

The influence on recognition of spoken words that are misperceived

WILLIAM P. WALLACE and JACQUELINE E. COLLINS
University of Nevada, Reno, Nevada

Two experiments were performed that involved an initial word-identification task in which the acoustic signal was degraded by either 50% or 60% compression of the recorded words. A control group was tested at the original recording rate (0% compression). The compression manipulation was successful in increasing the number of identification errors. During the second stage of the experiments, words were read at a normal rate. In Experiment 1, the subjects indicated whether each word had appeared on the prior list, and in Experiment 2, the subjects estimated the number of times each word had been read on the prior list. The false-negative rate decreased and frequency ratings increased as a function of the number of prior presentations (one, two, or three) of the target words. The most interesting result was that the effect of repetitions was present even for words that had been misperceived.

The issue addressed by the current research was the subsequent recognition of misperceived words. Misperceptions were manipulated by using a speech-compression technique to degrade the acoustical signal. The questions of interest in Experiment 1 were the following:

1. Will a word presented a single time on a study list, but misperceived, produce more positive recognition responses on a later test than a control word that was not presented on the study list?

2. Will words presented twice on a study list, but misperceived once or both times, and words presented three times on a study list, but misperceived once, twice, or three times, produce more positive recognition responses on a later test than words having fewer presentations but the same number of correct prior identifications?

3. If words related to "new" target words are misperceived on the identification list, will they still influence false recognitions to the new target words?

Experiment 2 was conducted to determine if presented words similarly influence frequency judgments (estimates of the number of times target words had been presented on the identification list).

There is experimental evidence to support the assertion that there are relatively sophisticated degrees of perceptual processing of stimuli that subjects are unable to identify or recognize (Bowers & Schacter, 1990; Fowler, Wolford, Slade, & Tassinari, 1981; Gardiner, 1988; Greenwald, Klinger, & Liu, 1989; Jacoby & Whitehouse, 1989; Marcel, 1983; Schacter, 1987; Watkins & Gibson, 1988). These researchers have described the provocative phenomena in terms such as unconscious memory, memory without awareness, or implicit memory. They generally share the interesting feature of a denial or failure in

response to a direct memory or a perceptual probe accompanied by subsequent evidence of behavior that shows an influence of the prior experience.

Two examples of this line of research are briefly described. Eich (1984) had subjects shadow (repeat aloud and verbatim) an essay presented on the right speaker of a pair of stereophonic headphones. Word-homophone pairs were presented through the left speaker, with the words selected to bias the less frequent meaning of the homophone (e.g., *taxi-fare*). A recognition test (with each word spoken during the test) for the homophones presented on the unattended channel indicated that positive responses to old words were not significantly higher than false-positive responses to new words. However, a subsequent spelling test revealed that subjects were more likely to spell the homophone as biased by the modifier (e.g., *fare* rather than *fair*) if it had been presented on the unattended channel.

Merikle and Reingold (1990) reported four experiments in which a stimulus (word or nonword) was flashed on a screen. On any given trial, an item or a blank field was presented. Subjects first indicated whether a stimulus had been presented and were then required to make either a forced-choice recognition response or a lexical decision. Merikle and Reingold found that even when subjects failed to detect presented stimulus words (the same result did not hold for nonword stimuli), their subsequent recognition and discrimination performance was above chance. Although we may need to worry about how subjects interpret the rather contradictory nature of this task (e.g., the subject first responds that nothing had been presented and then is forced to indicate which of two target words had been shown), this research also satisfies the basic condition for demonstrating that items previously experienced can influence subsequent performance in the absence of direct evidence for detection or memory of the prior experience.

Correspondence should be addressed to William P. Wallace, Department of Psychology, College of Arts and Sciences, University of Nevada, Reno, NV 89557-0062.

The purpose of the present research was to examine whether spoken words *not* identified correctly when presented for study influence subsequent recognition in a manner consistent with the effects that occur when they are correctly identified. The experiment reported here adapted a procedure used by Wingfield (1975) to degrade the speech signal by time compression. The goal was to generate a pool of experimental words that would be *misperceived*. That is, immediate word-identification reports should permit categorization of each critical word on the study list according to whether subjects perceived the word presented or perceived a word other than the one presented. Basically, the interest was in the following experimental question: If word stimuli such as "oak" and "sample" are presented, and a subject misperceives these stimuli as "hope" and "several," will this subject later correctly recognize the stimuli "oak" and "sample" as having been presented on the study list? We believe the answer to this question, and variations of it, have relevance for the understanding of the perceptual processing that takes place during study.

EXPERIMENT 1

Method

Design. Experiment 1 was conducted in two stages. The first stage involved a word-identification task. The second stage followed immediately and involved a recognition test in which the subjects had to indicate as each word was spoken whether that word had been presented on the word-identification list. All stimuli were presented auditorially.

During Stage 1, separate groups of subjects either listened to the list of stimulus words at the normal speaking rate or listened to a degraded version of the list. The stimuli were degraded by either 50% or 60% time compression. The normal frequency characteristics of the signal were preserved, but one group (50% compression) listened to the recorded words played back at double the original recording speed and one group (60% compression) listened to the recorded words played back at 2.5 times the original recording speed. The 50% and 60% compression levels were selected because word intelligibility is only moderately impaired in this range (Foulke & Sticht, 1969).

