

## Direct comparison of auditory implicit memory tests

MAURA PILOTTI, ERIK T. BERGMAN, DAVID A. GALLO,  
MITCHELL SOMMERS, and HENRY L. ROEDIGER III  
*Washington University, St. Louis, Missouri*

In this experiment, we examined the degree to which four implicit tests and two explicit tests, all involving auditory presentation, were sensitive to the perceptual characteristics of the stimuli presented during study. Presenting stimuli visually decreased priming in all the implicit memory tests, relative to auditory presentation. However, changing voice between study and test decreased priming only in the implicit memory tests requiring identification of words degraded by noise or by low-pass filtering, but not in those tests requiring generation from word portions (stems and fragments). Modality effects without voice effects were observed in cued recall, but the opposite pattern of results (voice effects without modality effects) was obtained in recognition. The primary new finding is the demonstration that auditory memory tests, both explicit and implicit, differ in their sensitivity to the perceptual information encoded during study.

A primary concern in implicit memory research is the degree to which implicit memory tests are dependent on the perceptual characteristics of the information encoded during study. Perceptual or data-driven implicit memory tests tap into recent experiences incidentally (i.e., in the absence of instructions to recollect) by requiring subjects to identify stimuli that have undergone some form of degradation. Most research in this area has used implicit tests that employ verbal materials presented in a visual format. For example, subjects who have recently read the word *window* might be asked, in an ostensibly unrelated task, to complete a word stem (win\_) or fragment (w\_n\_o\_) with the first word to come to mind (word stem completion and word fragment completion) or to identify a word degraded by a mask (perceptual identification). In each case, priming is defined as subjects' identifying or completing previously studied words more often than non-studied words.

In general, visual implicit tests like those described above are sensitive to study-to-test changes in the perceptual characteristics of the studied words. For instance, Rajaram and Roediger (1993) found that four implicit memory tests (perceptual identification, word stem and fragment completion, and anagram solution) showed a decrement in priming following study-to-test changes in salient perceptual dimensions, such as modality of presentation (auditory study and visual test) and format (picture study and verbal test; see, also, Weldon, 1991).

In the last several years, researchers have extended the study of implicit memory to the auditory domain. Thus, a number of auditory implicit tests have been developed that are analogous to visual implicit tests: auditory perceptual identification (Church & Schacter, 1994; Goldinger, 1996; Schacter & Church, 1992; Sheffert, 1998b), auditory stem completion (Schacter & Church, 1992), and auditory fragment completion (Habib & Nyberg, 1997). In identification tests, subjects are required to identify words degraded by either white noise or low-pass filtering, whereas in stem or fragment completion tasks, portions of the original speech signals are replaced by silence, and the subjects' task is to produce the items from these tokens. Not surprisingly, research using these auditory implicit tests has been concerned with the extent to which such tests are sensitive to perceptual information. To this end, in several studies, the question was examined of whether priming would be affected by study-to-test modality changes, such as those involving visual presentation at study and auditory presentation at test (Habib & Nyberg, 1997; Jackson & Morton, 1984), and study-to-test voice changes, such as those occurring when a word is spoken by a female at study and then by a male at test (Church & Schacter, 1994; Goldinger, 1996; Jackson & Morton, 1984; Schacter & Church, 1992; Sheffert, 1998b). For the most part, these studies showed that both modality and voice changes reduce but do not eliminate auditory priming, relative to a condition in which such changes do not occur.<sup>1</sup> These findings suggest that, on auditory implicit tests, subjects utilize perceptual information, such as information about the modality of presentation (auditory vs. visual) and acoustic characteristics (e.g., perceptual details of a speaker's voice), as well as lexical information.

