

Identification times for phonemic components of graded complexity and for spelling of speech*

RICHARD M. WARREN

University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201

Identification time was measured for targets in speech in terms of the number of phonemes occurring between the target and response. Auditory targets in the main experiment were at three levels of phonetic complexity: (1) monosyllabic words or nonsense syllables, each containing four phonemes; (2) clusters (vowel + consonant) consisting of the second and third phonemes within these syllables; (3) individual phonemes within these clusters. Identification time was always shorter for (1) than for (2) or (3). The effect of semantic and grammatical context upon identification was small. Plosive consonants which did not begin a syllable generally could not be identified as isolated phonemic targets, although they could be identified readily as part of a cluster. Letter targets corresponding to the spelling of the auditory stimulus permitted identification of plosive consonants, and in general followed different rules than the phonemic targets. Additional observations concerning individual targets were made.

While speech may be characterized and transcribed as a succession of individual phonemes, it has been pointed out many times that perception of speech cannot proceed in this simple phonetic fashion—that the listener does not detect single phonemes which he then joins together to form clusters, syllables, words, phrases, and sentences. Among the recent arguments against the phoneme as a perceptual building block is that of Miller (1962), who stated that a process based upon Markovian chaining of phonemes would be extremely sensitive to errors: a mistake in one would cause errors in identification of subsequent phonemes which could not be corrected, causing the whole perceptual process to stall. Investigators at the Haskins Laboratories (Lieberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967) have shown that it would not be possible for all phonemes to be perceived as separate and discrete entities, since some speech sounds (notably some plosive consonants) have no invariant auditory characteristics and can be described only in terms of the acoustic environment provided by other speech sounds.

Illusory changes in perception of repeated words (the verbal transformation effect) have indicated that the size of the smallest perceptual units employed in verbal organization varies with age, and

that for young adults, the units seem to be larger than phonemes and smaller than meaningful words (Warren, 1968). Another illusion, the phonemic restoration effect, has suggested that a prior organization into clusters or words served as the basis for subsequent analysis into component phonemes. In this illusion, a listener is unable to tell which speech sound in a clearly spoken sentence has been replaced by an extraneous sound (such as a cough), since the cough is mislocalized, and every phoneme, including the one synthesized perceptually in keeping with the supraphonemic organization, seems equally clear (Warren, 1970).

If, as suggested above, perceptual identification of speech is initiated at a phonetically complex level, then it would be anticipated that identification time would be longer for simpler targets. If we take a single word in a sentence and assign one group of listeners the task of pressing a response key as soon as they detect the target words and assign other groups the task of pressing the key when they identify individual phonemes and phoneme clusters within the same word, we may trace the times required for recognition at different levels of stimulus complexity.

The present investigation is a preliminary study, testing the feasibility of this method. The variables investigated were limited in number, with the aim of determining the utility of the technique, and a few selected characteristics of temporal identification of speech. All Ss heard identical stimuli, but in the main experiments different groups had different targets. A comparison was made of the time required for identification of an entire word with the time required for individual

phonemes and phoneme clusters within that word. The time required for identification of target letters used in spelling was compared with the time required for identification of the corresponding speech sound. Additional observations of a preliminary nature were also made.

GENERAL METHOD AND PROCEDURE

Subjects

The 36 Ss in Experiment 1 and the 9 Ss in Experiment 2 were mostly staff, visiting scientists, and graduate students at Oxford University. The graduate students were paid 5 shillings. Each S was tested separately.

Stimuli

The nine sentences used as stimuli were all recorded at a normal speaking rate (average for each sentence about 12 phonemes/sec) on a Revox A77 tape recorder. The stimuli were played back on the Revox recorder and heard binaurally through a pair of matched S. G. Brown headphones at a peak voice level of 80 dB (re 0.0002 microbars). All Ss in both Experiments 1 and 2 heard the same stimuli presented in the same order; the only difference between experimental groups was in the designated response targets.

