

Phonemic identification in a phoneme monitoring experiment: The variable role of uncertainty about vowel contexts

DAVID A. SWINNEY and PENNY PRATHER
Tufts University, Medford, Massachusetts 02155

Subjects monitored for the syllable-initial phonemes /b/ and /s/, as well as for the syllables containing those phonemes, in lists of nonsense syllables. Time to detect /b/ was a function of the amount of uncertainty as to the identity of the vowel following the target consonant; when uncertainty was low, no difference existed between phoneme and syllable monitoring latencies, but when uncertainty was high, syllables were detected faster than phonemes. Time to detect /s/ was independent of uncertainty concerning the accompanying vowel and was always slower than syllable detection. The role of knowledge of contexts in a phoneme-monitoring task as well as the relative availability of phonemic information to the listener in this task are discussed.

The processes involved in the recognition and the identification of phonemic components of the speech signal have been examined from a number of perspectives in the past few years. One line of study, originally undertaken in an attempt to determine the "basic" psychological and acoustic units of speech perception, has used a monitoring task to examine the relative order of detectability of phonemes, syllables, and words in speech. Savin and Bever (1970) began this work by demonstrating that monitoring (detection) time for syllables is faster than that for phonemes in a list of syllables, a finding they took to be evidence that syllables are basic perceptual units and that phonemes are not. A number of subsequent studies, however, have forced considerable reinterpretation of these results (see, in particular, Foss & Swinney, 1973; Healy & Cutting, 1976; McNeill & Lindig, 1973). Foss and Swinney (1973), for example, demonstrated that not only are single syllables detected faster than phonemes, but that two syllable nonsense words are detected faster than single syllables (a finding that would, by parity of argument, require the two-syllable word to be the perceptual primitive in speech). In presenting this reductio demonstration, Foss and Swinney argued that the monitoring task does not directly reflect those processes involved in phoneme perception, but, rather, the process by which one gains conscious access to the phoneme, a process which they labeled phoneme identification. Such identification was argued to involve (in rough order): the perception of the speech sounds (extraction of phonetic information from the acoustic signal) followed by identification of the larger units (syllables, words, etc.) in which these speech sounds occurred, followed by the subsequent (conscious) identification of the phonemes comprising these units.

The distinction made between perception and identification of phonemes in a monitoring task is, we would argue, an important one. Monitoring for any target (phoneme or otherwise) is essentially a hypothesis-testing, matching-to-sample type of operation. It requires that a stimulus input be sufficiently identified so that a matching response can be made to it. While the perceptual process of extracting phonetic information from the acoustic signal is an important component of this process, it is not necessarily isomorphic with the identification of a particular linguistic segment (see, Foss & Swinney, 1973, for further discussion of this position). The distinction between perception and identification has been particularly useful for gaining some perspective on the apparent discrepancies that have resulted from the various uses of the phoneme monitoring task in investigations of units of speech perception (see, e.g., Ball, Wood, & Smith, 1975; Foss & Swinney, 1973; Healy & Cutting, 1976; McNeill & Lindig, 1973; Morton & Long, 1976; Rubin, Turvey, & VanGelder, 1976; Shand, 1976; Swinney, Note 1). Further, because the monitoring task has been used in a large number of experimental situations in which the relative availability of the target phonemes to direct access has been made an issue (see, e.g., Foss, in press; Morton & Long, 1976; Newman & Dell, 1978; Foss & Blank, Note 2; Marslen-Wilson, Note 3), it is important that this issue be further examined and that the factors which govern the relative availability of phonemes (and other units) in a monitoring task be detailed.