The word-identification list included words that were presented one, two, or three times, randomly dispersed throughout the list. It also included sets of three words that were associatively related to a critical target word that was not on the Stage 1 list but was on the Stage 2 list (e.g., *child*, *cry*, and *infant* for the target word *baby*). Words presented at normal speaking rates and the degraded stimuli presented for word identification provided opportunities to examine the relationships between number of presentations (of target words and words that are assumed to elicit target words) and subsequent recognition and between number of identifications (correct perceptions) and subsequent recognition.

Materials. The word-identification list consisted of 142 items. There were 3 filler words in the first 3 positions and 4 filler words in the final 4 positions. The remaining 135 positions were used for the critical items. Ninety were required for target words that would later appear on the Stage 2 recognition test. Fifteen of these words appeared a single time (15 positions), 15 appeared twice (30 positions), and 15 appeared three times (45 positions). The remaining 45 positions were occupied by words that were associatively related to specific target words. These target words appeared on the recognition test, but they were not presented on the Stage 1 list. There

were 15 of these target words, and for each, 3 associatively related words were presented on the word-identification list. Words were assigned to positions at random, with the restriction that a minimum of 5 words intervene between repetitions of any given word and between words from a common set of associates.

All of the stimuli were one-, two-, or three-syllable common words. There were two versions of the word-identification list. The two versions were different in the following ways with respect to the specific words assigned to the repetition categories and the converging associate categories:

1. For the repetition target words, a total of 60 words constituted the pool of items. These words were randomly assigned to four categories, and all target words appeared on the recognition test. The four categories represented presentation frequencies on the word-identification list, with 15 words presented zero times (new control words on the recognition test), 15 presented one time, 15 presented two times, and 15 presented three times. A second version of the identification list was created by randomly assigning each of these words to a different frequency category.

2. For converging associations, a total of 30 sets of words consisting of 1 target word and 3 associates constituted the pool of items. Fifteen of these sets were randomly selected to represent the experimental target words that appeared on the recognition test and the converging associates that appeared on the word-identification list. The target words for nonselected sets appeared on the recognition test as control words. For the second version of the word-identification list, converging associates that were not selected for the first version replaced those originally selected, thus reversing the roles of the experimental and the control target words on the recognition test.

The recognition-test list had 92 words. Sixty of the words, including words that had been presented 0, 1, 2, or 3 times during the word-identification task (15 words in each repetition category), provided the tests for the repetition conditions. Thus, 45 of these words were actually old words and 15 were new words. There were 30 additional new words on the test list to assess the effects of the prior presentation of converging associates on false-positive responses. Fifteen of these new words were "experimental" words, because sets of words associatively related to them were presented during the word-identification task. The remaining 15 were "control" words, since their sets of associatively related words were not presented during Stage 1. Two filler words that were not presented on the word-identification task occupied the first two positions on the recognition list.

Subjects. A total of 60 undergraduate students at the University of Nevada, Reno participated in the experiment as unpaid volunteers. Each subject was assigned randomly to one of six groups that were distinguished according to the version of the word-identification list the subject received and the extent of the signal degradation during word identification (0%, 50%, or 60% compression).

Procedure. All testing was done individually in a small laboratory room off the main corridor of the Psychology building. Each subject was provided with a brief general description of the experimental procedures and then completed a one-page self-rating form. Next, the word-identification task was explained, with no mention of the later recognition test. The subjects were told that they would hear a long series of words. After each word was spoken, the subjects were to say aloud the word they heard. They were informed that they would only have about 4 sec to give a response before the next word would be spoken.

The experimenter had an ordered list of the correct responses and made one of three notations by each item as the subject responded: a check for a correct identification, an "NR" for no response, or either a written out or an abbreviated spelling of the spoken misperception (e.g., "doctor" or "Dr." by the actual target word "answer"). The experimenter listened to the tape along with the subject and was aware that the words were more difficult

to understand at the compressed rates. However, the experimenter was naive regarding hypotheses about misperceived words and later recognitions. Also, it is not likely that an experimenter would be able to keep track of the various repetition-response contingencies with the list lengths and the response rates that were used in the present study.

At the completion of the word-identification task, the subjects were informed that they would next hear a long series of words, but this time they were to indicate as each word was spoken whether or not that word had been presented on the identification list. Again, they were informed that they would have about 4 sec to make a recognition decision.

The lists were recorded by a female speaker. The words for both the identification and the recognition lists were played back through stereophonic headphones with a Varispeech II audiocassette recorder. Varispeech II operates as a normal cassette playback for the 0% compression condition. For the 50% and the 60% compression groups, the Varispeech was adjusted to the appropriate compression setting, and the pitch correction was activated. The compression is achieved by a sampling process. A computer chip in the Varispeech recorder accomplishes the sampling of the speech signal, whereby very brief segments are alternately discarded and retained, with the retained samples abutted in time. It should be noted that the 4-sec response interval was constant across conditions and that only the stimulus words were compressed in the 50% and the 60% compression groups. All groups were given the recognition test under the normal 0% compression.