Procedure

After hearing the instructions appropriate to his experimental group, S was shown the response key, which was the "stop" key of a Sony TC-100 cassette recorder. The technique for measuring responses was described to S; the electrical output from the Revox recorder went simultaneously to their headphones and to the Sony recorder. When S detected the target and depressed the key, the termination of the Sony rerecording measured the time taken to detect the target. The output to the headphones was not influenced by the response key.

EXPERIMENT 1: PROCEDURE

Each S heard the following instructions: "This experiment deals with the reaction time to speech sounds. You are to press a key as soon as you hear a designated target, which will be either a single speech sound, a group of speech sounds, or a word. The spelling is unimportant—you should react only to the sounds, regardless of how they are spelled. Also, the meaning of the sentence is of no importance. You will be asked nothing concerning meaning, so concentrate on the sounds as they occur. To summarize, the purpose of this

*Supported by a University of Wisconsin-Milwaukee Graduate School research fellowship. The author thanks Professor Lawrence Weiskrantz for making the facilities of the Institute of Experimental Psychology, Oxford University, available during the experimental work and the writing of this manuscript.

Table 1
Median Identification Times (in Phonemic Units) for Target Phonemes and Letters in the Main Parts of Experiments 1 and 2

Stimuli	Experiment 1 (Four Groups of Nine Ss)						Experiment 2 (Nine Ss)							
	GpA:		(#)	GpD:		(#)	GpB:		(#)	GpC:		(#)		(#)
M1	GpA: /ston/		(0)	GpD: /o/		(3)	GpC: /t/		(∞)	GpB: /to/		(3)	"T"	(9)
M2	GpB: /bren/		(3)	GpA: /e/		(6)	GpD: /r/		(6)	GpC: /re/		(6)	"R"	(8)
M3	GpC: /lups/		(2)	GpB: /u/		(5)	GpA: /p/		(∞)	GpD: /up/		(3)	"P"	(6)
M4	GpD: /skiv/		(2)	GpC: /i/		(5)	GpB: /k/		(15)	GpA: /ki/		(4)	/i/	(5)
M5	GpA: "P"		(13)	GpD: "L"		(4)	GpC: "O"		(11)	GpB: "P"		(10)	/p/	(10)

experiment is solely to measure the speed at which you can respond to designated sounds and words."

Practice Sentences

The Ss were informed that the first few sentences would be practice sentences and that, while their performance would be scored, the major purpose was familiarization (this information was, in fact, quite true). All 36 Ss heard four practice sentences, presented in the order given below. When the targets were individual phonemes and phoneme clusters (but not a word), they were given two examples of words containing the target (these examples never had the phonemes preceding or following the target matching those actually present in the stimulus word). After S responded, he was required to name the word containing the target.

Sentence P1: "While many actors tried that evening, they found that they could not manage to calm the audience nor obtain applause for their arduous efforts." (Target, /l/.)

Sentence P2: "The air was cold and the sky filled with clouds, but the snow failed to fall despite all expectations." (Target, /n/.)

Sentence P3: "Many artists as well as scientists during the eighteenth century required private patrons to support their work with generous awards of funds." (Target, /tr/.)

Sentence P4: "The attempts of alchemists to transmute base metals into gold led to the development of modern chemistry." (Target word deliberately ambiguous, and could be either of the homophones "led" or "lead.")

The S was informed of any error he made and told the actual position of the target following his response.

Main Sentences

Following the four practice sentences, Ss were informed that the main part of the experiment was commencing. Each S was assigned randomly to one of four groups, each group containing nine Ss. All groups heard the stimulus sentences in the order listed below, but each group was given different targets within the same word in the sentence. For example, in Sentence M1, Group A had as target the entire word "stone," Group B had the

