Effects on recognition memory of misperceived spoken words following two attempts at initial identification

MARK T. STEWART and WILLIAM P. WALLACE
University of Nevada, Reno, Nevada

Wallace and Collins (1991) reported that the number of presentations of spoken words influenced the words' later recognition even if the words had been incorrectly identified by subjects when presented originally. In the present study, that experiment was repeated, but subjects were allowed to make a "second guess" during the initial word-identification task. The original results were replicated; target words misidentified on both first- and second-guess opportunities still influenced later recognition responses. A mechanism of prerecognition activation of word candidates is discussed.

When the time that it takes to recognize a spoken word is measured, the point of recognition often occurs before the end of the sound stimulus has been reached (see, e.g., Marslen-Wilson, 1985, 1987; Marslen-Wilson & Tyler, 1980; Tyler & Wessels, 1985). Thus, the real time that it takes for a spoken word to run its course may be divided into two components: a *prerecognition* component, corresponding to the time from stimulus onset to the recognition point, and a *postrecognition* component, corresponding to the time from the recognition point to the offset of the sound stimulus, and perhaps continuing into the next word.

A general approach in the study of preaccess processing has been to terminate or mask the stimulus during the preaccess stage. This interruption or interference then prevents the subject from carrying the processing to the normal conclusion level that results in the subject's giving a clear and confident report about the stimulus. The end result is that the subject fails at some level to provide the necessary responses that would lead one to conclude that he/she perceived or was consciously aware of the presented stimulus. In a variety of experiments, it has been shown that in spite of a failure on a direct test of performance, subjects show the influence of the prior experience on a number of indirect measures (e.g., Bowers & Schacter, 1990; Eich, 1984; Greenwald, Klinger, & Liu, 1989; Jacoby & Whitehouse, 1989; Marcel, 1983; Merikle & Reingold, 1990). These phenomena have been summarized with such descriptive labels as "perception without awareness," "unconscious perception," "subliminal perception," and "implicit memory."

Wallace and Collins (1991) had subjects perform two separate tasks in their investigation of the spoken word recognition process. In the first task, subjects were asked to make a single identification response to words presented one, two, or three times under degraded auditory conditions. In the second task, a list was presented that included both "old" words from the identification list and "new" words. Auditory conditions were not degraded in the second task. After each word was spoken, the subjects had to indicate whether it had been presented on the previous list. The interesting result was a "misperceived repetition effect" (i.e., recognition memory accuracy improved as a function of earlier repetitions even when the words had *not* been correctly identified). The authors offered an explanation of this phenomenon that was based on the acquired "familiarity" of target words due to their implicit activation during early stages of processing.

The present research represents an attempt to replicate Wallace and Collins (1991) under conditions designed to evaluate an alternative postrecognition interpretation of their data. Their procedures allowed subjects to make only one identification response for each presented word. It is possible, however, that subjects may have "consciously" considered more than one word before emitting a specific word-identification response, thereby suppressing an alternative because of task requirements. If this occurred, and if the actual target word *was* considered and rejected, the postulated mechanism (acquired familiarity) would also fit a postrecognition description. Our primary goal in the present experiment was to determine whether the "misperceived repetition effect" could be replicated when the identification task permitted a second chance at identifying each spoken word.

METHOD

Design and Materials

In this experiment, we used the word lists from and followed the basic design of the experiment reported by Wallace and Collins (1991, Experiment 1). In the first task, subjects listened to a series of 142 words and tried to identify each word as it was presented. In the second task, the subjects listened to a series of 92 words and indicated whether each

Correspondence should be addressed to M. T. Stewart, Department of Psychology, University of Nevada, Reno, NV 89557.

word had been presented on the first list. The words on the first list were played back at either a 50% or a 60% rate of compression, making the task of identification somewhat more difficult than would normally be the case. Words on the recognition memory test list were played back at the original speaking rate.

The word-identification series consisted of 7 filler words that occupied the first 3 positions and the final 4 positions of the 142-item list. An additional 90 positions were required for *target* words that were tested for recognition in the second task. Fifteen target words were presented once for word identification, 15 were presented twice, and 15 were presented three times. Repeated items were separated from their prior occurrences by a minimum of 5 other words. The word-identification list also included words that were associatively related to a specific target word presented on the recognition list. There were 15 such target words, each with 3 associates that were presented on the word-identification list.