Although previous research suggests that auditory implicit tests depend on perceptual characteristics in addition to lexical information, it is unclear whether the con-

---

We thank Mark Yen for his help in conducting the experiments and Sandy Hale and Erik Bergman for recording the stimuli. Correspondence concerning this article should be addressed to M. Pilotti, Department of Psychology, Washington University, One Brookings Drive, Campus Box 1125, St. Louis, MO 63130-4899 (e-mail: mpilotti@artsci.wustl.edu).

tribution of perceptual information is uniform across different tests. For instance, Schacter and Church (1992; Church & Schacter, 1994) found that study-to-test voice changes produced a larger decrement in priming in stem completion and identification of low-pass-filtered words than in the task of identification of words masked by noise. Alternatively, Sheffert (1998b) found significant effects of changing voice between study and test in both identification tasks, but only when words were presented in a degraded form at study (see, also, Goldinger, 1996). In contrast to experiments that introduced within-modalities changes, studies that changed modality of presentation between study and test (visual to auditory) found more consistent effects across different auditory implicit tests. For instance, Habib and Nyberg (1997) observed less priming on an auditory word fragment completion test following visual presentation, relative to auditory presentation. Jackson and Morton (1984) found the same pattern of results in the test of identification of words embedded in noise. These experiments, however, differ on many other dimensions besides the auditory implicit tests used, thus making cross-experiment comparisons difficult to interpret.

To overcome the difficulties in comparing findings from different studies, we examined the perceptual sensitivity of four auditory implicit tests in a single experiment: identification of words masked by noise or low-pass filtered and stem and fragment completion. The primary purpose of this investigation was to examine the effects of study-to-test modality and voice changes. Directly comparing four implicit memory tests that have the same subjects, materials, and encoding manipulations permits an evaluation of the notion that auditory implicit memory tests might differentially engage perceptual information. This notion is of theoretical importance because existent models of implicit memory phenomena (see Goldinger, 1998; Schacter & Church, 1992) predict that perceptual changes will affect auditory repetition priming, but they do not predict that the effects of these changes will be a function of the type of implicit memory tests used.

To examine the extent to which auditory implicit tests involve the processing of perceptual information, the experiment reported here included four conditions of presentation between study and test (visual, same voice, different voice, and nonstudied) and four auditory implicit memory tests (identification of low-pass-filtered words, identification of words presented in noise, word stem completion, and word fragment completion). The goal was to examine whether all four tests would be differentially affected by study-to-test modality changes and by study-to-test voice changes. The same-voice condition served as a reference for measuring the effects of these changes. It was hypothesized that any decrement in priming following study-to-test changes in modality, voice, or both would indicate that the incidental retrieval operations engaged by the implicit memory tests selected for this experiment rely on perceptual information (modality and/or voice) in addition to lexical information.

The encoding task remained the same across all the speech perception tests. We thought that an encoding task involving deep processing of the study words would be especially desirable in light of Schacter and Church's (1992) finding that semantic tasks, such as category judgments or pleasantness ratings, consistently produced slightly, albeit nonsignificantly, stronger study-to-test effects of voice changes than did a nonsemantic task, such as pitch ratings. Therefore, we had subjects rate the extent to which they knew the meaning of the study words on a 1–7 scale. We also used a larger stimulus set (300 words) than the one used by Schacter and Church (48 words), to increase the power to detect any effects.

Finally, we employed two explicit memory tests, cued recall and recognition, to examine whether retrieval operations in explicit and implicit tasks could be differentiated on the basis of their sensitivity to perceptual information. Cued recall served as the explicit version of the stem completion task, whereas recognition served as a rough explicit counterpart of the identification tasks.

Previous studies on the role of modality or voice change effects suggest that explicit and implicit tasks cannot be reliably dissociated on the basis of their sensitivity to perceptual information. For instance, Schacter and Church (1992) found that after a semantic encoding task (pleasantness and meaning ratings), changing voices between study and test significantly decreased priming in stem completion, whereas the same changes produced smaller (nonsignificant but consistent) decrements in cued recall and recognition performance. Goldinger (1996) found voice effects in both recognition memory and priming in the task of identification of words embedded in noise, with deleterious effects on both tasks from a change of voice between study and test. (The data from the short delay condition in Goldinger's Experiment 2 are discussed here; see also Palmeri, Goldinger, & Pisoni, 1993; Sheffert, 1998a, 1998b). Similarly, other researchers have found that modality changes hurt recognition memory (Hintzman, Block, & Inskip, 1972; Kirsner, 1974), but these outcomes have not been consistently replicated (e.g., Craik, Moscovitch, & McDowd, 1994). In the present experiment, using auditory tests, we observed the effects of changing voice and modality within a single experiment, so comparisons of implicit with explicit tests should be more secure.