consonant-vowel cluster /to/, Group C had the consonant /t/, and Group D had the vowel /o/ as target. In this sentence, as in all sentences, M1 through M4, the targets for each of the four groups were within the same word, so that a comparison of identification times for various components could be made with each other and with that for the entire word. The words containing the targets in Sentences M1 through M4 all consisted of three consonants and one vowel, with the vowel either as the second or the third phoneme. The type of target changed for members of each group with successive sentences, so that every S had each of the four types of targets once [i.e., vowel, consonant, vowel-consonant (or consonant-vowel) cluster, and word (or nonsense syllable)]. The assignment of each target is listed in Table 1. As in the practice sentences, Ss were given two examples of words (or nonsense syllables with Stimulus M4) containing the target when targets were individual phonemes or phoneme clusters. Before Stimulus M3, Ss were told that they would hear a string of English words which did not form a meaningful sentence, and before Stimulus M4, they were told that they would hear nonsense words consisting of English speech sounds. Ss were always informed whether or not their response was correct, and if wrong, were told the correct response before proceeding to the next stimulus. In the list which follows, the word containing the targets is preceded and followed by asterisks (*).

Stimulus M1 (target word in keeping with prior context): "The final place described by the planning group as a possible region for the building of the *stone* dam had a firm foundation of bedrock."

Stimulus M2 (target word not in keeping with prior context): "The wind had howled since seven o'clock, and the *brain* was numbed by the noise and chill which could not be kept outside."

Stimulus M3 (target word one of string of isolated words): "Mountain yet and steams if mansion roams as *loops* stove magazine by list try yes kind if red stride risk."

Stimulus M4 (target word one of string of nonsense syllables): "Snoyn frane versh wames slove bist teffs thrane smabe rasht

skeev dripe djave tchoom shafe prev foops royts nang trouse twave."

Before hearing the last stimulus sentence, M5, Ss were informed that they now had a new task, and the following instructions were read to them: "You are to react to the letters used in spelling the words you hear. Of course, letters do not always have the same sound, and may even be silent." If S indicated that the task was not understood completely, he was given the example of the word "rough," which would qualify as a correct response if the target letter were "R," "O," "U," "G," or "H." As in Sentences M1 through M4, each group of nine Ss had a different target.

Stimulus M5 (target letters "K," "L," "O," or "P"; the first occurrence of each target letter in the sentence is underlined): "The sea remained quite calm, and the crew and skipper wanted a cooling breeze after the warm days and nights which had stayed with them during the entire cruise."

EXPERIMENT 1: RESULTS AND DISCUSSION

The delay between the occurrence of the target and the response was scored in terms of the number of phonemes rerecorded beyond the target before this recording was terminated by S's response. Thus, considering Sentence P3 with the target cluster /tr/, if the rerecording continued to include part of the last phoneme in the word "to" as follows "... patrons to," and S stated, following his response, that the target was in the word "patrons," then S would receive a score of 5 for that word. The terminal phoneme was detected readily upon playback, and each response was played back at least twice before the score was assigned.

The scores for both correct and incorrect responses for the practice sentences, P1 through P4, are shown in Table 2, which lists all median identification times for correct and incorrect responses. The identification times, scored in phonemic units, are given in parentheses following the response words. Directly below this, the number of Ss (N) responding to each word is given, as well as the corresponding percentage of the total Ss in the experiment. The sum of all responses scored for each target does not equal 36, since some Ss did not give

Table 2
Median Target Identification Times (in Phonemic Units) for Correct and Incorrect Responses
With Practice Sentences in Experiments 1 and 2

Sentence	Experiment 1 (36 Ss)			Experiment 2 (9 Ss)		
	Target Phonemes	Correct Word	Incorrect Words	Target	Correct Word	Incorrect Words
P1	/l/	While (6) N = 20, 55%	Applause (10) N = 7, 19%	"L"	While (5) N = 7, 78%	Applause (11) N = 2, 22%
P2	/n/	And (10) N = 3, 8%	Snow (8) Cold (7) N = 25, 69% N = 1, 3%	"N"	And (9) N = 2, 22%	Snow (7) N = 5, 56%
P3	/tr/	Patrons (9) N = 8, 22%	Artists (5) Century (12) N = 17, 47% N = 2, 6%	"TR"	Patrons (13) N = 5, 56%	Artists (5) N = 2, 22%
P4	/led/	Led (6) N = 33, 92%		"B"	Base (7) N = 9, 100%	

any response for particular targets. It should be emphasized that scores for different sentences may not be compared directly, since many factors could influence the ease of identifying the particular target sound or sounds. These include familiarity with the word containing the target, the number of phonemes and syllables within this word, the temporal duration of the scored phonemic units following the target, and the semantic and grammatical context provided by the rest of the sentence. With regard to this last factor, it is interesting that not only prior context, but subsequent context can influence identification of a phoneme (Warren, 1970).

