

Vowel mutability and lexical selection in English: Evidence from a word reconstruction task

BRIT VAN OOIJEN

Laboratoire de Sciences Cognitives et Psycholinguistique, Paris, France

This study introduces a new paradigm for investigating lexical processing. First, an analysis of data from a series of word-spotting experiments is presented suggesting that listeners treat vowels as more mutable than consonants in auditory word recognition in English. In order to assess this hypothesis, a word reconstruction task was devised in which listeners were required to turn word-like nonwords into words by adapting the identity of either one vowel or one consonant. Listeners modified vowel identity more readily than consonant identity. Furthermore, incorrect responses more often involved a vowel change than a consonant change. These findings are compatible with the proposal that English listeners are equipped to deal with vowel variability by assuming that vowel identity is comparatively underdefined. The results are discussed in the light of theoretical accounts of speech processing.

In the continuing search to characterize the nature of spoken word recognition, several candidates have been put forward as likely sublexical building blocks for perceptual processing. For example, there is by now strong experimental evidence that the syllable plays an important role in the segmentation of continuous speech into discrete words. Language-specific findings have been documented for both French (see, e.g., Mehler, Dommergues, Frauenfelder, & Segui, 1981) and English (Cutler & Norris, 1988; also Cutler & Butterfield, 1990, 1992). Syllables are important, but they are unlikely to be the sole atomic units that feature in the varying requirements of everyday speech communication. There is also ample evidence for a phoneme-sized unit. For instance, data on single-segment misperceptions (see, e.g., Bond & Garnes, 1980; Browman, 1980; Garnes & Bond, 1980; MacKay, 1970; Motley, 1973), as well as the existence of language games such as pig Latin, puns, and rhymes, indicate that individual speech sounds play a role in speech perception. Accordingly, a phonemic level of representation has been incorporated into models of spoken word recognition such as the cohort model¹ (see, e.g., Marslen-Wilson, 1987, 1990; Marslen-Wilson & Welsh, 1978), TRACE (McClelland & Elman, 1986), and SHORTLIST (Norris, 1994).

All of these theoretical accounts of spoken word recognition ignore differences between types of phonemes, in particular, between the two major categories of consonants and vowels. Likewise, experimental studies investigating lexical access through manipulation of single segments have almost exclusively concentrated on consonants (e.g., Connine, Blasko, & Titone, 1993; Marslen-Wilson & Zwitserlood, 1989). For all of these psycholinguists, a segment is a segment. This uniformity is all the more remarkable in view of decades of phonetic research showing that phonemes differ widely on a variety of measures, with the well-known difference between vowels and consonants in categorical perception being only one example (see, e.g., Liberman, Mattingly, & Turvey, 1972; Liberman, Safford Harris, Hoffman, & Griffith, 1957; Studdert-Kennedy, Liberman, Harris, & Cooper, 1970). Given this, it would seem important to ask whether different types of phonemes are processed in the same or in different ways by adult listeners. And if this has consequences for lexical access, existing models would have to be revised, possibly quite fundamentally, in order to capture such a distinction. The aim of this study is to draw attention to an aspect of spoken word recognition that has not previously been considered and to present preliminary evidence indicating that vowels and consonants are indeed processed differently during the process of word recognition.

The distinction between vowels and consonants can be asserted on acoustic-phonetic, phonological, and psycholinguistic grounds. However, speech researchers are well aware that this distinction seems more clear-cut than it is. For instance, some phonemes, such as the semi-vowels /w/ and /j/, count as vocalic in that they are relatively steady state and periodic, but they count as consonantal in that they cannot function as syllabic nuclei. Some phonemes can lose one or indeed nearly all of their defining acoustic-phonetic properties when realized in running speech, as is the case with voiced plosives, which

This research was carried out while B.v.O. was at the MRC Applied Psychology Unit, Cambridge, United Kingdom; this research was financed by the Faculty of Letters, University of Leiden, The Netherlands. Further analyses and reporting of the findings took place at B.v.O.'s present address, with financial support from The Fyssen Foundation, Paris. B.v.O. wishes to express her thanks to James McQueen, Anne Cutler, and Dennis Norris for access to their data, constructive discussions, and important comments on this study. Critical reading by Vincent van Heuven, Jacques Mehler, Sieb Nooteboom, and an anonymous reviewer considerably improved an earlier version of the manuscript. Correspondence should be addressed to B. van Ooijen, Laboratoire de Sciences Cognitives et Psycholinguistique, 54, Boulevard Raspail, 75006 Paris, France (e-mail: brit@lscp.msh-paris.fr).

may lose their closure and/or their voicing and/or their release. Many examples can be found in the literature to show that phonemes as acoustic events are subject to variation. Strange, Jenkins, and Johnson (1983) showed that certain regions within the vowel space populated by tokens of a particular vowel type produced under varying conditions overlapped significantly with regions containing tokens of other vowel types. Variation in the acoustic properties of vowels as spoken by different speakers was investigated by Peterson and Barney (1952), who demonstrated that, for example, a sound intended by one speaker as an [ɛ] had the formant frequencies of another speaker's [ɪ]. To complicate matters, phonemes are not only variably produced, but they can also be variably perceived. Perceptual tests of vowel quality have demonstrated that the same test word was heard as *bit* by the majority of subjects when preceded by a sentence with an overall high *F1* pattern, but as *bet* when the preceding sentence had a relatively low *F1* pattern (Ladefoged, 1989; Ladefoged & Broadbent, 1952). Thus the same physical realization can correspond to different perceptual representations and hence to different abstract categories.

Of course, with 12 monophthongal and 8 diphthongal vowels, the English language is a particularly rich one in terms of vowel repertoire (Gimson, 1980; Maddieson, 1984); further, dialect variation in this language is mainly carried by vowels. Such sources of variety arguably complicate the task of discriminating between vowel identities, especially when a listener is required to respond promptly. This idea was confirmed in a series of studies using the phoneme-monitoring paradigm (Cutler, van Ooijen, Norris, & Sánchez-Casas, 1996; Norris, van Ooijen, & Cutler, 1992; van Ooijen, 1993, 1994; van Ooijen, Cutler, & Norris, 1991). In these studies, a persistent negative correlation appeared between reaction times (RTs) to vowel targets and physical duration: The longer a vowel, the faster it was detected. In the studies in which vowels and consonants were compared, such a relation was never found for consonantal targets. Taken together, these findings suggest that precise vowel identity can take time to establish itself. If this difference between vowels and consonants exists in phoneme monitoring, the question that presents itself is whether such differences may be indicative of a fundamental distinction in perceptual processing between vowels and consonants. This question, then, warrants further investigation with other perceptual tasks.