## METHOD

### Subjects

Two hundred twenty-eight undergraduates from Washington University participated in the experiment in partial fulfillment of course requirements. Thirty-six subjects were included in each test condition, with the exception of 48 subjects in the test of identification of low-pass-filtered words (to increase the power for this test).

### Procedure and Design

The experiment was presented as an investigation of auditory and visual perception, consisting of a series of brief tasks. It involved three phases: study, a 5-min break, and test. During study, the subjects were presented with three blocks of 50 trials (50 visually pre-

sented words, 50 words spoken in a male voice, and 50 words spoken in a female voice). The order of block presentation was counterbalanced across subjects in a Latin-square design. Each trial consisted of a 500-msec blank screen, a 150-msec warning signal, a visually presented number labeling the trial, a 500-msec blank screen, and the presentation of a stimulus word. A trial ended and the next trial began 8 sec after stimulus presentation. On any given trial, the subjects were asked to listen to or read a word and to rate its familiarity on a 7-point scale (from 7, *you recognize the word and are sure that you know the meaning of the word*, to 1, *you have never seen or heard the word before*). During study, words were always spoken in the clear at 75 dB (SPL) over headphones.

At test, the subjects first heard 20 filler stimuli for practice and then 300 randomly presented stimuli (150 studied and 150 nonstudied). All the stimuli were spoken in either the male or the female voice of the previous phase. The presentation of words was identical to the one described for the study phase, with the exception that the subjects initiated each trial by pressing the Enter key of the computer keyboard. In the implicit tests, the subjects were told that they would participate in a speech perception test that required either (1) identification of degraded words or (2) completion of word stems or fragments. Stimulus degradation was accomplished by either low-pass filtering words at 1 kHz or presenting words in continuous white noise at a signal-to-noise (S/N) ratio of -5 (the noise was 5 dB louder than the signal). Stems were obtained by isolating the initial syllable of each word,<sup>2</sup> and fragments were created by removing portions of the speech signal and replacing them with equivalent amounts of silence (see Habib & Nyberg, 1997). The subjects were to report their answers in a booklet with 320 numbered blanks. In the explicit tests, the subjects were given the same booklet to fill out, but with different instructions. In the stem completion test, the subjects, who were presented with stems, were told that their task was to recollect the words presented during study. If a stem was the beginning of a word encountered at study, their task was to write down the word in the booklet. Alternatively, if they thought that a stem did not belong to a studied word, they were to report an X in the same booklet. In the recognition test, the subjects, who were presented with words spoken in the clear, were asked to recollect the words of the study phase and to recognize those words among others never presented before. The subjects were to report their answers (Y for old and N for new) in the booklet. In both the recognition and the cued recall tests, the subjects were instructed not to guess.

A mixed factorial design was used. Study-to-test condition (visual, same voice, different voice, and nonstudied) was the within-subjects factor, and test type (identification of words embedded in noise, identification of low-pass-filtered words, stem completion, fragment completion, recognition, and cued recall) was the between-subjects factor.

**Stimuli**

The stimuli consisted of 300 bisyllabic and trisyllabic English words. Twenty additional words of the same general type were chosen to serve as fillers at the beginning of the test list. The stimulus words were of either low or medium frequency ( $M = 9.4, SD = 14.8$ ; Francis & Kučera, 1982) and had a mean familiarity rating of 6.7 ( $SD = 0.57$ ; Nusbaum, Pisoni, & Davis, 1984).

All 320 stimuli were recorded by two speakers, one male and one female, and were digitized at a sampling rate of 20 kHz on an IBM-compatible computer using a 16-bit analog-to-digital converter equipped with anti-aliasing filters. The amplitude levels of all the stimuli were digitally equated, using a software package specifically designed to modify speech waveforms. Intelligibility of words spoken in the clear was 99%.