One important aspect of this problem concerns the role that knowledge of contextual information plays in the relative detectability of phonemes and syllables. For example, when subjects are asked to monitor for a syllable in a list of syllables, they are

given the entire acoustic characterization for the target. However, when they monitor for a single syllable-initial phoneme, all of the requisite acoustic information may not be provided; there is a large amount of uncertainty about the identity of the contextual information which will accompany the target. It may be that such information is important in the identification process for the phoneme. While evidence exists (although some of it is controversial) which suggests that most initial phonemes have invariant, context-independent cues by which they can be perceived (see, e.g., Cole & Scott, 1974a, 1974b; Stevens & Blumstein, 1978), evidence also exists that the recognition of certain phonemes is often affected by surrounding contexts. In a number of task situations, the recognition of initial stop consonants (/b/, for example) appears to be far more influenced by the vowel and vowel transition that follow them than is the recognition of other phonemes (such as /s/ and /z/) (see, e.g., Cole & Scott, 1974a, 1974b; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Pisoni, 1978).

If invariant cues alone are the functional factors in phoneme identification, then the classically reported syllable superiority effect seems unlikely to be related to the potential confounding of the amount of information presented in the phoneme and syllable target conditions. However, it seems possible that the lack of knowledge concerning information contained in the vowel and transition accompanying a target consonant may be responsible for the observed phoneme-syllable monitoring differences. Savin and Bever (1970), in fact, first raised one form of this possible objection. However, they quickly dismissed it as a causative factor due to the fact that the syllable monitoring superiority effect held for conditions in which the targets involved the phoneme /s/ as well as the phoneme /b/. Because it has been argued that the recognition of /s/ does not depend on information provided by contextual information, Savin and Bever concluded that uncertainty over identity of the vowel accompanying the target consonant was not the basis for the observed syllable superiority effect.

This line of reasoning, however, deserves further examination. The manipulation utilized by Savin and Bever does not, in fact, tell us whether the phoneme-syllable difference obtained for /b/ is independent of uncertainty over identity of the vowel. It only tells us (following the Foss & Swinney, 1973, argument) that the identification of /s/ occurs subsequent to identification of the syllable in which it occurs, even if such identification is independent of contextual information. It seems important, then, to pursue an answer to the original question, particularly because the monitoring technique continues to be used to make arguments about speech perception, syllable processing, and lexical access in the face of this potentially confounding factor.

It may be true, of course, that the process of phoneme identification which is involved in the monitoring task is unaffected by uncertainty concerning accompanying phonemic contexts. Newman and Dell (1978) have argued for a position which can be interpreted to support this view; their paper demonstrates that the number of distinctive features shared between the phoneme target in a word and the initial phoneme of the word preceding the target-bearing word directly affects phoneme monitoring times. They argue that the phoneme monitoring task involves a direct, bottom-up match of distinctive feature from the given target to candidate phonemes; furthermore, they claim that their data suggests that this match occurs prior to identification of the word (or parts of the word) containing the target phoneme. Thus, phoneme identification is argued to be a truly bottom-up process, one in which contextual information accompanying a target phoneme is seen as having no specific or direct effect on the monitoring process. (They do, however, argue that the top-down use of contextual information could take place, but not to the exclusion of the bottom-up process.)

On the other hand, the assumption that knowledge of the identity of an accompanying vowel context is not responsible for the syllable superiority effect (and the related claim that phoneme monitoring involves a direct match of distinctive features) may be somewhat premature. There has been no direct examination of either of these claims; both are inferences (however reasonable in appearance) derived from somewhat indirect experimental evidence. As just one piece of potential counter evidence, Wood and Day (1975), using a selective attention paradigm, found that variation in an accompanying vowel does affect the forced choice identification of /b/ and /d/. While this was not, in fact, a monitoring task, it certainly suggests that knowledge of contexts may have an important role to play in such a task. It seems likely, for example, that when subjects are told to monitor for a /b/ (subjects are typically given an auditory example: "listen for the /bə/ as in the word /bat/" or "listen for the /bə/ sound") they will hold an auditory version of the target in mind, a situation necessitating the use of some particular accompanying vowel. Even in experiments where the specified target phoneme is presented visually, subjects typically report that they hold an acoustic target in mind. If so, it seems likely that the degree of match (or mismatch) between the vowel held in mind along with the specified target phoneme and the vowel accompanying the target in the stimulus materials may have a major effect on phoneme monitoring latencies. However, for reasons suggested above, such potential effects may not be functional for all types of phoneme targets.