Results and Discussion

Word-identification list. During the first stage of the experiment, the subjects had to identify each word as it was presented. Identification accuracy was very high for the subjects who did not have the speech signal degraded. The 0% compression group had a total of 38 identification errors out of 2,840 opportunities (142 words \times 20 subjects), for an average of 1.3% errors. Degrading the signal by time compression impaired performance on the identification task. The 50% compression group averaged 14.6% errors, and the 60% compression group averaged 26.7% errors. More than half of the misperceptions were words that had either the same initial phoneme or a similar letter or sound structure as the appropriate target words. Only 7% of the total errors resulted from a failure to respond.

When a single misperception occurred for words actually presented two or three times, the misidentification tended to occur on the earlier presentations, with the effect consistent across the two speech-compression conditions. For $R_{2/1}$ words (words presented twice and correctly identified only once), 78% of the misidentifications occurred on the first presentation of the word. For $R_{3/2}$ words (words presented three times and misidentified on one occasion), 59% of the misidentifications occurred on the first presentation and 33% occurred on the second presentation. When the subjects made two identification errors ($R_{2/0}$ and $R_{3/1}$), the same misidentification word was repeated 50% of the time for the 50% compression group and 36% of the time for the 60% compression group.

The 50% and the 60% compression groups were compared with respect to the identification errors that were made, as a function of repetitions and relationships of the

words on the study list. The 0% compression group was excluded from the comparisons, since the error rate was so low in this condition. A 2×3 analysis of variance (ANOVA) was used to compare error rates with words, as a function of compression and repetitions. Since opportunities for errors increased with repetitions, the identification-error scores were converted to errors per opportunity with an arc sine transformation, as has been recommended for dealing with proportions (Winer, 1962). Error rates were higher for the 60% compression condition ($M = .28$) compared with the 50% compression condition [$M = .15$; $F(1,38) = 17.19$, $MS_e = .19$, $p < .05$], and there was a significant effect for repetitions [$F(2,76) = 6.50$, $MS_e = .05$, $p < .05$]. The repetition effect apparently resulted from a drop in the error rate to the items presented three times, as the average error proportions were .24, .23, and .17 for words presented one, two, and three times, respectively. The interaction was not significant.

A second indication that the subjects did a little better in identifying repeated words than words presented for the first time comes from a comparison of the 15 sets of words presented three times each (45 items) with the 15 sets of three converging associates (also 45 items). A 2×2 ANOVA demonstrated that more errors were made at 60% compression ($M = 11.92$) than at 50% compression [$M = 7.02$; $F(1,38) = 12.99$, $MS_e = 36.97$, $p < .05$]. Also, more errors were made to words in the converging-associate sets ($M = 11.25$) than to words in the repetition sets [$M = 7.70$; $F(1,38) = 36.48$, $MS_e = 6.91$, $p < .05$]. The interaction was not significant.

Recognition errors. The proportion of recognition errors to target words that had been presented on the identification list and to new experimental and control words are presented in Table 1. In addition, the correct-negative responses to new words are presented in the R_0 column. It is apparent from Table 1 that false-negative error rates declined as a function of the number of prior presentations of target words. False-positive error rates to new words were higher when they were preceded by earlier presentations of associatively related words ($CV-E$), compared with the $CV-C$ words for which the sets of converging associates were excluded from the identification list. Separate ANOVAs were used to test the significance of the apparent differences among false negatives and false positives.

Table 1
Proportion of False Negatives to Presented Words
and False Positives to New Words

Compression Group	False Negatives				False Positives	
	R_0	R_1	R_2	R_3	$CV-C$	$CV-E$
0%	.91	.43	.19	.09	.09	.13
50%	.85	.38	.25	.11	.13	.20
60%	.80	.44	.25	.14	.20	.28

Note—Subscripts denote number of presentations in Stage 1. Entries in the R_0 column are correct negative response rates to words that were not presented on the word-identification list.

A 3 × 3 ANOVA on the false-negative errors to words having one, two, or three presentations during Stage 1 revealed only one significant effect. The number of recognition errors varied as a function of the number of repetitions on the study list [$F(2, 114) = 109.84, MS_e = 2.83, p < .05$]. Differences as a function of prior signal degradation and the interaction between these variables were not significant. Separate comparisons between R_1 and R_2 , R_1 and R_3 , and R_2 and R_3 were all significant. There were more false negatives for words with one study presentation than for words with two study presentations [$F(1, 114) = 80.17$] or with three study presentations [$F(1, 114) = 216.26$], and there were more false negatives for words with two study presentations than for words with three study presentations [$F(1, 114) = 33.08, MS_e = 1.83, p < .05$, for all comparisons].