The stimuli were organized in two study lists of 150, matched for frequency and familiarity. Each study list was then subdivided into three sets of 50 words, also matched for frequency and familiarity.

Three versions of each set were devised: printed words, words spoken in a male voice, and words spoken in a female voice. The test list included 20 filler stimuli, placed at the beginning of the list, and 300 stimuli (150 studied and 150 nonstudied). One version of the test list was spoken in a male voice, and the other in a female voice. A Latin-square design was used to counterbalance study lists, study conditions (visual, female voice, and male voice), and speakers across subjects, so as to create 12 unique combinations of study and test lists. The experiment lasted approximately 1 h.

**RESULTS**

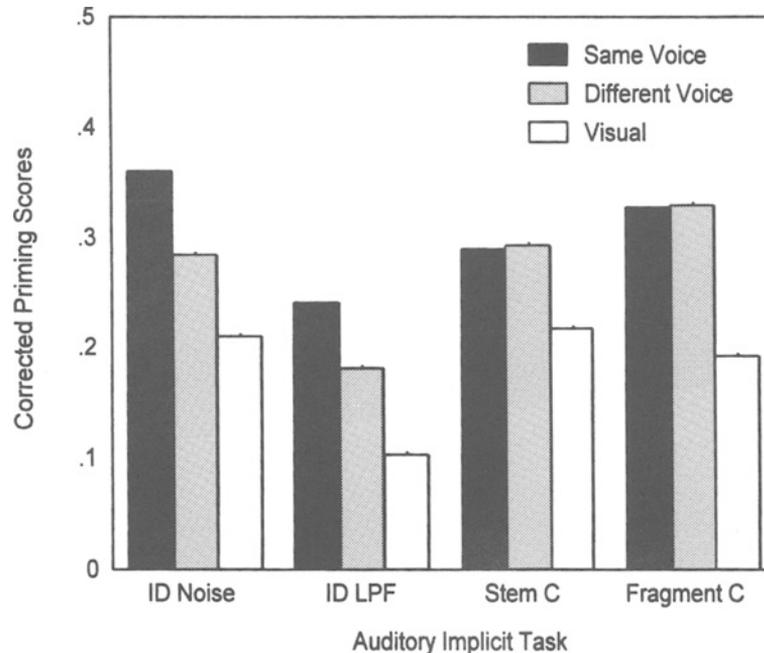
**Implicit Tests**

The proportions of studied and nonstudied words correctly identified/completed in each implicit test are shown in the top of Table 1. For each test, these proportions were submitted to a one-way repeated measures analysis of variance (ANOVA), with study-to-test condition (visual, same voice, different voice, and nonstudied) as the factor, and this analysis was followed by least significant difference (LSD) comparisons. LSD values (in Table 1) were used to determine the presence of priming and to assess

**Table 1**  
**Data from Six Auditory Memory Tests: Proportions Correct and Priming Scores for Auditory Implicit Tests and Proportions Correct and Proportions of False Alarms and Intrusions for Auditory Explicit Tests, in Four Study-to-Test Conditions**

Study-to-Test Condition	Proportion	Priming	LSD
Identification in Noise			
Same voice	.66	.17	
Different voice	.62	.13	.04
Visual	.58	.09	
Nonstudied	.49		
Identification of Low-Pass-Filtered Words			
Same voice	.48	.17	
Different voice	.44	.13	.03
Visual	.39	.08	
Nonstudied	.31		
Word Stem Completion			
Same voice	.48	.21	
Different voice	.48	.21	.03
Visual	.43	.16	
Nonstudied	.27		
Word Fragment Completion			
Same voice	.72	.13	
Different voice	.71	.12	.03
Visual	.67	.08	
Nonstudied	.59		
Recognition			
Same voice	.89		
Different voice	.84		.03
Visual	.89		
Intrusions	.08		
Cued Recall			
Same voice	.55		
Different voice	.54		.03
Visual	.51		
Intrusions	.05		

Note—The last column displays the least significant differences (LSDs) of the comparison between priming scores.