In order to directly examine the effects of knowl-

edge of (and/or uncertainty over) the identity of the vowel which accompanies the target phoneme in a monitoring task, a two-part experiment was designed. In this, subjects monitored for either a syllable-initial phoneme or the entire syllable containing that phoneme under three different experimental conditions. This was done for both the phoneme /b/ (Experiment 1a) and the phoneme /s/ (Experiment 1b). In one of the three experimental conditions, the target phoneme was always followed by the same vowel; in the second condition, it was followed by one of four different vowels; and in the third condition, it was followed by any one of eight different vowels.

Several aspects of the question about the role of contextual knowledge on phoneme and syllable monitoring can be examined with this design. First, if the identification of target phonemes takes place independently of consideration of contextual information provided by the accompanying vowel, the vowel uncertainty manipulations should create no differences in monitoring times to the phonemes. In addition, if phoneme identification takes place independently of (and prior to) the access and identification of the syllable in which the target is embedded (that is, via a direct perceptual feature match), we can expect to find phoneme monitoring to be faster than syllable monitoring. Finally, if the nature of the process involved in phoneme identification is uniform among phonemes, we can expect that phoneme-syllable monitoring differences for /s/ and for /b/ will be essentially the same for comparisons made within each of the three uncertainty manipulations.

On the other hand, if the phoneme identification process requires the information carried by the vowel and/or the consonant-vowel transition for a "decision," then we can expect that the vowel-uncertainty manipulation will affect phoneme monitoring times (in comparison to syllable monitoring times). Thus, if the basic syllable superiority effect is actually a function of the amount of information provided about the respective targets, then we can expect the phoneme-syllable monitoring difference to disappear when the vowel context is uniquely identified (the one-vowel condition). In addition, we can expect that the size of the phoneme-syllable monitoring difference will be a function of the amount of uncertainty held about the vowel context. Further, if the processes used to identify phonemes in this task differ in their dependence on contextual information (as has been suggested may be the case for /s/ and /b/), then we can expect to find smaller differences to be created by the uncertainty manipulation when the phoneme target is /s/ than when it is /b/.

EXPERIMENT 1a

Method

Subjects. Twenty-four undergraduates from Tufts University participated in this experiment in partial fulfillment of an introductory psychology course requirement.

Materials. The materials consisted of 56 lists of single-syllable nonsense words. The lists were from 5 to 12 syllables in length, with equal numbers of lists of each length. Thirty-two of the lists contained monitoring target materials and 24 did not (the latter were included to help prevent increased expectation for a target near the end of a list). The target syllables appeared with equal frequency in Positions 3-10 in the list, with two syllables always following each target syllable in any list.

All target syllables were of CVC construction, and the initial consonant was always /b/. Three different sets of target syllables were constructed, each with different numbers of vowels comprising the target syllables. In one, the vowel in the target syllable was always the same (/ɔ/) for each of the 32 experimental target syllables. In the second, there were four different vowels (/ɪ/, /e/, /u/, /ɔ/) which were used with equal frequency among the 32 experimental target syllables (eight exemplars each). In the third, eight different vowels (/i/, /ɪ/, /e/, /æ/, /u/, /o/, /ʊ/, /ɔ/) were used in the 32 experimental targets (four exemplars each). The final consonant in each of the three conditions was chosen from the following set of consonants: /v/, /f/, /ʒ/, /ʃ/, /b/, /p/, /ð/, /θ/, /g/, /k/. The distribution and use of final consonants was dictated by the single constraint that the resultant target syllable not form a word. This resulted in nearly identical distributions in the use of final consonants across all three conditions.