A 2 × 3 ANOVA on the false-positive errors to experimental words (their associates appeared in Stage 1) and to control words revealed significant main effects for both variables. The higher numbers of false positives with increasing levels of speech compression during the word-identification phase resulted in a significant main effect for the signal-degradation variable [$F(2, 57) = 5.49, MS_e = 6.96, p < .05$]. Also, there were significantly more errors to experimental words than to control words [$F(1, 57) = 7.50, MS_e = 3.61, p < .05$]. The interaction was not significant.

Degrading the speech signal in Stage 1 appeared to have a general effect of elevating the number of false-positive errors. That is, the poorer perception of the stimuli on which recognition decisions were based may have produced a general uncertainty and resulted in the subjects' establishing a lower criterion for accepting a word as old. This lower criterion might have compensated for poorer recognition of items actually presented earlier, accounting for the failure to find a difference between the 0%, the 50%, and the 60% compression groups in false negatives to repeated words. An analysis of the false-positive errors to the R_0 control words also revealed a significant effect for the time-compression variable, which is con-

sistent with this description [$F(2, 57) = 4.66, MS_e = 2.76, p < .05$].

Recognition-error contingencies. There were no real surprises in the performance of the subjects on the identification task and on the subsequent recognition task. The use of time compression to degrade the speech signal resulted in poorer identification of spoken words. Recognition performance was influenced by prior presentations of target words and of words related to target words.

The data of major interest involved whether presented target words that were not identified correctly in Stage 1 had any influence on recognition performance in Stage 2. The descriptions that follow involve contingency responses, hence caution should be exercised in interpreting these results because they contain elements of both subject selection and item selection. Items and subjects are not equally represented in the various contingency categories, such as the category of target words presented once and correctly identified versus the category of target words presented once and misidentified. Nevertheless, we believe that these data are informative, and the selection problems will be addressed in more detail at a later point.

For the following comparisons, the recognition errors are organized according to correct and incorrect identifications to target words during the word-identification task. Each target word has an actual presentation-occurrence number, either 0, 1, 2, or 3 actual presentations and 0 or 3 presentations of words associatively related to a later new test word. For each of these presentation frequencies, the designated word may be correctly identified from zero times up to the number matching the actual number of presentations. The proportion of false-negative errors and false-positive errors to the various contingency classifications are presented in Table 2. For comparison purposes, the R_0 column of correct-negative responses is repeated. The empty cells mean that the number of errors in that category were too few to provide meaningful comparisons (averaging fewer than one error for that category per subject). The number of identification errors for the

Table 2
Proportion of False Negatives and False Positives as a Function of Number of Correct Identifications

Compression Group	False Negatives									
	R_0	$R_{1/0}$	$R_{2/0}$	$R_{3/0}$	$R_{1/1}$	$R_{2/1}$	$R_{3/1}$	$R_{2/2}$	$R_{3/2}$	$R_{3/3}$
0%	.92 (20)	—	—	—	.42 (20)	—	—	.19 (20)	—	.09 (20)
50%	.85 (20)	.64 (20)	.61 (13)	—	.32 (20)	.28 (18)	—	.19 (20)	.08 (18)	.10 (20)
60%	.80 (20)	.64 (20)	.62 (17)	—	.36 (20)	.26 (20)	.20 (17)	.12 (20)	.07 (20)	.08 (20)
False Positives										
	$CV-C_0$	$CV-E_{3/0}$	$CV-E_{3/1}$	$CV-E_{3/2}$	$CV-E_{3/3}$					
0%	.09 (20)	—	—	—	.14 (20)					
50%	.13 (20)	—	.28 (14)	.14 (20)	.25 (20)					
60%	.20 (20)	—	.36 (17)	.28 (20)	.31 (19)					

Note—The entries in the R_0 column are correct negative response rates to new words. The numbers in parentheses indicate the number of subjects that made one or more responses for the contingency category. The subscripts indicate the number of presentations, with the number of correct identifications on those presentations indicated after the slash mark; for example, $R_{2/1}$ is the designation for those words that were presented twice but identified correctly on only one of the two occasions, and $CV-E_{3/2}$ indicates that only two of the three associates to a given target word were identified correctly in Stage 1. No entries are presented when the number of cases for that cell averaged fewer than one per subject.

0% compression group was too low to generate any contingencies other than in the categories in which the correct identifications matched the number of presentations.

There are two important points to note from the data in Table 2. First, there is the obvious relationship that recognition errors decline as a function of the number of prior presentations of the target words. The $R_{2/2}$ error rates are lower than the $R_{1/1}$ error rates, and the $R_{3/3}$ error rates are lower than the $R_{2/2}$ error rates. A comparison with Table 1 indicates that, in general, there were fewer errors with the more restricted sets (lower error rates with $R_{1/1}$ than with R_1 , $R_{2/2}$ than with R_2 , and $R_{3/3}$ than with R_3). Second, *there appears to be a very consistent influence on recognition errors of words that were presented on the word-identification list but that were not identified correctly.* For these comparisons, we must focus on the 50% and 60% compression groups for which sufficient numbers of misperceptions occurred. The correct-negative responses to words that were not presented (R_0) are higher than the false-negative responses to words that had been presented once ($R_{1/0}$) or twice ($R_{2/0}$) but were never correctly identified. Similarly, false-negative responses to words that had been presented once and were correctly identified ($R_{1/1}$) were higher than the false-negative responses to words that had been presented twice ($R_{2/1}$) or three times ($R_{3/1}$) but were correctly identified only once. The same pattern held for items presented twice and correctly identified both times ($R_{2/2}$), compared with items presented three times but correctly identified only twice ($R_{3/2}$). *The extra misperceived presentation of a target word during the study phase of the experiment appeared to make that target word more recognizable.*