**Figure 1.** Relative priming scores as a function of study-to-test condition (same voice, different voice, and visual presentation) and type of auditory implicit test (ID noise = identification of words embedded in noise, ID LPF = identification of low-pass-filtered words, Stem C = stem completion, and Fragment C = fragment completion).

whether modality or voice changes between study and test had a detrimental effect on priming (all the differences reported here are significant at the  $p < .05$  level). Performance in each implicit memory test was analyzed separately because the different baselines generally obtained in these tests make cross-test statistical analyses problematic (see Rajaram & Roediger, 1993).

**Identification of words embedded in noise.** The proportion of items identified correctly differed across conditions [ $F(3,105) = 36.22$ ,  $MS_e = 0.01$ ]. All the conditions displayed significant priming effects, with priming being greatest in the same-voice condition (.17), less when voice changed between study and test (.13), and least when modality changed between study and test (.09).

**Identification of low-pass-filtered words.** The overall ANOVA again revealed a significant effect of condition [ $F(3,141) = 63.09$ ,  $MS_e = 0.001$ ], and all the study conditions produced significant priming effects. As in the previous test, priming was greatest in the same-voice condition (.17), less in the different-voice condition (.13), and least in the visual condition (.08).

**Stem completion.** The effect of study-to-test condition [ $F(3,105) = 75.59$ ,  $MS_e = 0.001$ ] and the priming effects in each study condition were significant. In this task, however, equivalent priming was obtained in the same- and different-voice conditions (.21 priming in both conditions), although a change in modality did reduce the priming effect (.16 priming).

**Fragment completion.** The effect of study-to-test condition and the priming effects in each study condition

were significant [ $F(3,105) = 35.93$ ,  $MS_e = 0.001$ ]. As in stem completion, roughly equivalent priming was obtained in the same- and different-voice conditions (.13 and .12, respectively), although a change in modality reduced priming (.08).

The finding that both modality and voice changes did not eliminate priming in these implicit memory tests suggests that priming relies partly on abstract lexical information. Furthermore, the fact that modality changes had a detrimental effect on priming in all these tests indicates that priming is sensitive to salient perceptual information, such as modality of presentation. The results from these four tests, however, also indicate that auditory implicit tasks can be dissociated on the basis of their reliance on voice information. Voice changes had a detrimental effect on priming in tasks requiring identification of degraded stimuli (identification of low-pass-filtered words or words embedded in noise), whereas such changes did not reduce priming in tasks requiring generation from parts of words (stem and fragment completion).

In Figure 1, we present the relative priming scores of each task, in which priming is expressed as a function of the amount of possible priming (priming divided by 1.00 minus the baseline; see Snodgrass & Feenan, 1990; Rajaram & Roediger, 1993) to take base rate differences into account. The data in Figure 1 fully corroborate our conclusion, based on raw priming scores, that shifts in modality affected all four auditory implicit tests but that changes in voice affected only tests involving perceptual degradation of the entire stimulus.

### Explicit Memory Tests

The proportions of words recognized and recalled in the various conditions are shown at the bottom of Table 1. These data were submitted to a one-way repeated measures ANOVA, with study-to-test condition (visual, same voice, or different voice) as the factor, followed by LSD comparisons.

**Recognition.** The main effect of condition was significant [ $F(2,70) = 5.65$ ,  $MS_e = 0.001$ ]. The difference between the same- (.89) and the different- (.84) voice conditions was also reliable, although recognition performance in the visual and same-voice conditions did not reliably differ. Performance was, however, near ceiling, with .89 hits in both conditions.

**Cued recall.** A significant effect of study-to-test condition was obtained [ $F(2,70) = 3.76$ ,  $MS_e = 0.001$ ]. LSD comparisons showed a modality effect, with both auditory study conditions producing greater recall (.55 and .54) than the visual condition (.51). The effect of voice was not significant.