If we turn from a comparison of numerical scores and consider the errors in detection made by Ss, some interesting preliminary observations can be made which may be worth further investigation. The results with these practice sentences will be discussed briefly before taking up the main sentences which were constructed to permit quantitative comparisons and conclusions.

In Sentence P1, the target occurred within the very first word. We can see from Table 2 that, despite this early occurrence, over half of the 36 Ss (N = 20, or 55%) detected the target with a median latency of six units for those who responded correctly. An additional 19% detected the second occurrence of the target phoneme, leaving 25% who did not respond to either the correct or a wrong target. No false responses due to the intrusion of spelling were observed in connection with the silent letter "L" between the phonemes /l/ of "while" and "applause" (i.e., the words "could" and "calm"). In P2, only 8% of the Ss responded to the first appearance of the target /n/ in the word "and," but 69% responded when it appeared for the second time, in the word "snow." This difficulty in identifying the target in "and" will be discussed again in Experiment 2. In Sentence P3, a misleading stimulus was

provided, since the target cluster /tr/ was preceded by the same phonemes in reverse order (i.e., /rt/ in the word "artists"). Almost half (47%) of Ss responded to the reversed order, despite the prior specific instructions that the target sounds /tr/ must occur in the order given. Most Ss realized their error directly after responding and told the E that their response was wrong. In Sentence P4, responses to the word "led" were quick and with very few errors. This good performance took place despite the fact that the sentence was intentionally constructed so that Ss would anticipate a homophone with a different syntactic function (the noun corresponding to the metal "lead" rather than the actual target, the verb "led").

A less informal experimental design is used for the stimuli M1 through M4. In each of these stimuli, a single four-phoneme word contains all targets of the four separate groups (i.e., the entire word, the single vowel, the consonant other than the ones in the initial and terminal positions, and the cluster consisting of that consonant and the vowel). This use of targets within the same stimulus word permits direct comparison of the time required to identify these targets of different phonetic classification and complexity. The experimental results are summarized in Table 1, which gives the median identification time for the first occurrence of the designated target. The four groups in Experiment 1 are identified by separate letters (A, B, C, or D) preceding the listing of the target. The identification time is expressed in phonemic units and is given in parentheses following the target listing.

With Stimulus M1 ("The final place described by the planning group as a possible region for the building of the stone dam had a firm foundation of bedrock"), the group having as target the entire word "stone" was aided by the fact that the context preceding the target was selected so that the occurrence of this

word would be anticipated by the listener. The speed with which Ss responded apparently reflects success in achieving anticipation of the stimulus. It can be seen from Table 1 that the median identification time was 0 (i.e., the response occurred during the /n/ in "stone"). The identification time was significantly longer ($p < .001$) for each of the three groups having targets within the word "stone" (Mann-Whitney U test for these and subsequent similar comparisons of groups). The groups with the targets /o/ and /to/ each required three phonemic units for response. The plosive consonant /t/ by itself had a median identification time of ∞ (i.e., more than half of the Ss gave no response). This difficulty in detection of a plosive consonant was encountered with other stimuli, as we shall see. The identification time for /t/ was significantly greater than for /to/ ($p < .05$), but not significantly different from /o/.

With Stimulus M2 ("The wind had howled since seven o'clock, and the brain was numbed by the noise and chill which could not be kept outside"), the target word was deliberately made improbable by prior context, although its presence was justified metaphorically by the subsequent context. The identification time for the word "brain" was three phonemic units beyond the last phoneme in the target word. As in the case of the vowel and the two-phoneme cluster in M1, identification of /c/ and /re/ each took three units longer than identification of the entire word, as did the "liquid" or "semivowel" /r/. The score for identification time for the entire word "brain" was significantly less ($p < .001$) than for each of the other three targets.