There were 12 critical comparisons involving these contingency response sets in which words were different in the number of actual presentations but were equal in the frequency of correct identifications (e.g., R_0 vs. $R_{1/0}$ at 50% compression, $R_{1/1}$ vs. $R_{3/1}$ at 60% compression). Although the numerical differences in false-negative rates (or correct negatives for the R_0 sets) were not always large, they were consistently lower for contingency conditions that had additional misperceived presentations during study. In all 12 contrasts, the subjects were more likely to recognize a word as old for the condition in which the words had more actual presentations, although the number of correct identifications during the study presentation was equal in the conditions being compared. If these extra misperceived presentations had no effect, one would expect the numerical differences in the 12 comparison pairs to be equally divided with regard to favoring either the condition with the higher number of actual presentations (e.g., $R_{3/1}$) or the condition with the lower number of actual presentations (e.g., $R_{1/1}$). The probability that all 12 would favor one of the two categories is $p = .0005$.

The 12 critical comparisons were subjected to separate paired-sample t tests. These were regarded as *planned* comparisons, hence no adjustments were made to control the per experiment error rate (see Keppel, 1982). Thus, it should be noted that the probability of finding

one significant comparison by chance alone is quite high. The analyses of the transformed error proportions (arc sine transformations) revealed that only 3 of the 12 critical contrasts reached standard levels of statistical significance. With 50% compression, there was a significantly lower false-negative rate for $R_{1/0}$ than for R_0 [$t(17) = 2.17, p < .05$] and a significantly lower false-negative rate for $R_{3/2}$ than for $R_{2/2}$ [$t(17) = 3.49, p < .05$]. With the 60% compression, there was a significantly lower false-negative rate for $R_{3/1}$ than for $R_{1/1}$ [$t(16) = 2.23, p < .05$].

Stage 1 identification errors included a substantial number that were "structurally similar" to the appropriate target words with respect to sound and/or number of letters in common. This class of overt errors may take on special theoretical significance for the subsequent recognition task. It is possible at the time of initial identification that target words (or possibly even components of target words) may acquire some familiarity or frequency *implicitly*, that is, they may be elicited by the incorrect response word emitted on the identification task. Another possibility is that at the time of the recognition test, the test word may elicit the prior incorrect identification response *implicitly*, arousing a sense of familiarity. The familiarity could then mediate the recognition responses in a manner consistent with the data. If this type of implicitly based familiarity is responsible for the influence of misperceived identifications on subsequent recognition, then the effects should be most evident with the set of misperceptions that are structurally similar to their respective target words.

Three kinds of structurally similar misperceptions were considered: sound similarity, as judged from dictionary pronunciations; letter overlap, as defined by misperception words and target words sharing at least half of the same letters in approximately the same locations; and same initial phoneme. For each classification, misperception word and target word pairs were categorized as either high similarity (HS) or low similarity (LS). Of course, there was considerable overlap across these three classification procedures.

The total number of HS and LS misidentification errors for the $R_{1/0}$ words (words presented a single time and misperceived when presented) are shown in Table 3. The $R_{1/0}$ items provided the most meaningful contingency set for this analysis, since multiple misidentifications were not possible for words presented only a single time and subsequent false-negative errors were highest for this contingency set (compared with $R_{2/1}$ and $R_{3/2}$ sets). The data in Table 3 also indicate the number of subjects in each group who made both HS and LS errors and the proportion of false negatives associated with the different classification errors. Although it should be noted that the data in Table 3 are based on a relatively small proportion of the total responses made in these tasks, there did not appear to be any reliable variation in false-negative rates as a function of the similarity between the subjects' misperceptions and the actual target words. The largest differ-

Table 3
False Negative Rates to $R_{1/0}$ Words as a Function of "Similarity"
Between Misperception Responses and Target Words

	Similarity Classifications					
	First Phoneme		Letter Overlap		Sound Judgment	
	HS	LS	HS	LS	HS	LS
50% Compression						
Total misperceptions	19	22	28	19	20	26
False-negative rate	.69	.55	.65	.62	.53	.71
Number of subjects*	11	11	12	12	12	12
60% Compression						
Total misperceptions	29	44	40	37	19	42
False-negative rate	.65	.64	.61	.59	.55	.61
Number of subjects	15	15	15	15	13	13

*False-negative rates that appear in the table are based on the indicated number of subjects who made both HS and LS misperceptions.

ences appeared with the sound-judgment similarity classification, but the difference for the 50% compression group [$t(11) = -1.11$], the 60% compression group [$t(12) = -.44$], or the combination of these two groups [$t(24) = -1.16$] did not approach significance at $p < .05$.