Recognition and cued recall were similar to their implicit counterparts (perceptual identification and stem completion) in terms of the effect of voice. However, recognition differed from all the other tasks in not showing a modality effect, although this outcome has been reported by others with visual tests (e.g., Challis et al., 1993; Rajaram & Roediger, 1993).

## DISCUSSION

The primary findings to emerge from this experiment were that (1) visual presentation produced less priming than auditory presentation on all four auditory implicit tests, and (2) changing voices between study and test affected the two implicit tests in which whole words were presented in degraded form (low-pass-filtered words and words embedded in noise), but did not significantly decrease priming on the two tests in which only parts of words were presented at test (completion of word stems and fragments). These disparate findings across test types did not result from differences in the nonstudied baselines against which priming was assessed (see Table 1). Word identification in noise and word fragment completion, two tasks that had relatively high and similar baselines, were differentially sensitive to voice information. Similarly, identification of low-pass-filtered words and word stem completion, the tasks that had lower and roughly equivalent baselines, were also differentially sensitive to voice information.

The modality effects observed in all the implicit memory tests are consistent with earlier findings on auditory tests (see Habib & Nyberg, 1997; Jackson & Morton, 1984) and with findings of modality effects in visual implicit tests (e.g., Rajaram & Roediger, 1993). Such modality effects are consistent with several theoretical accounts of repetition priming in data-driven implicit memory tests. For instance, the presemantic perceptual

representation system account proposed by Schacter and Church (1992), the transfer-appropriate processing account proposed by Roediger, Weldon, and Challis (1989), and the episodic account proposed by Hintzman (1986) and Goldinger (1998) all predict that these data-driven implicit tests will show the modality effects observed in our experiment.

The dissociations based on voice changes that were observed in these implicit tests, however, are not directly implied by any of these accounts. The central question is why tests that involve degrading whole words show voice effects, whereas tests in which parts of words are omitted do not show similar effects. One idea is that stems and fragments are data-limited cues—that is, they provide listeners with less of the original speech signal than do the degraded words in identification tasks. In fact, although the original signal is altered in identification tasks, the word token is still a complete unit. Fundamental frequency is the primary cue for voice identification. This cue is not disrupted by low-pass filtering or by the addition of noise to the speech signal. Therefore, the two tests that preserved more voice information (fundamental frequency) showed voice effects. If our speculation is true, why did Schacter and Church (1992) find robust voice effects in stem completion? One major difference between our study and theirs is that Schacter and Church used a small test set (48 words) spoken by six speakers, whereas we used a large set (300 words) spoken by one speaker. Although data-limited cues may provide less of an opportunity for activating the perceptual information encoded at study than do complete word units, they still preserve some perceptual information (e.g., pitch). In an acoustically heterogeneous test set (i.e., fragments or stems produced by many speakers), the speakers' voices become salient features of the test material, thereby leading subjects to rely on voice information to generate words from incomplete tokens. In an acoustically homogeneous test set, as in our experiment, voice information is not a salient feature of the test material, and this may encourage subjects to use abstract lexical information as the primary means for generating words from incomplete tokens. Of course, given the differences in the two designs, another possible reason for this discrepancy in the findings is that our single test voice doubtless became quite familiar over the course of the experiment, relative to an experiment in which six different voices and far fewer stimuli were used. Future work is needed to unravel these possibilities.

The voice effects observed in the other implicit tasks also are not always consistent with earlier findings, which again could be due to task differences that may matter (note, however, that the experiment we report here is much more powerful than standard experiments in the literature, in terms both of the number of items used and of the number of subjects tested). The voice effects obtained in the task of identification of low-pass-filtered words replicate Church and Schacter's (1994) findings,

but those observed in the task of identification of words embedded in noise are difficult to reconcile with earlier findings. Schacter and Church (1992), Jackson and Morton (1984), and Sheffert (1998b) found only nonsignificant decrements in priming from study-to-test voice changes in this task, although all of them did observe such decrements. One plausible reason for the inconsistency between our results and those of Schacter and Church and of Jackson and Morton is that our S/N ratio was probably higher ( $-5$ ) than the one used in their studies (see baseline performance for evidence of the S/N ratio). As the level of background noise increases, it could be argued, voice-specific information would be drowned out and become largely unusable at test (see Schacter & Church, 1992, for a similar argument).<sup>3</sup>