Paradoxically, vowels may be acoustically salient segments relative to consonants. They are the loudest, longest, and most periodic portions of the speech stream (Crystal & House, 1988a, 1988b; Fry, 1979; Umeda, 1975, 1977); further, analyses of spontaneous misperceptions have shown that vowels, and especially vowels in stressed syllables, are less often misperceived than are consonants (Bond & Garnes, 1980; Garnes & Bond, 1980). Nevertheless, the above-mentioned phoneme-monitoring evidence strongly indicates that *precise* vowel identity is not always readily recognizable, suggesting that salient is

not synonymous with *unambiguous*. This in turn may have far-reaching consequences for the process of accessing word forms in the mental lexicon: Is it possible that the contribution of vocalic information to the activation of lexical items is less constraining (i.e., it activates a greater range of lexical candidates) than does the contribution of consonantal information? Evidence supporting just such an interpretation appeared in a post hoc analysis of data reported by McQueen, Norris, and Cutler (1994).

EVIDENCE FROM A LEXICAL RECOGNITION TASK: WORD SPOTTING

McQueen, Norris, and Cutler (1994) designed experiments to assess the performance of the SHORTLIST model of spoken word recognition (Norris, 1994). Their materials consisted of isolated nonsense strings in some of which real English words were embedded, such as the word *sack* in /sækkrək/, and the word *mess* in /dɒmɛs/. Subjects were asked to listen to these and to press a response button if the nonsense string began or ended with a real word, and then to say aloud, into a microphone, the word that they had spotted. For the purpose of the following post hoc analysis, observations were made across the three experiments as well as across the experimental and filler items of McQueen et al.'s (1994) study.

Since the materials were recorded specifically for experimental purposes, the speech was of a high quality. Thus the nonsense strings can be considered as consisting of clear, unambiguous phonemic tokens that should not elicit large numbers of incorrect responses. However, false alarms—that is, nonintended words—were reported in these experiments, and such cases mainly involved one of three kinds of single-segment perceptual changes: (1) from schwa to a full vowel, as in the case of the stimulus /bɒdən/, which elicited the false alarms *bad*, *bud*, and *bed*, or /lɔːfəɪb/ with *laugh*; (2) from one full vowel to another, for example, the stimulus /səprɛm/ was reported to contain the word *pram*, and /bælkən/ with its false alarm *ball*; (3) from one consonant to another, as in /kɪzpt/, to which some subjects responded with *kiss*.

Upon closer inspection, the pattern of incorrectly perceived segments showed remarkable asymmetries. First, responses involving a vowel change vastly outnumbered those involving a consonant change: Out of a grand total of 690 incorrect responses involving single phoneme changes, 649 were for vowels and 41 were for consonants. Out of a grand total of 810 stimulus items, 269 elicited single vowel changes. Of these, 57 stimuli elicited more than one vowel type, resulting in several real-word response alternatives, such as the stimulus /hɪpnɔːt/, which elicited *note*, *not*, and *nut*; /pɒnuːv/, which elicited the false alarms *pen*, *pun*, and *pan*; /fɛlps/, which elicited *fill*, *full*, and *fell*; and /trɛdɪʃ/, which elicited *tray* and *trod*. In contrast, only 28 out of the 810 stimuli elicited a consonant change, and never did the same stimulus elicit more than one alternative consonant type. In addi-

tion, of the 41 consonant change responses, 30 involved one feature only, as in the case of /mɛzməd/, with its false alarm *mess*; and /ləfaɪb/, with its false alarm *vibe*.

It is important to note that this absence of more than one consonant change alternative is not due to there being no other word alternatives available for the stimulus. For example, for /mɛzməd/ there would have been *fez*, *says*, *men*, *mesh*, and *met*; and for /ləfaɪb/, one could have had *file*, *fine*, *fight*, and so on. But these alternatives were never given; instead, it would seem that consonant alternatives were limited to those involving only one feature difference. Yet even in cases where a one-feature consonant alternative was available, as in /dɛvɪŋ/, which could have been changed to *lid*, often a (one-or-more-feature) vowel change was given instead; in this example, /dɛvɪŋ/ elicited *dove*.

So, a strong tendency was revealed in the false alarms to report a real word that differed from the stimulus string with respect to one vowel. These changes ranged from small, as in /ɛ/ → /æ/, to completely different with respect to stress, quality, and duration, as in /ə/ → /ɑ:/. The relatively few consonant changes, on the other hand, mainly involved one feature only, suggesting these changes might have been based on occasional low-level acoustic confusions. Unfortunately, several factors combine to prevent these data being subjected to detailed statistical analysis. First, there are methodological confounds: Since phonemic substitution was not a variable of interest in McQueen et al.'s (1994) study, different subjects heard slightly different nonsense strings; one group of subjects was instructed to detect words at the start of nonsense strings only; another group spotted words at the end of nonsense strings only; and a third group was unrestricted in terms of where to locate the embedded targets. Second, there are asymmetries in the patterns of substitutions that admit of several explanations: The majority of vowel changes (566 out of the total of 649) were from schwa to a full vowel, rather than from one full vowel to another (83 out of 649), and it is not possible to establish whether this was due to the selection of the materials or to some other, phoneme-specific, factor.

Nevertheless, the pattern of false alarms in this study is striking and suggests two things that merit further investigation. First, English listeners in the process of recognizing words appear to process phoneme-sized segments (rather than, for instance, always entire syllables). Second, within these phonemes, vowel identity seems to count as more mutable than consonant identity. What underlies this peculiar vowel–consonant asymmetry in perceived phoneme identity?

The present experiment was designed to explore further the idea that vowel identity is perceived as more mutable than is consonant identity during word recognition. The most stringent test of this idea is one that contrasts various strong (i.e., primary or secondary stressed, unreduced) vowels with various consonants. If mutability effects can be found for unreduced vowels, the results from the McQueen et al. (1994) study are less likely to be attributable to the special status of the reduction vowel

schwa. The task that was designed for this purpose is described below.

EVIDENCE FROM A NEW TASK: WORD RECONSTRUCTION

To assess the relative flexibility of vowels and consonants, a task is needed that meets the following requirements. First, it must ensure that lexical access takes place, which excludes phonetic decision tasks such as phoneme monitoring, for which the involvement of a prelexical versus a lexical level of processing is fervently debated (see van Ooijen, 1994, for a review of this debate). Second, the task must allow for a direct comparison of vowels and consonants within the same experimental item. This means that mispronunciation detection cannot be employed, because any real word that has both a mispronounced vowel and a mispronounced consonant easily becomes an unrecognizable stimulus. For the same reason of direct comparison, the word-spotting task, in which a real, intact word is embedded in a larger sequence, is not suited to address the issue under investigation. Furthermore, since we are interested in acoustic mutability, we need to force listeners to mutate directly their acoustic percepts, preferably without intervention of visual (orthographic) stimulation such as that which takes place in the cross-modal priming task.

The task designed to meet all of the above constraints is here termed “word reconstruction.” In this task, listeners hear word-like nonwords, such as *eltimate*, and are asked to press a response key as soon as they can identify a real word that resembles the nonword stimulus; in this example, *ultimate* or *estimate*. In addition, subjects are asked to report verbally which word they have recognized, and in the case of more than one word being recognized, which one sprang to mind first. In this way, one and the same nonword stimulus serves as a perceptual template via which more than one real word can be recognized. Thus an important feature of this paradigm is that the stimuli can be chosen to be ambivalent, and hence the change that is required to produce either alternative can be controlled.

