it is to be responded to via a lexical outlet. The shorter, on average, is a set of words, the more likely are at least some of its members to be responded to via the lexical outlet. Length effects indicate lexically based responding. Thus the model predicts that the set of items with the shortest mean length will be the set of items to show the most significant length effects. This prediction is highly counterintuitive, since, on the face of it, it would appear that the more variation in length, the more opportunity there is for length effects to show up. Nevertheless the race model's prediction, in our case, is clearly that [b] items, which show the significant length (and other postlexical) effects, should be on average shorter than other items. They are. Table 8 shows the mean lengths for items beginning with each of the target phonemes used in these experiments.[5] An analysis of variance on the measures shows that in both cases the effect of target phoneme was significant (for total item length $F(3,72) = 3.53$, $p < .02$; for CV length, $F(3,72) = 4.15, p < .01$). Word-nonword status, on the other hand, had no effect.

The race model therefore provides a coherent account of the pattern of results across our experiments. Items beginning with [b] have an initial phoneme which is slightly harder to detect than other target phonemes. They also tend to be shorter. For both reasons they are more likely to be responded to via the lexical outlet. Thus [b] items show not only length but also frequency effects. The race model predicts this pattern of results. The dual-code model does not.

CONCLUSION

Our experiments have shown that phoneme-monitoring responses can under certain circumstances be sensitive to lexical characteristics of the word bearing the phoneme target, and we have identified degree of task monotony as one of the factors which determine when such lexicality effects occur and when they do not. We have argued that a parsimonious explanation of why monotony has this effect can be constructed in terms of attention shifting between levels of processing, an explanation which is independently motivated by listeners' ability to concentrate selectively on lower or higher levels of processing of speech signals.

We have maintained further that these findings provide a basis for dis-

---

[5] The items for Experiments 3, 5 and 6 were derived from a single original recording, so items used in more than one experiment only have one length measurement.

TABLE 8
Mean Length (in milliseconds) of Items Used in Experiments 3,5, and 6

	Words	Nonwords	X
Total length			
[t]	539	531	535
[p]	499	486	493
[b]	446	442	444
[d]	492	478	485
X	493	483	
CV length			
[t]	338	339	338
[p]	291	314	303
[b]	264	260	262
[d]	311	256	283
X	301	290	

tinguishing between alternative theoretical descriptions of the relationship between prelexical and lexical processing in auditory comprehension. First, the simple model recently proposed by Foss and Gernsbacher (1983) can be rejected. Of the remaining models, we have argued that our findings are more elegantly accounted for within a model in which these two levels are strictly serially ordered, and in which attention automatically shifts between alternative outlets whenever it shifts between processing levels; an interactive model can, indeed, account for the same effects, but only via a modification which is ad hoc, and hence, we feel, undesirable.

Thus we would argue that the model which best accounts for the overall pattern of lexicality effects in phoneme monitoring is one in which the stages and sequence of comprehension processes do not change from situation to situation. Information flow is always bottom-up. Lexical processing can always only take place after prelexical processing. The target detection task required in phoneme monitoring can produce a response from a prelexical representation, and it can also produce a response based on information in the lexical entry. Note that, as we pointed out in the introduction, there is no inherent requirement that the prelexical representation upon which the target detection process operates contains labeled segments isomorphous with phonemes. Nor do we consider that the internal representation which the subject forms of the target is necessarily an abstract phonemic representation. Both such assumptions may seem superficially quite plausible and parsimonious, but no aspect of our model nor of our results actually requires them. Phonemic segmentation is something which speakers of languages using alphabetic writing