Stimulus M3 ("Mountain yet and steams if mansion roams as loops stove magazine by list try yes kind if red stride risk") was a string of isolated words with no contextual links between successive words, so that not only were prior semantic cues avoided (as in M2), but syntactic cues were lacking as well. Identification of the entire word

"loops" was significantly faster than was identification of each of the components ($p < .05$ for /up/, $p < .001$ for /u/ and for /p/). The cluster identification was faster than was identification of each of the components ($p < .005$ for /p/ and $p < .025$ for /u/). There was no significant difference between /p/ and /u/. We see that, for this word, identification time decreased consistently with increasing phonetic complexity—from individual phonemes, to cluster, to word. With the two previous stimuli (which were meaningful sentences), this smooth progression was not observed.

Stimulus M4 (Snoyrn frane verth wames slove bist teffs thrane smabe rasht skeev dripe djave tchoom shafe prev foops royts nang trouse twave") consisted of a string of nonsense syllables. As with M1 through M3, the entire syllable ("skeev") was significantly faster in identification time than the cluster and the two individual phonemes ($p < .001$ in each case). As with the other stimulus which was not a sentence, M3, the cluster, /ki/, was significantly different from its components: faster than /k/ ($p < .001$), faster than /i/ ($p < .005$). The vowel /i/ was significantly faster than the consonant /k/ ($p < .005$).

One of the unexpected observations in the main experiment thus far is that the plosive consonants in the second or third position of a four-phoneme syllable are identified with great difficulty unless they are included within a target containing a temporally contiguous vowel. This difficulty in identifying an isolated plosive does not seem to be found when the phoneme is in the initial position of a syllable. Foss (1969) studied identification time for /b/ which was always the initial sound of a word, and found that the time required to identify the occurrence of /b/ increased when the sound followed a word of low frequency in English. Foss and Lynch (1969) also used words beginning with /b/ and found that identification time was longer for self-embedded than for right-branching sentences. In these reports there was no evidence that Ss generally missed the plosive target, nor did our Ss miss in Experiment 2 (which will be discussed subsequently) when the target was /p/ at the beginning of a syllable (the second syllable in the word "skipper").

But, before moving to Experiment 2, there is one more sentence to discuss. The last stimulus in Experiment 1 tests spelling and was designed to determine if constituent letters could be identified during discourse and if a different identification system was involved than that employed for phonemic targets. Sentence M5 ("The sea remained quite

calm, and the crew and skipper wanted a cooling breeze after the warm days and nights which had stayed with them during the entire cruise") had a variety of targets with different relations to the acoustic nature of the corresponding stimulus. The target letter "L" in "calm" corresponds to a silent letter, so that identification required prior recognition of the word. The letter "K" required that Ss ignore three words having the phoneme /k/ which preceded the correct target word "skipper." The target letter "P" corresponded to the second plosive consonant in "skipper," but there were no preceding phonetically equivalent false targets as with the letter "K." The target letter "O" can have several pronunciations, so that it would be difficult for S to employ the sound of the letter in speech as a response cue without first identifying and spelling the word.

The success of Ss in identifying all letter targets in M5 as shown in Table 1 suggested a small additional study, Experiment 2, designed to compare identification of phonetic and letter targets in greater detail.

EXPERIMENT 2: PROCEDURE

Each of the nine Ss heard the following instructions: "This experiment deals with the time required to identify the spelling of spoken words. You are to press a key as soon as you hear a word spelled with a designated letter or group of letters. The sound of the word may be an unreliable guide to spelling—you should react to the designated letters regardless of how they are pronounced. The meaning of the words and sentences is of no importance. You will be asked nothing concerning the meaning, so concentrate solely on the spelling of the words heard. To summarize, the purpose of this experiment is solely to measure the speed with which you can identify target letters used in spelling the words you will hear."

Practice Sentences

All Ss heard the same sentences in the same order employed in Experiment 1, but the targets were as follows: Sentence P1, the letter "L"; Sentence P2, the letter "N"; Sentence P3, the letters "TR," in that order; and Sentence P4, the letter "B."