The materials consisted of one-, two-, or three-syllable common words. The recognition memory target words that actually appeared on the word-identification list were selected randomly from a pool of 60 words to create four groups of 15 words each. The 15 words in one group then served as the zero-repetition group ("new" control words on the recognition memory test), with the other three groups providing the words for the single-repetition condition, the double-repetition condition, and the triple-repetition condition. A second version of this initial list was created by randomly assigning each word set to a different repetition condition.

The lists also included word-association materials consisting of 30 sets of 4 words. Each set had a target word that appeared on the recognition memory test and 3 associates of the target word. Fifteen of these sets were randomly selected, and the 45 associates were presented on the word-identification list. Thus, the target words for these sets served as new *experimental* words on the recognition test, with the target words of the nonselected sets serving as new *control* words on the recognition memory test. In the second version of the word-identification list, the associates that were not selected for the first version replaced those originally selected, thus reversing the roles of the experimental and control words on the memory test.

Subjects

A total of 40 undergraduate students at the University of Nevada participated in the experiment. All subjects reported "normal" hearing and English as their first language. Each subject was randomly assigned to one of four groups that were based on the version of the word-identification list and the extent of the first-task signal degradation created by compressing the speech signal in time (50% or 60% compression).

Procedure

The two tasks were presented successively in a small, quiet laboratory room. For the first task, the subjects were told that they would hear a long series of common words, and that some words might be difficult to understand because they were "sped up." The task simply required subjects to say the word that they heard. The only departure from the procedure used by Wallace and Collins (1991) was that the subjects were now allowed to give a second response to identify another possible solution word they considered before they had decided on their first choice. This procedural change also resulted in a somewhat slower pace through the word-identification list (Wallace & Collins used a 4-sec rate). In the present experiment, after each word was presented, the subject made an identification response and was then prompted by the experimenter for a second response, resulting in an effective presentation rate of approximately 6–7 sec per item.

At the end of the word-identification task, the subjects were informed that they would hear a long series of words, but that this time they were to indicate, as each word was spoken, whether the word had been presented on the identification list. The stimulus words were not degraded in this presentation. The subjects were given 4 sec for each successive recognition memory decision.

The lists were recorded by a female speaker and played back through stereophonic headphones. The subjects heard a sample tape and were allowed to adjust the volume settings on their headsets to a comfortable level. A Varispeech II audiocassette recorder/player was used to degrade the speech signal. The Varispeech II was adjusted to either the 50% or the 60% compression setting, with the pitch correction acti-

vated. The compression is achieved by a sampling process, in which very brief segments are alternately discarded and retained, with the retained samples abutted in time. Finally, each experimental session was recorded, providing a reliability check for scoring responses.

RESULTS AND DISCUSSION

Word Identification

Not surprisingly, degrading the speech signal with the speech-compression procedure produced a reasonable number of word-identification errors. Excluding the filler words, there were 135 critical items on the word-identification list. The 20 subjects in the 50% compression group made a total of 479 identification errors out of 2,700 opportunities, yielding an error rate of 17.7%. On 58 of the occasions (12.1%) when a misidentification occurred, subjects made a correct response on the second try—that is, when they were reporting an alternative word that they had considered as a possible response. The 20 subjects in the 60% compression group made a total of 819 errors, yielding an error rate of 30.3%. On 80 of the occasions (9.8%) when a misidentification occurred, the subject made a correct response on the second try. The second guess following most of the first-guess errors (65.9%) was either a repetition of the specific first-guess error or no response.

A 2 × 3 analysis of variance (ANOVA) was used to compare error rates as a function of degree of speech compression and repetitions. Since opportunities for errors increased with number of repetitions, the identification-error scores were converted to errors per opportunity, and an arc sine transformation was applied following recommendations for dealing with proportions (Winer, Brown, & Michels, 1991). Error rates were higher for the 60% compression condition ($M = .26$) than for the 50% compression condition [$M = .12$; $F(1,38) = 60.00$, $MS_e = .08$, $p < .05$], and there was a significant effect for repetitions [$F(2,76) = 12.04$, $MS_e = .05$, $p < .05$]. The average proportions of errors as a function of repetitions were .24, .23, and .15 for words presented one, two, and three times, respectively. The interaction was not significant.