The word reconstruction task provides three dependent variables: RT, error rate, and proportion of vowel versus consonant responses. Three experimental conditions were constructed: a sound change condition in which subjects were unrestricted as to their choice of whether they changed one vowel or one consonant in the nonword, a vowel change condition in which listeners were instructed to make only single vowel changes, and a consonant change condition for consonant changes only. The predictions are as follows. If vowels are more mutable than consonants, then listeners in the first condition, who are free to choose, should be better at modifying vowels than consonants. This would be indicated by a higher proportion of vowel responses and/or faster RTs for these responses. In the second condition, listeners are restricted to vowel changes. According to the vowel mutability hypothesis, listeners should find this condition relatively

easy; this would be borne out by relatively fast RTs and/or high accuracy. In the third condition, listeners are restricted to consonant changes. If they treat vowels as the segments that potentially need adapting in terms of identity, then turning their attention to consonants should be relatively difficult. This should result in comparatively slow RTs and/or high error scores.

THE WORD RECONSTRUCTION EXPERIMENT

Method

Materials. Sixty nonwords were constructed as stimuli. Each of these could be turned into a real English word by changing one of its constituent vowels and into another word by changing one of its consonants, as in /'eltɪmət/ → *ultimate* and *estimate*. It proved impossible to construct a sufficient number of experimental nonwords in which, for each one, the possible vowel change word(s) had the same frequency of occurrence as the possible consonant change word(s). However, care was taken to ensure that the average frequency for the entire group of possible vowel change words versus that of the entire group of possible consonant change words (402 and 564, respectively) could be considered similar, in that both fell in the higher frequency range of >130 counts per million (Francis & Kučera, 1982). Furthermore, in half of the materials, the word that resulted from the vowel change (in the example above, *ultimate*) had a higher frequency of occurrence than the word that resulted from the consonant change (in the example above, *estimate*). In the other half the reverse was the case. Twenty-two nonwords had more than one possible vowel word, as in /'mɛtən/, which has the vowel words *mitten* and *mutton* versus the consonant word *melon*. Twenty-four other nonwords had more than one possible consonant word, as in /'ɛtɪk/, which has *epic* and *ethic*, versus *attic* as the only vowel word; the total number of these possible real-word alternatives was matched for the vowel versus the consonant words (102 vs. 114, respectively). In accordance with the lexical statistics of English (Cutler & Carter, 1987), most words had a strong-weak stress pattern; also, more words started with consonant-vowel (CV) than with VC. For any given nonword, the corresponding consonant change by necessity occurred in a different position within the word template than did the vowel change. For example, in the nonword /'tɪ:bəl/, which has a CVCVC template, the consonant /t/, which needs to be changed to form the real word *feeble*, occurs earlier than the vowel /i:/, which needs to be changed to form the corresponding vowel word *table*. In the nonword /'mɛləs/, however, this is the other way round: The vowel change required to form *malice* occurs earlier than does the consonant change that forms the real word *menace*. Vowel and consonant positions within the words were controlled as well as possible (in 35 of the 60 items, the consonant change occurred earlier than the vowel change; in the remaining 25, the vowel change preceded the consonant change). The uniqueness points of the vowel and consonant words were taken into account so that overall the number of vowel changes and consonant changes occurring before or after a uniqueness point was approximately the same. For the vowel changes, the distance between stimulus vowel and replacement vowel ranged from one to five features different, with an average of 2.49; for the consonant changes, it ranged from one to six features different, with an average of 2.97 (Jakobson, Fant, & Halle, 1951).

A further 70 nonwords served as filler items; 12 of these were used as practice items. The majority of the filler items were trisyllabic, and the phonemes that needed to be changed tended to occur toward the end of the nonword, for instance /'rɛkəgnɪz/ (*recognize*) and /'dæfədɪl/ (*daffodil*). There were two reasons for this. First, the purpose of having comparatively easy filler items was to maintain listeners' confidence and hence to increase the likelihood

of both frequent and fast responding. Second, it was hoped that the presence of a comparatively large number of phonemic changes located near word endings in fillers would discourage listeners from adopting a rhyming strategy for the experimental items, which could have biased the number of initial consonant changes.

The materials were recorded by a male native speaker of British English onto the left channel of a digital audiotape. To ensure an acoustic realization of each nonword that was as close as possible to both its alternatives, the speaker first pronounced both the vowel and the consonant real-word alternative, before each stimulus nonword, for example: "*table, feeble, teeble*." The nonwords were then spliced onto an experimental tape. On the right channel of this tape, timing pulses were aligned with each nonword onset. All experimental nonwords with their corresponding vowel and consonant real words are listed in the Appendix.

Design. The 60 nonwords were divided into three groups of 20. In each of these, 10 items had vowel words with a higher frequency than consonant words and 10 had consonant words with a higher frequency than vowel words. These groups were rotated across the three phoneme change conditions, resulting in a within-subject design. Thus all subjects heard all of the materials, but 10 heard a given set of 20 items in the vowel change condition, 10 were given the same set in the consonant change condition, and 10 heard it in the sound change condition.

Subjects. Thirty members of Cambridge University and of the Subject Panel of the Applied Psychology Unit in Cambridge, ranging in age from 18 to 35 years, received cash payment for participating in the experiment.

Procedure. Subjects were tested individually in a quiet room; they listened through closed headphones (Sennheiser HD-520) to the left channel of the tape, presented binaurally. Taped instructions informed them that they would be hearing nonwords. A *nonword* was defined as "a string of sounds which could belong to the English language, but does not, such as *namp*, *rostive*, and *jistle*." Subjects were asked to listen carefully to each nonword, to press a response key as soon as they had thought of a real-word substitution, and then to say this word. Depending on which condition the listeners were in, they were first asked to change only one consonant, only one vowel, or any one sound in the nonword. They were not made aware beforehand that they would be asked to switch their attention from one phonemic category to another after each third of the experiment. In the consonant change and the vowel change conditions, listeners were familiarized with their task by hearing examples of a consonant change and a vowel change, respectively. In the sound change condition, listeners heard examples of both a vowel change and a consonant change, but they were not made aware of this distinction. Listeners were informed that there could be more than one possible acceptable word for a given nonword, and that they had to say whichever word sprang to mind first. Finally, all participants were urged to concentrate on how the nonwords sounded, rather than on how they might be spelled.

Order of presentation of the three conditions was counterbalanced so that all listeners heard all nonwords in the same order, but 10 started with consonant change instructions, 10 with vowel change instructions, and 10 with sound change instructions. Stimulus presentation was controlled by the experimenter with a timeout of 10 sec; a small pilot study had served to establish this as a comfortable interstimulus interval. RT data collection and storage were under the control of a microcomputer running the TSCOP RT software (Norris, 1984). In addition, subjects' verbal responses were taped with a Marantz CP-430 audiocassette recorder and a Realistic 33-1060 electret condenser microphone. In all, the experiment took about 30 min.