The discrepancy between Sheffert's (1998b) results and ours, however, cannot be explained in the same manner. In her study, the S/N ratio was also  $-5$ . One major difference between these experiments is that Sheffert used high-frequency words, whereas we used low- and medium-frequency words. Therefore, the difference between our results and Sheffert's (1998b) findings may be that high-frequency words require less processing during encoding than do low- and medium-frequency words; if so, high-frequency words may provide less opportunity for encoding perceptual information and, therefore, would be less likely to reveal effects of voice. Looked at another way, additional cues may be needed to identify low-frequency words, and voice helps serve this purpose. This explanation is supported by Sheffert's (1998b) finding that words presented in noise at study produced larger study-to-test voice effects on priming than did words presented in the clear in identification tasks involving either noise or low-pass filtering. Her findings suggest that presenting words in noise at study enhances the processing of the perceptual features of high-frequency words, thereby improving the encoding of voice information that is later displayed on identification tasks.

In our experiment, we also investigated whether explicit tasks, such as recognition and cued recall, would show effects that were similar or dissimilar to those from the implicit memory tests selected for comparison. We found that modality, but not voice changes, lowered both cued recall and the priming effects observed in stem completion. This parallel effect is not surprising. Explicit tests that depend on data-limited cues often show perceptual effects caused by the need to identify the cue before it can be used to guide a meaning-based search (e.g., Challis et al., 1993; Weldon, Roediger, & Challis, 1989).

The recognition results are more perplexing. We found that shifting modality between study and test did not affect recognition, whereas changing voice between study and test reduced recognition. Therefore, recognition memory was dissociated from the implicit memory tests that involved completion of word portions and from explicit word stem cued recall, both of which revealed modality but not voice effects. The voice effects observed in the recognition task replicated earlier findings that suggest that voice information is useful in auditory recognition

memory (Goldinger, 1996; Palmeri et al., 1993; Sheffert, 1998a, 1998b). However, unlike some earlier findings (Hintzman et al., 1972; Kirsner, 1974) that showed modality effects in recognition memory, we found no evidence of an effect of modality shifts between study and test. Similar findings have been reported by Craik et al. (1994) and Challis et al. (1993). Although the issue of which conditions foster modality effects in recognition memory must await future resolution, it is worth noting that these data agree with a growing body of evidence indicating that perceptual variables can differentially affect priming on perceptual implicit memory tests and recognition memory. Some variables, such as modality, affect priming and not recognition memory in the same experiment, whereas other variables (such as size or left/right reflection of pictures) have no effect on priming and yet have large effects on recognition memory (e.g., Rajaram, 1996). These outcomes pose interesting puzzles with respect to the relation between perceptual priming and recognition memory (see Rajaram & Roediger, 1997, for a discussion).

In sum, our experiment revealed varying effects of modality and voice on six tests. Implicit tests were dissociated from one another on the basis of voice changes, whereas explicit tests were dissociated from each other by both voice and modality shifts. These findings suggest that understanding the differences among auditory tests (and the experimental conditions instantiated in the tests, such as single or multiple speakers) is a prerequisite for comprehensive theories of auditory priming.