Main Sentences

Following the four practice sentences, Ss were informed that the main part of the experiment was commencing. Stimuli were the five sentences employed in Experiment 1, presented in the same order but with targets that were different from those of the earlier experiment (except for M4). All nine Ss in this experiment had the

same targets. These targets were: Stimulus M1, the letter "T" (except in the word "the"); Stimulus M2, the letter "R"; Stimulus M3, the letter "P"; Stimulus M4, the phoneme /i/; Stimulus M5, the phoneme /p/. Before Sentence M1, Ss were told to respond to the letter "T" in any word except "the." Before Stimulus M3, Ss were told they would hear a string of English words which did not form a sentence, and before Stimulus M4 they were told that they would hear nonsense words consisting of English speech sounds. They were also told before Stimulus M4 that they now had a new task, and the following instructions were read to them: "In the next part of the experiment, you are to react directly to speech sounds rather than their spelling. Press the key when you hear the designated phoneme or speech sound in a word. Remember, you are now to react solely to sounds in speech."

EXPERIMENT 2: RESULTS AND DISCUSSION

The results with letter targets for practice sentences are summarized in Table 2, which allows a direct comparison with the data appearing alongside for Experiment 1. When targets in the two experiments corresponded to the same stimulus item (Sentences P1, P2, and P3), median identification times for Ss were quite similar and may be considered equivalent. But if we compare the errors in identifying targets, spelling seems superior to speech sound search for these three sentences. A smaller percentage of Ss allow the target to go by without responding (P1 and P2), and there is a decrease in the tendency to react to a prior false target (less metathesis in P3). However, this superiority of letter identification does not reach statistical significance in any individual stimulus using a 2 by 2 contingency table (Finney et al, 1963).

It should be noted that the majority of Ss cannot identify either /n/ or "N" in the word "and," suggesting some special perceptual status for this very familiar conjunction.

The plosive consonants are missed less frequently when searched for as a letter rather than a phoneme. With Stimulus M1, /t/ was missed by six out of nine Ss, but "T" was missed by only one out of nine (this difference between Experiments 1 and 2 is significant at $p < .025$ by 2 by 2 contingency table). The requirement that "T" in the word "the" not be responded to was followed successfully by all Ss. With Stimulus M3, /p/ was missed by five out of nine, and "P" by only one out of nine (this difference just misses significance). The liquid consonant target in M2 was not

missed by any Ss, whether searched for as /r/ or as "R." Turning to medians of identification times as shown in Table 1, phonemic identification of /r/ was significantly faster than the "R" in spelling ($p < .025$ by Mann-Whitney U test). For the plosives of M1 and M3, the median of the relative times required for these types of search is reversed, as might be anticipated from the great difficulty in identification of these items as phonemes, with spelling faster than phonemic identification. However, these differences do not reach statistical significance for either sentence.

Stimulus M4 allows a direct comparison of the performance of the different Ss in the two experiments, since the phonemic target /i/ was identical for both groups. We see that the median identification times were identical (five units).

For the last stimulus, M5, identification of /p/ in "skipper" takes just the same time as when the target is the letter "P." It is interesting that /p/ was missed by only one S in "skipper" in Experiment 2, while in Experiment 1, six Ss missed identification of the same phoneme in "loops" ($p < .025$ by 2 by 2 contingency table). As discussed earlier, there is other evidence that the position of a plosive at the beginning of a syllable facilitates identification.

CONCLUSIONS AND SUMMARY

It seems that the comparison of identification times for targets within speech can be used to elucidate the temporal course of speech perception and the units or chunks employed in perceptual processing. The term, "identification time," seems preferable to "reaction time" since it seems doubtful if the sounds of speech can be perceived and reacted to directly. It appears that perceptual synthesis of speech into syllables must precede analysis into the component items. In all cases for which comparison can be made in the present study, identification time for single-syllable words or nonsense syllables was faster than the identification time for the component phonemes or phoneme clusters. As would be anticipated, identification time for words was decreased with prior semantic context leading to the increased probability of occurrence of the target word. However, the influence of grammatical context and of word frequency (rather unexpectedly) could not be demonstrated. Indeed, a nonsense syllable could be identified as readily as a word within a sentence. Once the word was

identified, the additional time required for identification of a given class of components was remarkably constant.