Results

The data were analyzed as follows. First, listeners' verbal responses were scored and inappropriate RTs were

discarded. These included intrusions (vowel responses given in the consonant change condition or vice versa), mistakes (responses that involved a substitution of more than one phoneme), and false alarms.² Standard analyses of variance with subjects and with words as random factors were carried out on the remaining RTs (measured from the onset of each nonword) and on the error scores for all three phoneme change conditions. In addition, the verbal responses were used for a combined analysis of the vowel change and the consonant change conditions with a view to the number of intrusions in each. Finally, the sound change condition was analyzed on its own in order to compare the proportion as well as the RTs of the consonant versus the vowel responses.

Reaction times. Table 1 shows the mean RTs (in milliseconds) with their corresponding error rates for the three phoneme change conditions. There was a significant main effect of phoneme change condition [$F_1(2,54) = 13.39, p < .001$; $F_2(2,110) = 12.99, p < .001$]. T tests showed that the RT difference of 459 msec between the sound change and the vowel change condition was significant [$t_1(29) = 5.02, p < .001$; $t_2(56) = 4.16, p < .001$]. As expected, the difference of 654 msec between the sound change and the consonant change conditions was also significant [$t_1(29) = 4.4, p < .001$; $t_2(56) = 5.47, p < .001$]. The RT difference of 195 msec between the vowel change condition and the consonant change condition was in the predicted direction but did not reach significance. There was no effect of the order in which the three phoneme change conditions had been presented.

Error rates. The consonant change condition not only had the slowest RTs, but also the highest number of incorrect or missing responses; that is, there was no speed-accuracy tradeoff. One-tailed Wilcoxon signed ranks tests on the error scores showed a significant difference of 26% between the consonant change condition and the sound change condition ($z_1 = 4.63, p < .001$; $z_2 = 4.99, p < .001$) as well as a significant difference of 14% between the consonant change condition and the vowel change condition ($z_1 = 3.58, p < .001$; $z_2 = 2.25, p < .02$). The 12% difference between the vowel change condition and the sound change condition was significant by subjects only ($z_1 = 3.90, p < .001$). There was no effect of the order in which the three conditions had been presented.

Phonemic intrusions. Table 2 shows the percentage of vowel intrusions in the consonant change condition (i.e., instances where listeners said, e.g., the vowel word *ultimate* instead of the consonant word *estimate* in response to the stimulus /'eltɪmət/) and of consonant intrusions in the vowel change condition. One-tailed Wilcoxon

Table 2
Percentages of Consonant Intrusions in the Vowel Change Condition and Vowel Intrusions in the Consonant Change Condition

Intrusion	Change Condition	
	Vowel	Consonant
Consonant	7.2	
Vowel		14.5

signed ranks tests showed that the difference of 7.3% between these proportions was significant ($z_1 = 3.36, p < .001$; $z_2 = 2.44, p < .01$). Thus listeners reported both vowel and consonant changes in the conditions where they had been explicitly asked to limit themselves to one category only, but these spontaneous slip-ups were more frequently inappropriate vowel changes in the consonant change condition. This is in fact in accord with informal remarks made by listeners during the course of the experiment. Many reported while being under instructions to change consonants that "vowels kept coming into their head." For both the vowel and the consonant changes, in about half of the cases (49%), the intrusion was higher in frequency than the correct response alternative; in 51% of the cases, the intrusion was lower in frequency. Thus the spontaneous slip-ups were not due to a frequency bias.

Consonant versus vowel responses. Let us now turn to a closer inspection of the sound change condition, in which listeners were unrestricted as to their choice of phonemic category. Table 3 shows the mean RTs (in milliseconds) of the vowel change responses and the consonant change responses in this condition. The vowel change responses were made faster than the consonant change responses, and this difference of 327 msec was significant by subjects³ [$F_1(1,27) = 5.34, p < .03$]. Also, for those stimuli that had a vowel change word with a higher frequency than the corresponding consonant change word, vowel change responses were more often chosen; the same was true for the stimuli that had a consonant change word with a higher frequency than the corresponding vowel change word: In these cases, consonant change responses prevailed. This frequency effect was significant by subjects [vowel change words higher frequency, $F_1(1,27) = 20.59, p < .01$; consonant change higher frequency, $F_1(1,27) = 10.60, p < .01$] but not by items, indicating that the effect was attributable to a few extremely highly frequent vowel and consonant change words. One-tailed Wilcoxon signed ranks tests on the proportion of vowel versus consonant responses in this condition showed no significant difference (vowels, 43%; consonants, 57%; $z_1 = 0.55$; $z_2 = 0.83$). That is, listeners were equally likely to opt for a vowel change as for a consonant change. Taken together, these findings indicate that both vowel words and consonant words were recognized on the basis of the nonword stimuli, but vowel words were easier to recognize than consonant words.

Control analyses. One possibility is that vowel changes were made faster than were consonant changes because the perceptual distance between a phoneme in a non-

Table 1
Mean Response Times (RTs, in Milliseconds) and Error Rates for the Three Phoneme Change Conditions: Sound Change, Vowel Change, and Consonant Change

Change Condition	RT	Error Rate
Sound	1,758	16%
Vowel	2,217	28%
Consonant	2,412	42%

Table 3
Mean Response Times (RTs, in Milliseconds) of the Vowel Change Responses and the Consonant Change Responses in the Sound Change Condition

RT	Responses	
	Vowel Change	Consonant Change
	1,595	1,922

sense word and its replacement phoneme in the appropriate real word was generally smaller for vowels than for consonants. Distance between phonemes can be expressed in terms of features, and this approach has been assumed to be perceptually valid (see, e.g., Cole, 1973; Cole & Scott, 1972; Miller & Nicely, 1955; Mohr & Wang, 1968; but see Wang & Bilger, 1973, for a critical review). Since the issue under investigation is a comparison between vowels and consonants, the use of features is problematic. Due to the inherent acoustic-phonetic differences between vowels and consonants, the features most suitable to describe one category are not the same as those that most aptly represent the other category. With this in mind, however, the following analyses were conducted.

Inspection of the data showed that 19 out of the 60 nonword stimuli yielded both a unique vowel response and a unique consonant response. For example, *metton* elicited both *mitten* and *melon* responses, but no others, such as, in this case, *mutton*. For each of these 19 nonwords, the corresponding vowel and consonant words were scored in terms of distance from the vowel or consonant in the stimulus. This was done with two distinctive feature systems, 1 and 2 (Tables 4A–4B and 5A–5B). Feature System 1 (Jakobson et al., 1951) uses different features for vowels and consonants and can be considered the most appropriate one for each category on its own. In addition, an attempt was made to equate vowels and consonants at least in terms of feature definitions, using a matrix with the same features for both, called Feature System 2 (Jakobson & Halle, 1956). For example, the stimulus /'wɛdəʊ/ had a vowel change from /ɛ/ to /ɪ/ (for *widow*), which corresponds to a distance of 1 on Feature System 1, and a consonant change from /w/ to /m/ (for *meadow*), corresponding to a distance of 2 on the same feature system.