Each of the three sets of 32 experimental target syllables was placed in the 32 experimental lists, which were combined in random order with the 24 nontarget lists. The three resultant sets of 56 nonsense word lists were each recorded separately on audio tape. The recording was made by a practiced male speaker. Within each list, syllables were heard at a rate of approximately one syllable every .5 sec. There was a 2-sec delay placed between lists. In each set, half of the 32 target syllables were designated as phoneme targets and half were designated as syllable targets. This was accomplished in a manner which resulted in exemplars of each type of target condition (created by the vowel and final consonant manipulations) being equally represented in these two target conditions. Thus, the syllable and phoneme targets were assigned to equivalent sets of experimental materials (16 exemplars each).

In those lists containing targets, a signal that was inaudible to the subjects was placed coincident with the beginning of the target syllable. This signal initiated a timing mechanism in a PDP-8/I computer, which was terminated by the subject's response. The placement of the timing signal was performed by moving the tape manually across the tape-recorder head until the onset of the target sound was detected. Signal placements were checked oscilloscopically in a post hoc examination to determine placement accuracy. All signal placements were found to be accurate to within 4 msec, with no systematic differences in placement variation existing across experimental conditions (i.e., all placement errors were both small and equally distributed across experimental conditions). The presentation and reaction time recording mechanism were under control of a PDP-8/e computer (see Onifer, Hirshkowitz, & Swinney, 1978, for a description of the apparatus and software package). The recorded reaction times were accurate to within $\pm .5$ msec.

Procedure. The subjects were told that they would hear lists of nonsense syllables, each list preceded by the specification of a target for that list. They were told that the target would be either a syllable in the list or the initial phoneme of some syllable in the list, and that the specified target would not actually be in all of the lists. The subjects were instructed to press a button (on

Table 1
Mean Reaction Time (in Milliseconds) to Syllable and Phoneme Targets in Each of the Three Variability Conditions for Experiments 1a and 1b

Monitoring Target	Number of Vowels Accompanying Target (Vowel Uncertainty)		
	1	4	8
Phoneme /b/	283	298	315
Syllable Containing /b/	274	275	277
Phoneme /s/	426	430	430
Syllable Containing /s/	383	387	382

which they rested a finger at all times) as soon as they heard the specified target in the list. In all cases involving a phoneme target, the target was specified by the phrase "listen for the /bə/ sound." The stimulus materials were presented binaurally over headphones. Each subject listened to all three experimental tapes. All six possible orders of presentation of these three tapes were used. Four subjects participated in each of these six order conditions.

Results

Mean reaction times were calculated for each of the vowel manipulations (one, four, or eight vowels) in each of the monitoring target conditions (phoneme and syllable), for each subject. The mean of these data, calculated across all subjects, is given in the first half of Table 1 (those materials involving /b/). The data were submitted to an analysis of variance, with number of vowels and monitoring target as main effects. There was a significant main effect for monitoring target [$F(1,23) = 6.38, p < .02$], a significant main effect for number of vowels [$F(2,46) = 3.19, p < .05$], and a significant interaction between vowels and monitoring target [$F(2,46) = 5.91, p < .01$].

Planned *t* tests were computed to examine the individual phoneme-syllable differences for each of the different vowel conditions in the significant interaction. There was no significant difference between phoneme and syllable monitoring conditions when the vowel was held constant throughout the conditions [one vowel; $t(23) = .73$], but the four- and eight-vowel conditions each displayed significant phoneme-syllable differences [$t(23) = 2.66, p < .005$, $t(23) = 6.93, p < .0005$, respectively, on one-tailed tests].

EXPERIMENT 1b

Method

This experiment was identical in every respect to Experiment 1a, except that 24 new subjects participated in the study and the initial phoneme on target syllables was always /s/ (and, hence, the designated monitoring target was the phoneme /s/).