The false-positive data did not permit tests that were comparable to the tests for repeated words. The contingency classifications for the associatively related sets of three words (assumed to elicit experimental test words implicitly during the word-identification phase) differentiated sets that were equal in the number of actual presentations of exemplars but were different in the number of exemplars that were misperceived. The only comparison similar to the false negatives with the repetition contingencies reported above was the potential contrasts between $CV-C_0$ and $CV-E_{3/0}$, in which control test words (those not preceded by presentations of associatively related words) are contrasted with experimental test words for which none of the prior associatively related words had been perceived correctly. Unfortunately, there were only 16 cases from the 600 opportunities (40 subjects in the 50% and the 60% compression groups and 15 exemplars per subject) that met the conditions for this contingency. The data did not permit a meaningful comparison to determine whether misperceived associates of recognition-test words influenced false-positive rates in the way that misperceived prior occurrences of recognition-test words influenced correct-positive response rates.

For the remaining $CV-E$ contingencies, sets differed with regard to the number of words that were correctly identified in Phase 1. In these cases, if a misperceived associate exerts the same influence on false-positive rates as does a correctly perceived associate, then there should be no differences across the various contingency classifications. There are six critical comparisons apparent in Table 2. The four pairwise comparisons involving the $CV-E_{3/1}$ contingencies were in the direction showing *more* false positives when only one of the three converging associates in a set was correctly identified during the study,

compared with when two or three were correctly identified. An arc sine transformation was applied to the proportions of false positives, and the six critical comparisons were tested with paired-sample t tests. The only significant comparison was between $CV-E_{3/2}$ and $CV-E_{3/3}$ with 50% compression [$t(19) = -2.42, p < .05$]. A significantly higher false-positive rate to experimental test words in the 50% compression group occurred in those cases in which all three eliciting words in a set were correctly perceived, compared with when only two of the three were correctly perceived.

Response contingencies and item/subject selection biases. The data reported on response contingencies are susceptible to potential item- and subject-selection biases. Although these problems cannot be eliminated, some data are available that suggest that selection factors were not critical. First, with regard to item selection, one may argue that the items that were misperceived on the word-identification task may have been inherently more difficult items. However, this selection bias might lead one to expect poorer recognition performance with additional misperceived presentations. That is, items presented twice and misidentified on one of the presentations might be viewed as more difficult items than those presented only once and correctly identified. Yet the results showed better recognition performance for words with the higher presentation frequencies even when the words were not correctly identified on the extra presentations.

To differentiate words according to recognition difficulty, items that served as the repetition target words were ranked on the basis of number of Stage 2 recognition errors made by the 0% compression group. These same words were then ranked on the basis of the number of Stage 1 identification errors made by the 50% compression group and on the number of Stage 1 identification errors made by the 60% compression group. Six separate Spearman rank-order correlation coefficients (recognition difficulty based on Stage 2 performance of the 0% group and identification difficulty based on Stage 1 performance of the 50% and the 60% groups) were computed for the combinations of the R_1 , R_2 , and R_3 words and the 50% and the 60% compression groups. The correlations ranged from $\rho = -.14$ to $\rho = +.16$. Words that were more difficult to identify when spoken under degraded conditions were not intrinsically more difficult to recognize.

Concerning a subject-selection bias, one could argue that subjects who make more identification errors are poorer at all tasks, including memory tasks, and that because of selection factors, these subjects contribute relatively more data to the item categories with identification failures (e.g., $R_{2/1}$ words compared with $R_{1/1}$ words). Note again that if this produces a bias, it works to underestimate the effect that an extra misperceived target-word presentation improves recognition accuracy. The first thing that each subject did at the start of the experiment was to complete a one-page self-rating questionnaire. This questionnaire asked the subjects to provide information on both their high school GPAs and their SAT or ACT

Table 4
Mean Frequency of Occurrence Estimates as a Function of Repetitions and Prior Presentations of Associatively Related Words for 60% Compression Group

Repetitions				Converging Associates	
R_0	R_1	R_2	R_3	$CV-C$	$CV-E$
.38	1.13	1.43	2.39	.43	.42

Note—Subscript designations are the same as in Experiment 1.

test-score performance. Although there was no follow-up check on the accuracy of the information reported by the subjects, Pearson correlation coefficients relating these measures of academic achievement to recognition and word-identification performance did not produce significant correlations for any of the listening groups.

EXPERIMENT 2

Experiment 2 was conducted as a further check on the influence of misperceived words on a subsequent memory task. A frequency-judgment test was used during Stage 2. Experiment 1 indicated that presentation frequency increased recognition accuracy even though words were misperceived on the extra presentations. Experiment 2 was designed to determine if this phenomenon could be replicated in a frequency-judgment task—that is, would extra misperceived target-word presentations increase frequency-of-presentation estimates to those target words? The same lists that were used in Experiment 1 were used in this experiment; however, during the critical test stage, the subjects had to estimate the number of times each word had been presented on the identification list. Only the 60% compression condition was included in this extension.