systems are trained to do as they learn to read; speakers of the same languages who have not been trained to read find phonemic segmentation extremely difficult (Morais, Cary, Alegria, & Bertelson, 1979). It is certainly not clear whether the perception of speech involves segmentation into phonemes; at least some language groups show clear evidence of segmenting speech into syllables (Cutler et al., 1983, 1986; Mehler et al., 1981). Foss and Gernsbacher (1983) assumed that results from phoneme-monitoring experiments address the nature of the normal prelexical access code; we feel, however, that the task does not directly speak to the nature of this code, for the simple reason that phoneme monitoring is not itself a component of normal language comprehension. As an extrinsic task it can thus be modeled as tapping into the comprehension process at any of several points. Inferences about where these points are can only be drawn on the basis of results which suggest that monitoring responses are affected by certain factors which are unambiguously associated with a given level of the comprehension process. In our model, one of the points at which phoneme monitoring taps into the comprehension process is at or above the lexical level, since RT may be affected by factors which can only be effective after lexical access has taken place (e.g., word-nonword differences); and one of the other points is a prelexical point, since RT may also be affected by such factors as closeness of match between target and stimulus. But all the factors which have been postulated as exercising prelexical effects would apply equally well to representations at a number of different levels—acoustic, phonetic, phonemic, syllabic, lexical, or many others; none of them is uniquely associated with a representation in terms of phonemes. The phoneme-monitoring task cannot be claimed to be a direct window onto normal comprehension processes. But it does allow us to draw strong inferences about the general framework within which such processes are most satisfactorily modeled.

APPENDIX: MATERIALS

Experiments 1 and 2

Words, tas, pis, dot, mat, doux; pote, pic, cap, gaffe, toc, tard, belle, but, poche, date; glas, clos, gras, creux, bru, pre, dru, trop, drap, tri

Nonwords. da, ti, co, na, gu; poc, pite, cac, gasse, tote, tal, berre, buc, posse, dac; cla, glo, cra, greu, pru, bre, tru, dro, tra, dri

Experiments 3 and 6

Words, butt, bid, bag, bite, bark, boat, bit, bus, beg, book, bud; dash, deaf, dim, dull, dog, dive, dam, deep, dove, dice, daze

Nonwords. bul, bef, boott, bup, buv, bip, bam, bipe, bart, boke, bim; dut, dipe, dake, dack, deg, dit, duss, dop, dite, dag, deech

Experiment   4

   Words, car, dye; bus, dull, dash, tack, beg, deaf, dog, pop, goat, poke, tell, peg, tip, bid; tree, draw, blue, plough, clay, glee, claw, glow, cry
   Nonwords. dar, kye; duss, bul, tash, dack, deg, bef, pog, dop, pote, goke, pell, teg, bip, tid; dree, traw, plue, blough, glay, clee, glaw, cloh, gry

Experiment 5: (As Experiment 3, plus:)

   Words, tack, tart, type, teach, toad, tape, take, tip, tell, tight; poach, pipe, put, pace, poke, peg, pop, pup
   Nonwords. teep, toach, tace, taze, tid, teg, tive, tash, tark, tice; pite, pookk, pape, pote, pell, pog, pud, poad

Experiment   7

   Words, bosse, pope, bague, pape, biche, pipe, bonne, pomme, bache, panne, dame, lache, dague, tape, digue, tic, due, tube, dur, tulle
   Nonwords. posse, bope, pague, bape, piche, bipe, ponne, bomme, pache, banne, tame, dache, tague, dape, tigue, die, tuc, dube, tur, dulle

   Note. Item means for all experiments are available from the first author on request.

REFERENCES

Carr, T. H., Davidson, B. J., & Hawkins, H. L. (1978). Perceptual flexibility in word recognition: Strategies affect orthographic computation but not lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 674-690.

Cutler, A., & Fodor, J. A. (1979). Semantic focus and sentence comprehension. *Cognition,* 7, 49-59.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1983). A language specific comprehension strategy. *Nature (London),* 304, 159-160.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language,* 25, 385-400.

Cutler, A., & Norris, D. (1979). Monitoring sentence comprehension. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett.* Hillsdale, NJ: Erlbaum.

Dell, G. S., & Newman, J. E. (1980). Detecting phonemes in fluent speech. *Journal of Verbal Learning and Verbal Behavior,* 19, 608-623.

Elman, J. L., & McClelland, J. L. (1984). Speech perception as a cognitive process: The interactive activation model. In N. J. Lass (Ed.), *Speech and language* (Vol. 10). New York: Academic Press.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language.* New York: McGraw-Hill.

Foss, D. J. (1969). Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning and Verbal Behavior,* 8, 457-462.

Foss, D. J., & Blank, M. A. (1980). Identifying the speech codes. *Cognitive Psychology,* 12, 1-31.

Foss, D. J., & Dowell, B. E. (1971). High-speed memory retrieval with auditorily presented stimuli. *Perception & Psychophysics,* 9, 465-468.