The timing signal, as before, was placed coincident with the minimal detectable onset of the target phoneme. Placement of the timing signal was examined oscilloscopically and found to be accurate to within 4 msec. Errors in placement, as before, were distributed equally in each of the three experimental conditions. There was no difference in the distribution of timing signal errors between syllables used for phoneme and syllable targets.

Results

Mean reaction times were calculated for each of the vowel conditions in each of the monitoring target conditions (syllable vs. phoneme), for each subject. The means of these data computed across all subjects are given in the second half of Table 1. The individual data were submitted to an analysis of variance, with number of vowels and monitoring target conditions as main effects. While the monitoring target main effect was significant [$F(1,23) = 4.98, p < .03$], neither the manipulation of the number of vowels nor the interaction of vowels with monitoring target were significant [$F(2,46) = .087, F(2,46) = .71$, respectively].

Discussion of Experiments 1a and 1b

The results for Experiment 1a provide relatively straightforward evidence that increases in uncertainty concerning the nature of the vowel accompanying the stop consonant target /b/ in a phoneme monitoring task increase the reaction time to detect that phoneme in a monitoring task. As one might expect, reaction time to detect a syllable, which is exactly specified in each case, does not differ in the three vowel manipulation conditions. It is, of course, the syllable-phoneme comparisons involved in the interaction of the monitoring target and the vowel manipulation factors that is of prime interest here. In these, it is particularly important to note that while the often-reported syllable superiority effect was obtained in the eight-vowel variability condition of this study, there was no significant difference in reaction time to monitor for phonemes and syllables when the vowel was held constant throughout the experiment (the one-vowel condition). Further, the size of the difference in phoneme-syllable detection time for the intermediate (four-vowel) uncertainty condition falls between that for the one-vowel and eight-vowel conditions. In short, at least with the phoneme /b/, it appears that the detection/identification process involved in phoneme monitoring is affected by knowledge of the identity of the vowel accompanying the target phoneme.

On the other hand, the results of Experiment 1b suggest that the identification process for the phoneme /s/ in a phoneme monitoring task is quite different from that for /b/ (Experiment 1a). Regardless of whether the target phoneme /s/ was accompanied by the same vowel, four different vowels, or eight different vowels in the experiment, reaction time to monitor for that phoneme was identical. Thus, these results suggest that in some cases (such as for /s/), identification of a phoneme in a monitoring task is independent of uncertainty concerning identity of the vowel (and the transition to that vowel) accompanying the target phoneme.

It should be noted that the data argue against the view that the lack of an effect for the vowel uncer-

tainty conditions on the /s/ targets is simply due to inaccurate placement of the timing signal (as might be thought possible due to variability in the rise of the energy or amplitude envelope for /s/). Reaction times to detect the /s/ phoneme are remarkably similar for each of the three vowel uncertainty conditions (averaging 426, 430, and 430 msec, respectively), a fact which would argue against variability in placement of the timing signal. Likewise, variance in reaction times for each of these three vowel uncertainty groups is nearly identical and quite small (.0001, .00016, and .00014, respectively). Finally, it should be noted that there were no differences in the amount of error involved in the timing signal placement to the syllable and the phoneme target conditions involving /s/. With respect to this, it can be seen that reaction times to the syllables containing /s/ are extremely similar across the vowel manipulation conditions, as was the case for the /s/ phoneme targets. Interestingly, the variance in these syllable conditions compares favorably to the variance found for the syllable targets containing /b/. Thus, it appears that the reliable lack of any difference among the three vowel manipulations for the /s/ conditions can be argued to be due to the failure of the manipulation (the vowel variations) to induce such an effect, and not a result due merely to large variability in signal placement or some similar methodological problem.