Method

The lists, materials, and procedures in this experiment were identical to those in Experiment 1. The only change occurred in the response requirements during Stage 2 of the experiment. Instead of a "yes-no" recognition response, the subjects were asked to indicate the number of times they thought each word had been presented on the identification list. The Stage 1 identification list was presented at the 60% compression level for all subjects in this experiment, since this was the optimal condition for producing misperceptions. An additional 20 undergraduate students at the University of Nevada participated in this study. None of these subjects were involved in Experiment 1.

Results and Discussion

The results will be reported briefly, since they were consistent with Experiment 1 in most every respect. For the identification list, the subjects had an average of 32.7% misperceptions. There were no major discrepancies between the performance of these subjects and the 20 subjects in the 60% compression condition from Experiment 1. There was a significant effect of repetitions on error rates [arc sine transformations; $F(2,38) = 9.72$, $MS_e = .03$, $p < .05$], and this appeared to result from a decrease at the R_3 category. Average error proportions were .33, .34, and .24 for words presented one, two, and three times, respectively. There were significantly more identification errors to the CV words than to the R_3 words [$F(1,19) = 7.26$, $MS_e = 12.81$, $p < .05$].

The mean frequency judgments as a function of number of prior presentations of target words (R_0 , R_1 , R_2 , and R_3) and converging associates are presented in Table 4. Frequency judgments increased as a function of number of repetitions, with the effect of repetitions significant [$F(2,38) = 30.84$, $MS_e = .28$, $p < .05$]. Separate comparisons between R_1 and R_2 , R_1 and R_3 , and R_2 and R_3 were conducted. The frequency judgments were higher for R_2 than for R_1 [$F(1,19) = 10.65$, $MS_e = .08$, $p < .05$], for R_3 than for R_1 [$F(1,19) = 78.15$, $MS_e = .20$, $p < .05$], and for R_3 than for R_2 [$F(1,19) = 42.65$, $MS_e = .22$, $p < .05$]. The only major discrepancy between this experiment and Experiment 1 occurred with the converging associates. The overall frequency judgments to $CV-E$ words and $CV-C$ words were virtually identical. It was only with the subset of $CV-E$ words, for which all three converging associates were identified correctly ($CV-E_{3/3}$), that the frequency estimates to $CV-E$ words ($M = .54$) exceeded the frequency estimates to $CV-C$ words ($M = .43$), but the difference was not significant.

The mean frequency judgments to words for the various repetition contingency categories based on Stage 1 identification responses are presented in Table 5. It is apparent from Table 5 that the frequency-judgment data are consistent with the recognition data from Experiment 1. In general, the frequency of presentation estimates increased as actual presentation frequency increased, even when the additional presentations were misperceived. There were 10 critical comparisons involving these contingency response sets. In all but the $R_{1/1}$ versus the $R_{2/1}$

Table 5
Mean Frequency of Occurrence Estimates as a Function of Number of Correct Identifications for the 60% Compression Group

Repetitions									
R_0	$R_{1/0}$	$R_{2/0}$	$R_{3/0}$	$R_{1/1}$	$R_{2/1}$	$R_{3/1}$	$R_{2/2}$	$R_{3/2}$	$R_{3/3}$
.38 (20)	.65 (20)	.70 (20)	.78 (13)	1.27 (20)	1.22 (20)	1.47 (16)	1.84 (20)	2.26 (20)	2.82 (20)
Converging Associates									
$CV-C_0$	$CV-E_{3/0}$	$CV-E_{3/1}$	$CV-E_{3/2}$	$CV-E_{3/3}$					
.43 (20)	—	.36 (20)	.35 (20)	.54 (19)					

Note—The numbers in parentheses indicate the number of subjects who made one or more responses for the contingency category. Subscript designations are the same as in Experiment 1.

comparison, the frequency estimates were higher for conditions that included misperceived presentations (e.g., $R_{1/0}$ vs. R_0). The empirical-chance probability that at least 9 out of 10 comparisons would favor one of the classification outcomes is .0215. Separate paired-sample t tests indicated that 3 of these differences were significant: for R_0 versus $R_{1/0}$ [$t(19) = 3.52, p < .05$], for R_0 versus $R_{2/0}$ [$t(19) = 2.71, p < .05$], and for $R_{2/2}$ versus $R_{3/2}$ [$t(19) = 2.12, p < .05$].

GENERAL DISCUSSION

The results of the present experiment showed repetition effects on recognition performance and frequency judgments even when repeated words were misperceived. The data are consistent with other reports in the literature showing that stimuli that are masked, that are too faint to be perceived, or that are nonrecognizable for some reason, influence later performance (e.g., Eich, 1984; Greenwald et al., 1989; Jacoby & Whitehouse, 1989; Marcel, 1983; Merikle & Reingold, 1990).

It has been demonstrated in a running recognition series that words associatively related to later new target words increase the false-positive recognition errors made to those target words (Bryant, 1990; Underwood, 1965, 1983; Wallace, 1968). One interpretation of this effect is that the new target words were elicited "implicitly" as associative responses when related words were presented, and, as a result, had a familiarity value greater than new control words. It is this implicit repetition or familiarity that accounts for subjects being more likely to decide that the target words were old.