Foss, D. J., & Gernsbacher, M. A. (1983). Cracking the dual code: Toward a unitary model of phoneme identification. *Journal of Verbal Learning and Verbal Behavior,* 22, 609-632.

Foss, D. J., Harwood, D. A., & Blank, M. A. (1980). Deciphering decoding decisions: Data and devices. In R. A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, NJ: Erlbaum.

Foss, D. J., & Swinney, D. A. (1973). On the psychological reality of the phoneme: Perception, identification and consciousness. *Journal of Verbal Learning and Verbal Behavior,* 12, 246-257.

Goldstein, L. (1980). Bias and asymmetry in speech perception. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen and hand.* New York: Academic Press.

Healy, A., & Cutting, J. (1976). Units of speech perception: Phoneme and syllable. *Journal of Verbal Learning and Verbal Behavior,* 15, 73-83.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown Univ. Press.

Marslen-Wilson, W. D. (1980). Speech understanding as a psychological process. In J. C. Simon (Ed.), *Spoken language generation and understanding.* Dordrecht: Reidel.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology,* 18, 1-86.

McNeill, D., & Lindig, K. (1973). The perceptual reality of phonemes, syllables, words, and sentences. *Journal of Verbal Learning and Verbal Behavior,* 12, 431-461.

Mehler, J. (1986). *A comparison of English and French speech processing.* Paper presented to NATO Advanced Research Workshop, The Psychophysics of Speech Perception, Utrecht, The Netherlands.

Mehler, J., Dommergues, J.-Y, Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior,* 20, 298-305.

Miller, G. A., & Nicely, P. (1955). Analysis of perceptual confusions among English consonants. *Journal of the Acoustical Society of America,* 27, 338-352.

Mills, C. B. (1980a). Effects of context on reaction time to phonemes. *Journal of Verbal Learning and Verbal Behavior,* 19, 75-83.

Mills, C. B. (1980b). Effects of the match between listener expectancies and coarticulatory cues on the perception of speech. *Journal of Experimental Psychology: Human Perception and Performance,* 6, 528-535.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition,* 7, 323-331.

Morton, J., & Long, J. (1976). Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior,* 15, 43-51.

Newman, J. E., & Dell, G. S. (1978). The phonological nature of phoneme monitoring: A critique of some ambiguity studies. *Journal of Verbal Learning and Verbal Behavior,* 17, 359-374.

Norris, D. G. (1986). Word recognition: Context effects without priming. *Cognition,* 22, 93-136.

Pachella, R. G. (1974). The interpretation of reaction time in information-processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* Hillsdale, NJ: Erlbaum.

Peterson, G. E., & Lehiste, I. (1976). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America,* 15, 45-51.

Rubin, P., Turvey, M. T., & Van Gelder, P. (1976). Initial phonemes are detected faster in spoken words than in non-words. *Perception & Psychophysics,* 19, 394-398.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception, Pt. II: The contextual enhancement effect and some tests and extensions of the model. *Psychological Review,* 89, 60-94.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General,* 110, 474-494.

Samuel, A. G., & Ressler, W. H. (1986). Attention within auditory word perception: Insights from the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception and Performance,* 12, 70-79.

Segui, J., & Frauenfelder, U. (1986). The effect of lexical constraints upon speech perception. In F. Klix & H. Hagendorf (Eds.), *Human Memory and Cognitive Capabilities: Mechanisms and Performances.* Amsterdam: North-Holland.

Segui, J., Frauenfelder, U., & Mehler, J. (1981). Phoneme monitoring and lexical access. *British Journal of Psychology,* 72, 471-477.

Steinheiser, F. H., & Burrows, D. J. (1973). Chronometric analysis of speech perception. *Perception & Psychophysics,* 13, 426-430.

Stemberger, J. P., Elman, J. L., & Haden, P. (1985). Interference between phonemes during phoneme monitoring: Evidence for an interactive activation model of speech perception. *Journal of Experimental Psychology: Human Perception and Performance,* 11, 475-489.

Treisman, A., & Squire, R. (1974). Listening to speech at two levels at once. *Quarterly Journal of Experimental Psychology,* 26, 82-97.

Wang, M. D., & Bilger, R. C. (1973). Consonant confusions in noise: A study of perceptual features. *Journal of the Acoustical Society of America,* 54, 1248-1266.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science,* 167, 392-393.