Finally, it should again be noted that in all conditions (except for that in which there was no variation in the vowel accompanying the initial phoneme /b/ -- the one-vowel condition) reaction times to monitor for phonemes (both /s/ and /b/) were significantly slower than those to monitor for the syllables containing those phonemes. Thus, the general syllable-superiority effect reported in the literature replicates in certain of the present experimental conditions. What is of greatest importance, however, is the fact that the relative size of this effect can be manipulated by varying the listener's uncertainty concerning identity of the vowel which accompanies the monitored-for target, *for some types of monitoring targets*. For other types of targets, the syllable superiority effect holds regardless of uncertainty over the vowel accompanying the target. This difference, which reflects differences in the nature of these processes used to obtain access to /b/ and /s/ in a monitoring task, requires further discussion. However, one further aspect of these results requires attention first. It might be thought that there are still some circumstances in which a subject could monitor for such initial stop-consonant targets as /b/ in the abstract, that is, independently from vowel contexts. A subject might, for example, be able to ignore variance in the target vowel if such variance were specifically brought to his or her attention. In order

to examine this possibility, a second experiment was run. This experiment was identical to Experiment 1a except that the subjects were explicitly told that the monitored-for phoneme (/b/) would not occur in the exact form given in the prelist target designation in the four- and eight-vowel experimental conditions.

EXPERIMENT 2

Method

The materials and procedures were identical to those in Experiment 1a, with the exception that, in conditions involving four- and eight-vowel variations, subjects were explicitly told that the phoneme target (/b/) could occur in different syllables (such as "bap," or "bok," or "boof") throughout the experiment so that it might not always sound exactly like "/bə/." Subjects were urged to focus upon the /b/ portion of the syllable in the phoneme monitoring conditions.

Subjects. Twenty-four subjects from the University of Texas at Austin participated in the experiment in partial completion of an introductory psychology course requirement.

Results

Table 2 displays mean reaction times for each of the three vowel conditions in each of the two monitoring target conditions. Individual data for each subject were submitted to an analysis of variance in which number of vowels and monitoring target (syllable vs. phoneme) were main effects. The main effect for monitoring target was significant [$F(1,23) = 6.72, p < .02$], as was the interaction of number of vowels and monitoring target [$F(2,46) = 6.32, p < .01$]. The main effect for variability was not significant [$F(2,46) = .10$].

Discussion

The results of Experiment 2 provide a replication of the important aspects of Experiment 1a. Even when subjects are made explicitly aware of the variation in the values which the vowel context accompanying a monitored-for phoneme /b/ may take, they apparently cannot use that information. The latencies obtained in this experiment are somewhat faster than those in Experiment 1a, a fact which most likely has to do with individual differences in the two groups of subjects that were run, and with the

Table 2
Mean Reaction Time (in Milliseconds) to Syllable and Phoneme Targets in Each of the Three Variability Conditions for Subjects Who Were Instructed About Variance in Vowels (Experiment 2)

Monitoring Target	Number of Vowels Accompanying Target (Vowel Uncertainty)		
	1	4	8
Phoneme /b/	270	286	300
Syllable Containing /b/	273	265	262

“special” attention the subjects were accorded in the instruction period of Experiment 2.

GENERAL DISCUSSION

Overall, the data show that knowledge of the identity of the vowel accompanying an initial /b/ affects the time to detect that phoneme in a monitoring task. Furthermore, this effect holds regardless of whether the listener is aware of the uncertainty manipulation or merely is dealing with it at some unconscious level. In either case, knowledge appears to be quickly induced as to the range of values that accompanying vowels will take in each experimental condition. Thus, the conclusion drawn by Savin and Bever (1970) (and the subsequent assumptions made by a large number of authors) about the role that uncertainty about vowel context plays in producing the monitoring latency differences between syllables and phonemes appears to be incorrect, at least for some types of phonemes. It should be noted that, while these results are only (and are only meant to be) generalizable to the target phoneme /b/, this phoneme has traditionally been one of the most commonly used targets in phoneme monitoring studies. (However, the commonalities reported in the literature about the degree to which the various stop consonants are affected by contextual information does lead us to suspect that other stop consonants may produce results similar to those obtained for /b/.)