For Feature System 1, the mean distance between stimulus phoneme and response phoneme was somewhat larger for the vowels (1.95) than for the consonants (1.53). A two-tailed Wilcoxon signed ranks test showed that this difference was close to significance ($p < .07$). For System 2, the mean distance was slightly larger for consonants (2.16) than for vowels (1.84), but not significantly so. To see whether RTs varied as a function of feature distance (in the sense that faster responses may go with smaller differences), linear regression analyses were performed on the factors of RT and number of features different. These yielded no significant correlations, irrespective of feature system. Thus, insofar as it is at all justified to equate distance between vowels with distance between consonants, the distance between stimulus phoneme and response phoneme was, if anything, larger for vowels than

for consonants. This rules out the possibility that vowels were easier to change because of a smaller distance between stimulus vowel and response vowel than between stimulus consonant and response consonant. Note that such an argument also seems unlikely in view of the observed asymmetry in the data, with vowel intrusions occurring significantly more often than consonant intrusions. Indeed, some of these vowel intrusions (18%) involved a larger feature distance between stimulus and response vowel than the corresponding, correct consonant response would have (e.g., *cammon* → *cōmmon* instead of *cannon*).

A final control analysis was motivated by the observation that the stimuli in this experiment contained more consonantal than vocalic positions. Of the 60 nonwords, only 3 had more vocalic than consonantal positions (e.g., /'ɛzi:/ = VCV); 19 had an equal number (e.g., /'nɛni:/ = CVCV), and a majority of 38 had more consonantal than vocalic positions (e.g., /'lɒvəl/ = CVCVC). This imbalance in the materials directly reflects the characteristics of English: There are more words starting with CV than with VC, so vocalic nuclei tend to be surrounded by consonants, and therefore the average word will have more consonants than vowels.⁴ One might argue that, since the subjects' task may be interpreted as "spot the odd sound out and correct it," it would not have taken subjects long to discover that, per item, a vowel stood a proportionally higher chance of being a candidate for replacement. If this were true, the proportion of vowel versus consonant responses obtained in the condition where listeners could choose freely should be a function of the proportion of vowel versus consonant positions per item: the fewer vowels in the stimulus, the more often a vowel response will be found. In order to check this, all experimental items were scored in terms of these proportions.

Table 4A
Distinctive Feature Matrix for Feature System 1 (Vowels)

Vowels	Vowels												
	i:	ɪ	ɛ	æ	ʌ	ɑ:	ɔ	ɔ:	ɜ:	u:	eɪ	aɪ	oʊ
Round	-	-	-	-	-	-	+	+	-	+	-	-	+
Low	-	-	-	+	-	+	-	+	-	-	-	+	-
High	+	+	-	-	-	-	-	-	-	+	-	-	-
Back	-	-	-	-	+	+	+	+	+	+	-	-	+
Tense	+	-	-	-	-	+	-	+	+	+	+	+	+
Diphthong	-	-	-	-	-	-	-	-	-	-	+	+	+

Note—After Jakobson, Fant, and Halle (1951).

Table 4B
Distinctive Feature Matrix for Feature System 1 (Consonants)

Consonants	Consonants													
	p	b	t	d	k	g	s	f	v	m	n	l	r	w
Syllabic	-	-	-	-	-	-	-	-	-	+	+	+	+	-
Nasal	-	-	-	-	-	-	-	-	-	+	+	-	-	-
Voiced	-	+	+	+	-	+	-	-	+	+	+	+	+	+
Continuant	-	-	-	-	-	-	+	+	+	+	+	+	+	+
Anterior	+	+	+	+	-	-	-	+	+	+	+	+	-	+
Coronal	-	-	+	+	-	-	+	-	-	-	+	+	+	-
Sonorant	-	-	-	-	-	-	-	-	-	+	+	+	+	+

Note—After Jakobson, Fant, and Halle (1951).

Table 5A
Distinctive Feature Matrix for Feature System 2 (Vowels)

Vowels	Vowels												
	i:	ɪ	ɛ	æ	ʌ	ɑ:	ɔ	ɔ:	ɜ:	u:	eɪ	aɪ	oʊ
Consonantal	-	-	-	-	-	-	-	-	-	-	-	-	-
Vocalic	+	+	+	+	+	+	+	+	+	+	+	+	+
Diffuse	+	+	-	-	-	-	-	-	-	+	-	-	-
Compact	-	-	-	+	-	+	+	+	-	-	-	+	-
Grave	-	-	-	-	+	+	+	+	+	+	-	-	+
Flat	-	-	-	-	-	-	+	+	-	+	-	-	+
Tense	+	-	-	-	-	+	-	+	+	+	+	+	+
Voice	+	+	+	+	+	+	+	+	+	+	+	+	+
Continuant	+	+	+	+	+	+	+	+	+	+	+	+	+
Strident	-	-	-	-	-	-	-	-	-	-	-	-	-
Nasal	-	-	-	-	-	-	-	-	-	-	-	-	-

Note—After Jakobson and Halle (1956).

Table 5B
Distinctive Feature Matrix for Feature System 2 (Consonants)

Consonants	Consonants													
	p	b	t	d	k	g	s	f	v	m	n	l	r	w
Consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	-
Vocalic	-	-	-	-	-	-	-	-	-	-	-	-	+	+
Diffuse	+	+	+	+	-	-	+	+	+	+	+	+	+	-
Compact	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Grave	+	+	-	-	+	+	-	+	+	+	-	-	-	+
Flat	-	-	-	-	-	-	-	-	-	-	-	-	-	+
Voice	-	+	-	+	-	+	-	-	+	+	+	+	+	+
Continuant	-	-	-	-	-	-	+	+	+	-	-	+	+	+
Strident	-	-	-	-	-	-	+	+	+	-	-	-	-	-
Nasal	-	-	-	-	-	-	-	-	-	+	+	-	-	-

Note—After Jakobson and Halle (1956).

For example, the stimulus /'mɛləs/ had a vowel-consonant ratio of 2/5 = 0.4, and the ratio of vowel-consonant responses for this item was 6/9 = 0.66. A correlational analysis was then conducted on the factors of vowel-consonant ratio and vowel-consonant responses per item. This yielded no significant correlation [$r(58) = .13, p < .3$]. Thus the higher number of vowel responses is not an artifact of there being fewer vowel positions than consonant positions in the majority of the stimuli.