The detection of phonemes probably does not involve access to an echo-box short-term store for search directed towards the trace of speech sounds corresponding to the target phoneme, nor does it correspond to a teasing apart of joined features so that the sensory input corresponding to these components can be identified. It seems likely that there is a scanning of the intact syllable or cluster in a nondestructive search for structural "elements." An illusion already mentioned briefly, the phonemic restoration effect, is pertinent here. When a listener employs the context of a sentence to fill in the speech sound replaced by a cough, he cannot detect which phoneme was physically absent, even when he knows that a sound is missing and listens to the stimulus several times. This supports the hypothesis that phonemic components cannot be identified directly, but can only be inferred on the basis of higher organizational levels.

The great difficulty in identification of plosive consonants as individual phonemes when not located at syllable boundaries is quite interesting. Placing the plosive in the initial position of a syllable facilitates identification, as does extending the target to include both the plosive and a contiguous vowel. Also, identification is much easier if accomplished indirectly by changing the target from the phoneme to the corresponding letter used in spelling.

The trouble with direct acoustic identification of the plosives probably follows from the fact that these phonemes have no invariant acoustic descriptions; indeed, they can be described acoustically only in terms of the adjacent phonemes (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). It would appear that this lack of acoustic specificity causes perceptual difficulties when the listener organizes speech into syllables and then scans these syllables for the presence of these consonants. However, when the plosive is part of a more complex phonetic target including a vowel (e.g., /ki/), then it becomes possible for the listener to specify the auditory characteristics of his target in advance, a strategy which appears to simplify the perceptual task.

The separate nature of the letter identification and the phoneme identification systems is indicated by the following: (1) Location of letters corresponding to plosive consonants offers

no difficulty for stimuli in which direct identification of these phonemes is not possible for most Ss; (2) a silent letter (such as "L" in "calm") can be identified with ease; (3) a letter with several possible pronunciations (such as "O") can be identified readily; (4) the presence of false phonetic targets does not interfere with letter identification (i.e., with the target "K," three words pronounced with /k/—"calm," "quite," and "crew"—did not produce any false responses when they preceded the correct target in "skipper").

Finally, it should be emphasized that the findings in this study should be considered as preliminary results obtained with a new technique.¹ Only a small number of stimuli and targets have been investigated, and the Ss, consisting mainly of graduate students, staff, and visiting scientists, may not have been typical of other groups in their responses, especially in regard to spelling by virtue of their presumed high literacy.

REFERENCES

- FINNEY, D. J., LATSCHA, R., BENNETT, B. M., and HSU, P. *Tables for testing significance in a 2x2 contingency table*. London: Cambridge University Press, 1963.
- FOSS, D. J. Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning & Verbal Behavior*, 1969, 8, 457-462.
- FOSS, D. J., & LYNCH, R. H., JR. Decision processes during sentence comprehension: Effects of surface structure on decision times. *Perception & Psychophysics*, 1969, 5, 145-148.
- LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-461.
- MILLER, G. A. Decision units in the perception of speech. *IRE Transactions on Information Theory*, 1962, IT-8, 81-83.
- WARREN, R. M. Verbal transformation effect and auditory perceptual mechanisms. *Psychological Bulletin*, 1968, 70, 261-270.
- WARREN, R. M. Perceptual restoration of missing speech sounds. *Science*, 1970, 167, 392-393.

NOTE

1. After this paper was written, an article by H. B. Savin and T. G. Bever appeared ("The nonperceptual reality of the phoneme," *Journal of Verbal Learning & Verbal Behavior*, 1970, 9, 295-302), which is highly relevant to and in accord with the present study. They employed sequences of nonsense syllables, and reported that responses to an entire syllabic target took less time than responses to the initial consonant or medial vowel within the syllable.

(Accepted for publication August 31, 1970.)