Contrasting the results of /b/ with those obtained for /s/ (in which the syllable superiority effect was undisturbed by manipulations of vowel context uncertainty) leads us to several conclusions with respect to the phoneme monitoring task. First, and most obvious, are the methodological considerations: assumptions made about the nature of mental processes involved in any recognition/identification task must be carefully examined before conclusions are drawn from use of such tasks. Experimental designs which employ phoneme and syllable monitoring must consider factors such as the degree to which contextual information must be provided in order to present equivalent knowledge of various targets to the listener. Summarizing results obtained across a variety of targets produces the risk that different (individual) processes will be lost from view in the grouping procedure.

The results also lead us to reject the notion that phoneme identification takes place based on a strictly bottom-up perceptual match of invariant features that are carried independently of knowledge provided by contextual information. Rather, we take these results to support the basic premise of Foss and Swinney (1973) that phoneme monitoring is an identification process that takes place subsequent to phonetic perception. However, while Foss and Swinney

argued that identification of phonemes took place only following the identification of the larger, more meaningful units in which they occurred, the current results (those demonstrating that nonsense syllables are not detected more rapidly than some phonemes when the vowel accompanying the phoneme is uniquely identified) suggest that some modification is required in that hypothesis. We would argue that, rather than always proceeding from the largest unit, identification of information in the monitoring task actually proceeds from the most easily *confirmable* unit—that unit which provides a sufficient amount of verifiable information to allow for an initial decision on the part of the listener (see Healy & Cutting, 1976, for discussion of some related concepts). The nature (and size) of such a “minimal confirmable unit” will vary greatly depending on task and materials conditions. When phoneme monitoring involves real words and sentences, the automaticity and speed of the sentence comprehension processes are such that words (or perhaps the entire sentence) will be the first confirmable units (for all of the reasons argued by Foss & Swinney, 1973). On the other hand, when the materials are nonsense syllables, as in the present study, there is no higher order structure or automatized process (such as that which is present in the highly practiced routines which provide for access to existing lexical representations) to aid in confirming the identity of information derived directly from perceptual analysis. When sufficient contextual information is provided, decisions to confirm the presence of some phonemes (such as /b/) are as rapid as decisions to confirm the presence of nonsense syllables containing that phoneme. This is because, with no existing higher order structure to aid in the confirmation of either one of these “nonsense” targets, it takes an equal amount of time to confirm these units as long as equivalent amounts of information are provided about each target. (Note that, as with nearly every study using this paradigm, the CVC syllable can be uniquely identified by the initial CV information; no foils are used in which nontarget syllables differ from target syllables in all but the final consonant.) Information carried by the accompanying vowel (or transition) is important in providing a processing base for the identification of /b/. Thus, when uncertainty exists as to the features provided by the vowel or transition, confirmation of the identity of the consonant must wait until the syllable containing the target is identified and analyzed into its “abstract” phonemic units.

While, for the reasons suggested earlier, it was expected that /s/ would be unaffected by the vowel manipulation, it might seem surprising that detection of the phoneme /s/ always takes longer than detection of the syllable beginning with /s/. That is, one might wonder why a perceptual unit which has invar-

iant, context-independent cues cannot be identified faster than some larger unit of which it is a part. Viewed in the current framework, the answer appears to involve facts about the nature of confirming information that is required by the identification process; whatever the invariant cues associated with /s/, they do not appear to be sufficient to meet the criteria for rapid and easy confirmation in this task. Without attempting to push the argument too far, in the absence of any substantiating data, we suggest that the problem may be related to difficulties in uniquely distinguishing /s/ from signal noise. While the various invariant features which have been detailed for /s/ provide unique cues in comparison to other phonemes (see, e.g., Cole & Scott, 1974a), they appear likely to be highly confusable with noise in the signal. Thus, if a subject's identification criteria are relatively stringent, they may not allow for a confident identification response to an /s/ in the absence of other, confirming information. On the other hand, once the syllable containing /s/ is identified, it can be decoded by rule into its component phonemes, and thus provide sufficient evidence that /s/ was present in the signal. We speculate that the relative detectability of /s/ (compared to a target syllable) in language materials will be greater than its relative detectability in nonsense syllables, due to the confirming evidence provided by the content and structure of the language which would precede the target in such a situation.