DISCUSSION

This study has raised the issue of whether vowels and consonants differ with respect to how they are processed in spoken word recognition. For this purpose, data from three existing word-spotting experiments (McQueen et al., 1994) were reconsidered. In these experiments, listeners' false alarms revealed a predominance of vowels over consonants, in that (mostly weak) vowels were the most likely segments to be changed for word recognition (e.g., /səprɛm/ → *pram*). Subsequently, a word reconstruction experiment was designed to test the hypothesis that English listeners will assume vowel identity to be more mutable than consonant identity. The results of this experiment showed that (strong) vowels were substituted more readily than were consonants to turn word-like nonwords into real words. Even when listeners were explicitly instructed to focus on consonants, responses

based on vowel changes interfered; the opposite—that is, of consonant changes interfering with instructions to change vowels—occurred significantly less often. Taken together, these results support the hypothesis. The finding that unwarranted vowel changes outnumbered unwarranted consonant changes could be taken as an indication that listeners are so accustomed to vowels being mutable that they “automatically” come up with vowel alternatives.

The majority of the vowels that were changed in this experiment were strong, unreduced vowels. This suggests that the pattern of false alarms in the McQueen et al. (1994) study cannot be explained in terms of the schwa being a special case of mutable vowel and therefore uniquely liable to perceptual modification. Rather, perceptual flexibility would seem to be an aspect of vowel perception that applies to English vowels in general, irrespective of their rhythmic status. The finding that a tense vowel such as /i:/, one of the most extreme vowels with respect to articulatory position, was readily changed to /eɪ/ (*teeble* → *table*) suggests that vowel flexibility does not result from sheer acoustic confusability (which was in any case not expected, because of the quality of the speech and the quiet listening conditions). Thus the observed ease of vowel changes seems not to result from low-level perceptual confusions.

Nor can the observed ease of vowel changes be explained in terms of the number of vowels versus consonants in the language. One might say that there are fewer vowels than consonants in English, so the set size is smaller and therefore easier to “run through.” However, first of all, with 20 vowels (including diphthongs) and 24 consonants, English does not in fact display a large difference between the two sets of phoneme types. In addition, not all consonants occur in all word positions; for example, /ŋ/ never occurs initially, and neither /h/ nor /r/ occurs finally in British English (Gimson, 1980). Furthermore, in the condition where listeners were free to choose, the proportion of vowel versus consonant responses was similar. If smallness of set size per se prevailed, one would have expected a predominance of vowel choices. Such a predominance did show up in the RTs for the vowel responses, which were significantly faster than those for the consonant responses. However, both vowel and consonant RTs *separately* were faster in this condition than in the two restricted conditions. If set size alone determined speed of response, there would be no reason for subjects to be faster from one condition to the next, since set size of the phoneme types never changed. The fact that the sound change condition yielded the fastest RTs indicates that having plenty of choice does not slow down overall decision time nor decision time for the two categories separately.

What, then, is the most likely explanation for the findings in this study? It is proposed that the observed mutability of vowels is indicative of a mechanism for dealing with expected uncertainty about precise vowel identity. The need for such a mechanism may be motivated by the variability of acoustic segments, both with respect to pro-

duction (due to such factors as speaker or dialect variation, speech rate, etc.) and with respect to perception (due to, for instance, phonemic inventory, speech quality, etc.). In view of all of these sources of uncertainty, listeners appear to be prepared to postpone their decision of exactly which vowel type they heard. For lexical access, the consequence of a mechanism that treats vowel information as less fixed than consonant information is that vowel information as specified in the acoustic–phonetic input can vary considerably from vowel identity in a given lexical entry.

The issue of how good a match must be between acoustic–phonetic input and lexical knowledge has been investigated in several studies with a variety of techniques, such as cross-modal priming, gating, phonetic categorization, and phoneme monitoring. For example, Marslen-Wilson and Zwitserlood (1989) have argued on the basis of findings obtained with the cross-modal priming task that a word-initial, consonantal mismatch of only one feature (e.g., *bleasant* instead of *pleasant*) is enough for the sequence not “to enter normally into the on-line decision space” (p. 583); that is, not to speed up lexical decision responses to a later occurring associated target word (see also Marslen-Wilson, 1993).⁵ Connine et al. (1993) showed that lexical decisions to a word such as *service* did undergo facilitation from an earlier presented nonword *zervice*, but not from a form like *gervice*, suggesting that permissibility of mismatch is a matter of feature distance. In a subsequent study investigating the consequences of phonemic ambiguity, Connine, Blasko, and Wang (1994) found not only that an ambiguous word-initial consonant nevertheless permitted lexical activation, but also that it led to multiple lexical hypotheses. That is, after presentation of an ambiguous token such as *d/time*, both associated words *penny* and *clock* were activated. The authors concluded that “ambiguity that results in competing specifications of a phonetic feature does not preclude the mapping of either feature to a compatible lexical representation” (p. 633).

It remains to be determined precisely what degree of mismatch the recognition system can tolerate, and under what circumstances. However, it seems that the most substantial deficiency in our knowledge at present is that all of this research has focused on consonants. There is an obvious historical reason for this: The cohort model of word recognition has always stressed the sequential nature of the speech input, with the beginning of an acoustic–phonetic sequence determining the word-initial cohort (see, e.g., Marslen-Wilson, 1987, 1990; Marslen-Wilson & Welsh, 1978). In English, the typical word-initial structure is CV; therefore it is logical to concentrate on degree of word-initial (consonantal) mismatch. Consequently, information on degree of vocalic mismatch has to be inferred from what we know about consonantal segments. However, the evidence presented in this study suggests that such an inference is by no means warranted. Instead, it may well be the case that the recognition system is designed in such a way as to allow a greater amount

of vocalic mismatch than consonantal mismatch. Clearly, further research is required to address this possibility.

What might be the implications of comparative vowel mutability for theoretical accounts of speech processing? In order to account for the data in the present study, a model would need to incorporate a mechanism for accomplishing lexical selection despite a potentially considerable mismatch between input and lexical entry. In addition, performance would have to vary according to whether this mismatch resided in a vocalic or in a consonantal segment. Vocalic segments could have a comparatively flexible goodness-of-fit and hence not restrict the set of likely word candidates in the same way as consonantal segments, and this relative freedom would lead more rapidly to retrieval of a lexical item. There are models that would seem capable, in principle, of incorporating such a vowel–consonant distinction. In the SHORTLIST model (Norris, 1994), for example, word recognition is achieved in two stages. In the first, all potential lexical candidates starting at every phoneme in the input are “shortlisted” into a small interactive activation network. Subsequently, these candidates compete for recognition. The initial activation level of a candidate is determined by the degree to which it matches the acoustic–phonetic input. Given a certain tolerance of phonetic mismatch in the first stage, candidates that differ slightly from the input with respect to vowel identity would nevertheless be shortlisted. In this way, the relative perceptual flexibility for vowels as exhibited by listeners in the two experiments discussed would be captured.