Subjects were better at identifying repeated words than words presented for the first time, and this was confirmed by contrasting the words presented three times (45 occurrences) with the 15 sets of three target-word associates (also 45 items). A 2 × 2 ANOVA demonstrated that more errors were made at 60% compression than at 50% compression [the mean error rates were .26 and .13, respectively; $F(1,38) = 28.70$, $MS_e = 22.44$, $p < .05$]. Also, more errors were made in response to the single occurring associates than to the repetitions [the mean error rates were .24 and .15, respectively; $F(1,38) = 119.65$, $MS_e = 2.51$, $p < .05$]. The interaction was not significant.

Recognition Memory Errors

The proportion of recognition errors in response to words repeated one, two, or three times on the identifi-

Table 1
Proportion of False-Negative Responses to Presented Words and False-Positive Responses to New Words

| Compression Group | Contingency | | | | | |
| | False Negative | | | | False Positive | |
	R_0	R_1	R_2	R_3	CV-C	CV-E
50%	.90	.44	.25	.07	.11	.16
60%	.84	.44	.24	.07	.18	.22

Note—R_0 are control words for R_1, R_2, and R_3, which refer to the number of repetitions on the identification list. CV-E, new experimental words; CV-C, their respective controls. Entries in the R_0 column are rates of correct-negative responses to words that were not presented on the word-identification list.

cation list (R_1, R_2, and R_3) and to new experimental words having converging associates (CV-E) to words presented on the identification list are presented in Table 1. The false-positive error rates for appropriate control words (CV-C) for the CV-E words are also presented. The control words (R_0) for R_1, R_2, and R_3 are presented in the first column, but it must be noted that the negative responses ("the word was not presented earlier") to words in this set are *correct* responses. These data appear quite consistent with the Wallace and Collins's (1991) data: the "false-negative errors" declined as a function of number of prior presentations, and false-positive error rates for responses to new words were higher when the new words had been preceded by presentation of associatively related words in the first list.

A 2 × 3 ANOVA of the false-negative errors revealed that recognition memory errors decreased as a function of increasing repetitions [$F(1,76) = 131.92$, $MS_e = 2.34$, $p < .05$]. Neither the compression level nor the interaction between repetitions and compression level had a significant effect on the number of false negatives.

A similar analysis was used to compare false positives to CV-E words and CV-C words. The only significant effect from that analysis resulted from the higher number of errors in response to the CV-E words [$F(1,38) = 6.57$, $MS_e = 1.60$, $p < .05$].

Recognition Error Contingencies

Wallace and Collins (1991) demonstrated that presentations of words on one or more occasions influenced subsequent recognition memory even when the words were incorrectly perceived when originally presented. The critical question in this experiment was whether this phenomenon was replicable when misperceptions for which subjects made a correct second guess were excluded from the comparison. Although this occurred for only about 10% of the errors, the exclusion of these responses had the effect of reducing all of the critical pairwise contrasts. The error-contingency data (errors given identification failures on both first and second guesses during the word-identification task) are presented in Table 2. For purposes of comparison, the R_0 column of correct-negative responses is repeated. Empty cells indicate that there were too few errors in that category to provide meaningful comparisons (averaging fewer than one error per subject).

The important contrasts involve those for which the number of correct identifications during the word-identification task (Stage 1) was constant but the number of actual Stage 1 presentations varied [e.g., words presented once and incorrectly identified ($R_{1/0}$) vs. words presented two times and incorrectly identified on both occasions ($R_{2/0}$)]. Wallace and Collins (1991) were able to make 12 of these critical comparisons: all 12 showed lower false-negative rates for the member of the pair with the higher actual presentation frequency, with 3 of the 12 contrasts significant at $p < .05$. Although in several cases the absolute differences in error rates were quite small, the overall consistency of the results seemed impressive (see Brown, 1991). A similar pattern of consistency was evident in the present data, but the overall results were less dramatic than in the previous study. Five relevant pairwise contrasts were possible with the 50% compression condition and 10 pairwise contrasts were possible with the 60% compression condition. Twelve of these 15 critical comparisons were consistent in showing that false-negative error rates were lower when extra "misperceived" presentations of the target words had taken place. The exceptions occurred for R_0 versus $R_{1/0}$ and $R_{1/1}$ versus $R_{2/1}$ at 50% compression and $R_{1/1}$ versus $R_{2/1}$ at 60% compression. If additional misperceived presentations had no effect, one would expect the numerical differences in the 15 comparisons 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.,