Since a fair percentage of the misperceptions in the present experiment were words that either sounded like the target words or had similar letter patterns, it could be argued that the presented target words gained familiarity through the indirect route of implicit occurrence as part of the response to the specific word generated as the incorrect identification. However, there may be reason to question this interpretation. In the case most similar to the present procedures, Wallace (1968) found that sound- or letter-pattern similarity did not increase false-positive responses (an effect occurred only when subjects were given an orienting task that focused on letters or spelling). Also, data from Experiment 1 indicated that recognition errors were not systematically related to similarity between target words and overt errors emitted during the word-identification phase of the experiment. In summary, if word responses occur implicitly as aroused by perceived words (a postaccess mechanism), they would appear to do so normally on a semantic and associative basis, and not on a structural or sound basis. Very few of the misperceptions in the present experiment were words associatively or semantically related to the target words.

Recent theoretical accounts of spoken-word perception that are concerned with the on-line processing of the sound stimulus have incorporated a mechanism that could im-

part familiarity to several word candidates in addition to the word selected for response (Cole & Jakimik, 1980; Marslen-Wilson, 1987; Norris, 1986). For example, a word *cohort* is assumed to be activated with the initiation of the sensory signal corresponding to the spoken word (Marslen-Wilson, 1987). The size of the cohort and the degree of activation of word candidates constitute a dynamic process as the sound stimulus unfolds over time until the sensory data (and other available data) have reduced the cohort sufficiently for a subject to make an identification decision. If the eventual target word is a candidate from the active cohort in the early processing stages, then the necessary familiarity underlying the later recognition decision would have been established. Early in the processing of the sound corresponding to the letter string *sharp*, "sharp" is likely to be considered as a response candidate (certainly more than a control word such as "land"). Thus, even if the subject ultimately decides that the presented word is "several" or "church," the target word "sharp" may have acquired a measure of familiarity, and this process could be repeated on subsequent presentations, further incrementing the familiarity value of "sharp." According to this application of the model, the critical familiarity increments are a by-product of the *preaccess* processing that occurs. It should be noted that this line of reasoning also may lead to the expectation that similarity between misperceptions and target words (at least as classified according to sound or initial phoneme) should be relevant to subsequent recognition, provided that it is reasonable to "postdict" activation and duration of the target word in the cohort from the specific misidentification response emitted by the subject. Given the fast presentation rates during word identification and the dynamic nature of the cohort, this may not be realistic with the present state of development of the model.

REFERENCES

- BOWERS, J. S., & SCHACTER, D. L. (1990). Implicit memory and test awareness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 404-416.
- BRYANT, D. J. (1990). Implicit associative responses influence encoding in memory. *Memory & Cognition*, *18*, 348-358.
- COLE, R. A., & JAKIMIK, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, NJ: Erlbaum.
- EICH, E. (1984). Memory for unattended events: Remembering with and without awareness. *Memory & Cognition*, *12*, 105-111.
- FOULKE, E., & STICHT, T. G. (1969). Review of research on the intelligibility and comprehension of accelerated speech. *Psychological Bulletin*, *72*, 50-62.
- FOWLER, C. A., WOLFORD, G., SLADE, R., & TASSINARY, L. (1981). Lexical access with and without awareness. *Journal of Experimental Psychology: General*, *110*, 341-362.
- GARDINER, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, *16*, 309-313.
- GREENWALD, A. G., KLINGER, M. R., & LIU, T. J. (1989). Unconscious processing of dichoptically masked words. *Memory & Cognition*, *17*, 35-47.
- JACOBY, L. L., & WHITEHOUSE, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, *118*, 126-135.

- KEPPEL, G. (1982). *Design and analysis: A researcher's handbook*. Englewood Cliffs, NJ: Prentice-Hall.
- MARCEL, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology*, **15**, 197-237.
- MARSLER-WILSON, W. (1987). Functional parallelism in spoken word recognition. *Cognition*, **25**, 71-102.
- MERIKLE, P. M., & REINGOLD, E. M. (1990). Recognition and lexical decision without detection: Unconscious perception? *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 574-583.
- NORRIS, D. (1986). Word recognition: Context effects without priming. *Cognition*, **22**, 93-136.
- SCHACTER, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 501-518.
- UNDERWOOD, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, **70**, 122-129.
- UNDERWOOD, B. J. (1983). *Attributes of memory*. Glenview, IL: Scott, Foresman.
- WALLACE, W. P. (1968). Incidental learning: The influence of associative similarity and formal similarity in producing false recognition. *Journal of Verbal Learning & Verbal Behavior*, **7**, 50-54.
- WATKINS, M. J., & GIBSON, J. M. (1988). On the relation between perceptual priming and recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 477-483.
- WINER, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.
- WINGFIELD, A. (1975). Acoustic redundancy and the perception of time-compressed speech. *Journal of Speech & Hearing Research*, **18**, 96-104.

(Manuscript received August 13, 1990;
revision accepted for publication March 5, 1991.)