CONCLUSION

If the acoustic–phonetic input does not map immediately onto a lexical entry, British English listeners relax their constraints on precise vowel identity, suggesting that they possess a perceptual mechanism for dealing with expected vowel variability that operates in the case of uncertainty about a lexical candidate. This finding cannot be adequately explained by such factors as word frequency, full versus reduced vowel quality, phonological distance as measured by distinctive features, distribution of vowels and consonants across words, or total number of vowels and consonants in the language’s repertoire. Nor is the observed effect of vowel mutability an artifact of the specific experimental task. For instance, a possible argument would be that the word reconstruction paradigm focuses attention on individual segments and therefore encourages awareness of single phonemes in an unnatural fashion. However, the effect was first observed in McQueen et al.’s (1994) word-spotting task. In this task, subjects’ attention was by no means drawn to single phonemes, nor to the possibility of the acoustic input being in need of modification—the task was simply to identify an embedded word of the language. Thus the observations of the word-spotting study together with the data from the word reconstruction experiment converge to suggest that, at least in British English, listeners

rely on a perceptual mechanism by which acoustic information about vowels is treated differently from acoustic information about consonants: In the case of uncertainty, vowel identity is assumed to be more mutable. Furthermore, the mutability of vowels indicates that comparatively salient segments are not necessarily comparatively easy to identify. In spontaneous misperceptions, which typically appear in noisy circumstances, consonants are more likely to suffer from distortion. This is not surprising: They are comparatively weak and short sounds, and as such are harder to pick up if the signal is not clear. If, on the other hand, the signal is clear enough to hear all of the segments, but does not produce an immediately interpretable percept, vowels are treated as the more mutable segments. Thus it is hoped that this study will serve as a starting point for the proposal that there is a difference between vowels versus consonants as lexical building blocks in spoken word recognition.

Finally, this study has demonstrated that the task of word reconstruction usefully extends the repertoire of paradigms available to investigate the process of spoken word recognition. It can be considered as an approximation of what listeners have to do in everyday speech communication, namely, to match up a less than perfect signal with abstract lexical representations. Several lines of research exploiting this task suggest themselves. For instance, the question of whether vowel mutability is related to vowel richness of a particular language seems important; in the present study, these two factors are confounded in that only one vowel-rich language has so far been tested. Among the many things that remain to be determined is the question of whether vowels are intrinsically mutable or mutable under the influence of number of neighbors in a language's phonemic repertoire.

REFERENCES

- BOND, Z. S., & GARNES, S. (1980). Misperceptions of fluent speech. In R. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, NJ: Erlbaum.
- BROWMAN, C. P. (1980). Perceptual processing: Evidence from slips of the ear. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen and hand* (pp. 213-230). New York: Academic Press.
- COLE, R. A. (1973). Listening for mispronunciations: A measure of what we hear during speech. *Perception & Psychophysics*, **13**, 153-156.
- COLE, R. A., & SCOTT, B. (1972). Distinctive feature control of decision time: Same-different judgments of simultaneously heard phonemes. *Perception & Psychophysics*, **12**, 91-94.
- CONNINE, C. M., BLASKO, D. G., & TITONE, D. (1993). Do the beginnings of words have a special status in auditory word recognition? *Journal of Memory & Language*, **32**, 193-210.
- CONNINE, C. M., BLASKO, D. G., & WANG, J. (1994). Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence in context. *Perception & Psychophysics*, **56**, 624-636.
- CRYSTAL, T. H., & HOUSE, A. S. (1988a). Segmental durations in connected-speech signals: Current results. *Journal of the Acoustical Society of America*, **83**, 1553-1573.
- CRYSTAL, T. H., & HOUSE, A. S. (1988b). Segmental durations in connected-speech signals: Syllabic stress. *Journal of the Acoustical Society of America*, **83**, 1574-1585.
- CUTLER, A., & BUTTERFIELD, S. (1990). Durational cues to word boundaries in clear speech. *Speech Communication*, **9**, 485-495.
- CUTLER, A., & BUTTERFIELD, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory & Language*, **31**, 218-236.
- CUTLER, A., & CARTER, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, **2**, 133-142.
- CUTLER, A., & NORRIS, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 113-121.
- CUTLER, A., VAN OOIJEN, B., NORRIS, D., & SÁNCHEZ-CASAS, R. (1996). Speeded detection of vowels: A cross-linguistic study. *Perception & Psychophysics*, **58**, 807-822.
- FRANCIS, W. N., & KUČERA, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- FRY, D. B. (1947). The frequency of occurrence of speech sounds in Southern English. *Archives Néerlandaises de Phonétique Expérimentale*, **10**.
- FRY, D. B. (1979). *The physics of speech*. Cambridge: Cambridge University Press.
- GARNES, S., & BOND, Z. S. (1980). A slip of the ear: A snip of the ear? A slip of the year? In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen and hand* (pp. 231-239). New York: Academic Press.
- GIMSON, A. C. (1980). *An introduction to the pronunciation of English*. London: Edward Arnold.
- JAKOBSON, R., FANT, G., & HALLE, M. (1951). *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT Press.
- JAKOBSON, R., & HALLE, M. (1956). *Fundamentals of language*. The Hague: Mouton.
- LADEFOGED, P. (1989). A note on 'information conveyed by vowels.' *Journal of the Acoustical Society of America*, **85**, 2223-2224.
- LADEFOGED, P., & BROADBENT, D. E. (1952). Information conveyed by vowels. *Journal of the Acoustical Society of America*, **24**, 629-637.
- LAHIRI, A., & MARSLÉN-WILSON, W. D. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, **38**, 245-294.
- LAHIRI, A., & MARSLÉN-WILSON, W. D. (1992). Lexical processing and phonological representation. In D. R. Ladd & G. J. Docherty (Eds.), *Second conference on laboratory phonology*. Cambridge: Cambridge University Press.
- LIBERMAN, A. M., MATTINGLY, I. G., & TURVEY, M. T. (1972). Language codes and memory codes. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 307-333). Washington, DC: V. H. Winston.
- LIBERMAN, A. M., SAFFORD HARRIS, K., HOFFMAN, H. S., & GRIFFITH, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, **54**, 358-368.
- MACKAY, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia*, **8**, 323-350.
- MADDIESON, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
- MARSLÉN-WILSON, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, **25**, 71-102.
- MARSLÉN-WILSON, W. D. (1990). Activation, competition, and frequency in lexical access. In G. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148-172). Cambridge, MA: MIT Press.
- MARSLÉN-WILSON, W. D. (1993). Issues of process and representation in lexical access. In G. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The Second Sperlonga Meeting* (pp. 187-210). Hillsdale, NJ: Erlbaum.
- MARSLÉN-WILSON, W. D., & WARREN, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes and features. *Psychological Review*, **101**, 653-675.
- MARSLÉN-WILSON, W. D., & WELSH, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, **10**, 29-63.
- MARSLÉN-WILSON, W. D., & ZWITZERLOOD, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 576-585.