REFERENCE NOTES

1. Swinney, D. *The effect of target size and organization upon reaction times during sentence processing*. Paper presented at Midwestern Psychological Association, Chicago, May 1973.
2. Foss, D. J., & Blank, M. A. *On the alchemy of speech perception during sentence processing: Dross into gold or the reverse?* Paper presented at Midwestern Psychological Association, Chicago, 1977.
3. Marslen-Wilson, W. D. *Sequential decision processes during spoken word recognition*. Paper presented at Psychonomic Society, San Antonio, 1979.

REFERENCES

- BALL, F., WOOD, C., & SMITH, E. When are semantic targets detected faster than visual or acoustic ones? *Perception & Psychophysics*, 1975, 17, 1-8.
- COLE, R., & SCOTT, B. Distinctive feature control of decision

- time: Same-different judgments of simultaneously heard phonemes. *Perception & Psychophysics*, 1972, 12, 91-94.
- COLE, R., & SCOTT, B. Toward a theory of speech perception. *Psychological Review*, 1974, 81, 348-374. (a)
- COLE, R., & SCOTT, B. The phantom in the phoneme: Invariant cues for stop consonants. *Perception & Psychophysics*, 1974, 15, 101-107. (b)
- DIEHL, R., ELMAN, J., & MCCUSKER, S. Contrast effects on stop consonant identification. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, 4, 599-609.
- FOSS, D. J. Deciphering decoding decisions: Data and devices. In R. A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, N.J.: Erlbaum, in press.
- FOSS, D., & SWINNEY, D. On the psychological reality of the phoneme: Perception, identification, and consciousness. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 246-257.
- HEALY, A., & CUTTING, J. Units of speech perception: Phoneme and syllable. *Journal of Verbal Learning and Verbal Behavior*, 1976, 15, 73-83.
- LIBERMAN, A., COOPER, F., SHANKWEILER, D., & STUDDERT-KENNEDY, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-461.
- MCNEILL, D., & LINDIG, K. The perceptual reality of phonemes: Syllables, words, and sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 419-430.
- MORTON, J., & LONG, T. Effects of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior*, 1976, 15, 43-51.
- NEWMAN, J. E., & DELL, G. S. The phonological nature of phoneme monitoring: A critique of some ambiguity studies. *Journal of Verbal Learning and Verbal Behavior*, 1978, 13, 359-374.
- ONIFER, W., HIRSHKOWITZ, M., & SWINNEY, D. A miniprocessor PDP8/e-based system for investigations of on-line language processing: Automated program for psycholinguistic experiments (APPLE). *Behavior Research Methods & Instruments*, 1978, 10, 307-308.
- PISONI, D. Speech perception. In W. Estes (Ed.), *Handbook of learning and cognitive processes*. Hillsdale, N.J.: Erlbaum, 1978.
- RUBIN, P., TURVEY, M. T., & VAN GELDER, P. Initial phonemes are detected faster in spoken words than in spoken nonwords. *Perception & Psychophysics*, 1976, 19, 394-398.
- SAVIN, H., & BEVER, T. The nonperceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 295-302.
- SHAND, M. Syllabic vs. segmental perception: On the inability to ignore "irrelevant" stimulus parameters. *Perception & Psychophysics*, 1976, 20, 430-432.
- STEVENS, K. N., & BLUMSTEIN, S. E. Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 1978, 64, 1352-1368.
- WOOD, C., & DAY, R. Failure of selective attention to phonetic segments in consonant-vowel syllables. *Perception & Psychophysics*, 1975, 17, 346-350.

(Received for publication January 9, 1979;
revision accepted November 20, 1979.)