Table 2
Proportion (P) of False-Negative Responses and False-Positive Responses, and Number of Subjects Who Made One or More Responses, as a Function of Correct Identifications

| Repetition Contingency | Compression Group | | | |
| | 50% | | 60% | |
	P	No. Subjects	P	No. Subjects
	False Negative			
R_0	.90	20	.84	20
$R_{1/0}$	.92	20	.70	20
$R_{2/0}$	.78	12	.67	17
$R_{3/0}$			.42	11
$R_{1/1}$	.35	20	.31	20
$R_{2/1}$	.45	17	.37	20
$R_{3/1}$			.19	16
$R_{2/2}$	.15	20	.09	20
$R_{3/2}$	.10	16	.08	20
	False Positive			
$CV\text{-}C_0$	.11	20	.18	20
$CV\text{-}E_{3/0}$				
$CV\text{-}E_{3/1}$	.17	14	.16	19
$CV\text{-}E_{3/2}$	.11	20	.20	20
$CV\text{-}E_{3/3}$	.19	20	.27	20

Note—R_0, correct-negative response rates for new words. Subscripts indicate, first, the number of presentations, and second the number of correct identifications on those presentations. Words incorrectly identified on the first response, but correctly identified on the second response, were treated as correct identifications. Blank cells indicate that the number of cases in that contingency category averaged fewer than one per subject.

$R_{1/1}$). The "chance" probability that 12 or more of the 15 comparisons would all favor one of the conditions is .035.

The 15 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, 1991). The analyses of the transformed error proportions (arc sine transformations) revealed that only two contrasts were significant at $p < .05$. With 60% compression, the false-negative rate for $R_{3/0}$ was significantly lower than the correct-negative rate for R_0 [$t(10) = 2.37, p < .05$], and the false-negative rate for $R_{3/1}$ was significantly lower than that for $R_{2/1}$ [$t(15) = 2.39, p < .05$]. It may be of interest to note that if second-guess responses are ignored, 13 of the 15 comparisons show lower "false-negative" response rates for the category member with the higher number of actual presentations, and 6 of the pairwise contrasts reach the .05 significance level.

The false-positive data were analyzed to determine whether the error rates were higher when the associatively related words in the first list had been correctly perceived. Six comparisons were possible, and for five of these the false-positive rates were higher when a higher number of target-word associates had been correctly identified in Stage 1. The only exception occurred with the 50% compression groups and the contrast between CV-$E_{3/1}$ and CV-$E_{3/2}$. However, none of these comparisons among CV-E pairs for the 50% or 60% compression groups was significant.

The magnitude of the effect of misperceived presentations on later recognitions was reduced when the class of word-identification errors, for which subjects were able to make a correct target identification on a second guess, was excluded from the contingency analysis (approximately 10% of the errors in the present study). Overall, the pattern of results remained consistent with and appear to support Wallace and Collins's (1991) interpretation of a prerecognition mechanism of target-word activation.

REFERENCES

Bowers, J. S., & Schacter, D. L. (1990). Implicit memory and test awareness. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* 16, 404-416.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin,* 109, 204-223.

Eich, E. (1984). Memory for unattended events: Remembering with and without awareness. *Memory & Cognition,* 12, 105-111.

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. (1991). *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.

Marslen-Wilson, W. D. (1985). Speech shadowing and speech comprehension. *Speech Communication,* 4, 55-73.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition,* 25, 71-102.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition,* 8, 1-71.

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.

Tyler, L. K., & Wessels, J. (1985). Is gating an on-line task? Evidence from naming latency data. *Perception & Psychophysics,* 38, 217-222.

Wallace, W. P., & Collins, J. E. (1991). The influence on recognition of spoken words that are misperceived. *Memory & Cognition,* 19, 498-506.

Winer, B., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design.* New York: McGraw-Hill.