- McCLELLAND, J. L., & ELMAN, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1-86.
- MCQUEEN, J. M., NORRIS, D. N., & CUTLER, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 621-638.
- MEHLER, J., DOMMERGUES, J.-Y., FRAUENFELDER, U., & SEGUI, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning & 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.
- MOHR, B., & WANG, W. S.-Y. (1968). Perceptual distance and the specification of phonological features. *Phonetica*, **18**, 31-45.
- MOTLEY, M. T. (1973). An analysis of spoonerisms as psycholinguistic phenomena. *Speech Monographs*, **40**, 66-71.
- NORRIS, D. (1984). A computer-based programmable tachistoscope for nonprogrammers. *Behavior Research Methods, Instruments, & Computers*, **16**, 25-27.
- NORRIS, D. (1994). SHORTLIST: A connectionist model of continuous speech recognition. *Cognition*, **52**, 189-234.
- NORRIS, D., VAN OOIJEN, B., & CUTLER, A. (1992). Speeded detection of vowels and steady-state consonants. In *Proceedings of ICSLP '92* (pp. 1055-1058).
- PETERSON, G. E., & BARNEY, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, **24**, 175-184.
- STRANGE, W., JENKINS, J. J., & JOHNSON, TH. L. (1983). Dynamic specification of coarticulated vowels. *Journal of the Acoustical Society of America*, **74**, 695-705.
- STUDDERT-KENNEDY, M., LIBERMAN, A. M., HARRIS, K., & COOPER, F. S. (1970). The motor theory of speech perception: A reply to Lane's critical review. *Psychological Review*, **77**, 234-249.
- UMEDA, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America*, **58**, 434-445.
- UMEDA, N. (1977). Consonant duration in American English. *Journal of the Acoustical Society of America*, **61**, 846-858.
- VAN OOIJEN, B. (1993). Detection of vowels and consonants by humans: Effects of minimising auditory memory load. In *Proceedings of Euro-speech '93* (Vol. 2, pp. 1507-1510).
- VAN OOIJEN, B. (1994). *The processing of vowels and consonants*. Doctoral dissertation, University of Leiden, The Netherlands.
- VAN OOIJEN, B., CUTLER, A., & NORRIS, D. (1991). Detection times for vowels versus consonants. In *Proceedings of Eurospeech '91* (Vol. 3, pp. 1451-1454).
- 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.

NOTES

1. However, in more recent work, Marslen-Wilson and Warren (1994) have argued in favor of a sublexical featural level. As speech is heard, some set of features is extracted and mapped onto the level of lexical representations. On this view, phonetic decision is a strictly postlexical task.
2. False alarms were spontaneous keypresses followed by subjects either falling silent or alerting the experimenter to the fact that they had made a mistake. The majority of these false alarms occurred in the consonant change condition; as subjects testified afterward, these were cases where they had immediately thought of a real word but realized too late that it had a different vowel rather than a different consonant.
3. Significance by items was probably not reached owing to restrictions imposed by the nature of the analysis. In order for a meaningful comparison to be made, for each item both vowel RTs and consonant RTs must be contrasted. For 18 out of the 60 items, however, only one response category (mostly the vowel alternative) was used. After removal of these items, the remaining number was not significant [$F_2(1,41) = 1.69, p < .20$]. This effect could perhaps have been significant over a larger item set.
4. Correspondingly, in a study of the frequency of occurrence of speech sounds, Fry (1947) estimated that in running English speech, vowels account for about 40% and consonants for about 60% of the phonemes uttered.
5. However, it should be noted that, together with such an intolerant processing environment, phonologically regular variation does not count as mismatching information (Lahiri & Marslen-Wilson, 1991, 1992).

APPENDIX

Materials of the Word Reconstruction Experiment

Stimuli	Vowel Change Words	Consonant Change Words
Note that British English is nonrhotic, so a stimulus like "pessenger" is pronounced as /'pɛsəndʒə/, not as /'pɛsəndʒər/.		
/'pɛsəndʒə/	passenger	messenger
/'ɜ:mi/	army	early
/'dɔ:pə/	dapper, deeper, doper, dipper	pauper, torpor, daughter, dormer
/'bitəl/	battle, bottle, beetle	little, whittle
/'nɒtɪs/	notice	novice
/'ezi/	easy	any
/wɛsk/	whisk	west, desk
/'mɪtən/	mitten, mutton	melon
/'wɛdə/	widow	meadow
/mə'li:n/	malign	marine
/'i:ftə/	after	easter
/di'zɔ:t/	dessert	resort
/'rɪskəl/	rascal	fiscal
/'arrənd/	errand	island, ironed
/'lɒvəl/	level	novel
/'ti:bəl/	table	feeble
/'sɪvə/	sever, server, savour	liver, river, simmer, sinner
/tɪsk/	task, tusk	risk, disk
/'fɛsəl/	fossil	vessel
/'i:kə/	acre, ocre	eager, eater
/'fɒŋgə/	finger	longer

APPENDIX (Continued)

Stimuli	Vowel Change Words	Consonant Change Words
/ˈseɪfəl/	social	racial, facial
/ˈetɪk/	attic	epic, ethic
/ˈmɛləs/	malice	menace
/əbˈzɔ:v/	observe	absolve
/ˈnɔ:zi/	nosey, noisy	naughty
/ˈɛltɪmət/	ultimate	estimate
/ˈfɪðə/	farther, father, feather, further	wither, hither, dither, fitter
/ˈʃeɪvəl/	shovel	level
/ˈdɑ:ti/	dirty, deity, ditty, dorthy, doughy	party, hearty, tarty, darky
/ˈɑ:fən/	often, orphan	arson
/gru:n/	grown, green, grain, grin	group, groom, gruel, groove, croon, gran
/ɪˈleɪt/	elate, elite	elope
/ˈkæməŋ/	common	canon, gammon
/ˈfeɪʃən/	fashion	nation
/ˈwi:tə/	water, waiter, wetter, whiter, witter	meter, liter, heater, neater, teeter, beater, seater, weeder, weaker
/ɪˈlʊf/	aloof	enough
/ˈsɔ:veɪ/	survey	sorbet
/ˈnɛni/	nanny, ninny	many, penny
/ˈmædəl/	muddle, middle, medal, model	saddle, paddle, madden
/ˈnɪvəl/	naval, novel	civil, nibble, nipple
/ˈkɒlədʒ/	college	courage
/ˈsæmə/	summer, simmer	hammer, salmon, sadder, satyr
/ˈrɒpə/	wrapper	copper
/ˈli:kəl/	local	legal
/dɪˈsi:d/	decide	deceive
/ˈmɛtə/	matter, meter, martyr, motor	better, letter, wetter, setter, mitre, mutter
/ˈmɒri/	marry, merry, morrow, murrey	sorry, lorry
/ˈsɪnət/	sonnet	senate, synod
/əˈpaʊnt/	appoint	account, amount
/ˈrɔ:zɪn/	raisin, rosin	cousin
/ˈfəʊtəs/	foetus	focus, lotus
/ˈwɛməŋ/	women, woman	lemon
/ˈfɪnɪʃ/	famish	finish
/əˈdɔ:n/	attain, atone	adorn
/ˈi:pən/	open	even, eaten
/ˈmæʃən/	mission	passion, fashion
/ˈsəpraɪm/	supreme	surprise
/ˈfɪktəri/	factory	victory
/ˈpɔɪdən/	pardon	poison

(Manuscript received June 28, 1995;
revision accepted for publication February 8, 1996.)