## REFERENCES

- CHALLIS, B. H., CHIU, C., KERR, S., LAW, J., SCHNEIDER, L., YONELINAS, A., & TULVING, E. (1993). Perceptual and conceptual cueing in implicit and explicit retrieval. *Memory*, *2*, 127-151.
- CHURCH, B. A., & SCHACTER, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *20*, 521-533.
- CRAIK, F. I., MOSCOVITCH, M., & MCDOWD, J. M. (1994). Contributions of surface and conceptual information to performance on implicit and explicit memory tasks. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *20*, 864-875.
- FRANCIS, W., & KUČERA, M. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- GOLDINGER, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *22*, 1166-1183.
- GOLDINGER, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251-279.
- HABIB, R., & NYBERG, L. (1997). Incidental retrieval processes influence explicit test performance with data-limited cues. *Psychonomic Bulletin & Review*, *4*, 130-133.
- HINTZMAN, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411-428.
- HINTZMAN, D. L., BLOCK, R. A., & INSKEEP, N. R. (1972). Memory for mode of input. *Journal of Verbal Learning & Verbal Behavior*, *11*, 741-749.
- JACKSON, A., & MORTON, J. (1984). Facilitation of auditory word recognition. *Memory & Cognition*, *12*, 568-574.
- KIRSNER, K. (1974). Modality differences in recognition memory for words and their attributes. *Journal of Experimental Psychology*, *102*, 579-584.
- NUSBAUM, H. C., PISONI, D. B., & DAVIS, C. K. (1984). Sizing up the

- Hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research in speech perception* (Progress Rep. No. 10). Bloomington: Indiana University Press.
- PALMERI, T. J., GOLDINGER, S. D., & PISONI, D. B. (1993). Episodic encoding of voice attributes and recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 309-328.
- PILOTTI, M., GALLO, D. A., & ROEDIGER, H. L., III (in press). Effects of hearing words, imaging hearing words, and reading on auditory implicit and explicit memory tests. *Memory & Cognition*.
- RAJARAM, S. (1996). Perceptual effects on remembering: Recollective processes in picture recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 365-367.
- RAJARAM, S., & ROEDIGER, H. L., III (1993). Direct comparison of four implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 765-776.
- RAJARAM, S., & ROEDIGER, H. L., III (1997). Remembering and knowing as states of consciousness during retrieval. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 213-240). Mahwah, NJ: Erlbaum.
- ROEDIGER, H. L., III, WELDON, M. S., & CHALLIS, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3-41). Hillsdale, NJ: Erlbaum.
- SCHACTER, D. L., & CHURCH, B. A. (1992). Auditory priming and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 915-930.
- SHEFFERT, S. M. (1998a). Contributions of surface and conceptual information to recognition memory. *Perception & Psychophysics*, **60**, 1141-1152.
- SHEFFERT, S. M. (1998b). Voice-specificity effects on auditory word priming. *Memory & Cognition*, **26**, 591-598.
- SNODGRASS, J. G., & FEENAN, K. (1990). Priming effects in picture fragment completion: Support for the perceptual closure hypothesis. *Journal of Experimental Psychology: General*, **119**, 276-296.
- WELDON, M. S. (1991). Mechanisms underlying priming on perceptual tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 526-541.
- WELDON, M. S., ROEDIGER, H. L., III, & CHALLIS, B. H. (1989). The properties of retrieval cues constrain the picture superiority effect. *Memory & Cognition*, **17**, 95-105.

## NOTES

1. One exception to this claim might be when word identification occurs in noise. Although Schacter and Church (1992) argued that the task of identification of words embedded in noise did not produce voice effects, in the majority of their experiments using such a task, the same-voice condition produced more priming than did the different-voice condition. Jackson and Morton (1984) and Sheffert (1998b) also found a nonsignificant decrement in priming as a function of study-to-test voice changes in the same test. The replication of this pattern across different experiments appears to us to be a good reason for concluding that identification of words embedded in noise is sensitive to study-to-test voice changes. However, voice effects are generally small.

2. We isolated the first syllable of each of the words making up our stimulus set. To ensure that each word-initial syllable (stem) was unambiguously unique, in some isolated cases, we added the sound of the phoneme immediately following the syllable.

3. Recent data collected in our laboratory (Pilotti, Gallo, & Roediger, in press) support this conclusion. Following a clarity-rating task of study words spoken in the clear, we found that an S/N ratio of -5 produced reliable voice effects (proportion of correctly identified words: same voice = .61; different voice = .57; nonstudied = .44), whereas an S/N ratio of -10 did not (proportion of correctly identified words: same voice = .28; different voice = .28; nonstudied = .21).

(Manuscript received September 15, 1998;  
revision accepted for publication April 19, 1999.)