Voice-specificity effects on auditory word priming

SONYA M. SHEFFERT

Indiana University, Bloomington, Indiana

This research explores the nature of the memory traces that support spoken word identification. Specifically, do voice-specificity effects in implicit memory depend on information in a perceptual representational system or, alternatively, on the similarity of study and test exemplars? Memory for words and voices was assessed with two perceptual identification tests—the identification of words in noise and the identification of low-pass filtered words—after two encoding conditions (identification of words in noise and of words in the clear). At test, a word was presented in the same voice as at study or in a different voice. The data from the two experiments showed that study-to-test changes in voice reduced priming and that voice-specificity effects were greatest when the type of processing engaged at study overlapped with that required at test. Taken together, the results implicate the goodness of the processing match between encoding and test as the primary determinant of voice-specificity effects on perceptual identification tests and support the hypothesis that both voice and word information is represented within a single episodic memory system.

The aim of this research was to explore the nature of the memory traces that support spoken word identification. A specific goal was to determine the extent to which a speaker's voice (i.e., acoustic information other than information about the identity of a word) influences performance on spoken word identification tests.

One method for assessing a listener's incorporation of voice information into the memory for word occurrences is to use an implicit memory task. Unlike explicit tasks, which require subjects to consciously recollect a study episode, implicit tasks do not directly query subjects about a previous event. Memory for the experience is tapped by measuring the effects of *priming*, a facilitation or bias in responding to a test item that is attributable to information acquired in a previous study episode. Unconscious memory tapped by these procedures is called implicit memory (for reviews, see Moscovitch, Goshen-Gottstein, & Vriezen, 1994; Roediger & McDermott, 1993; Schacter, 1987).

Details about the processes or systems that underlie performance on implicit tests of verbal memory can be revealed by manipulating the physical format of the to-be-remembered item between the study and test phases of an experiment. Unlike the experiments that employ radical shifts in the form of presentation between study and test (i.e., cross modality), which consistently find effects of specific transfer, studies varying format within modality

(e.g., voice) produce less consistent effects. As a result, there is considerable uncertainty at the empirical level, making it difficult to adjudicate between the competing theoretical views. In the following section, I review several key studies from the auditory domain that vary voice across word repetitions and describe the two major theoretical perspective used to account for the findings.

Voice-Specificity Effects on Perceptual Identification Tests

Several experiments report that priming effects on implicit tasks depend on voice matches across study and test items. For example, Church and Schacter (1994) found that study-to-test changes in voice lowered identification performance of low-pass filtered words. Their experiment had three phases: a study phase and two test phases (one implicit, one explicit). Specifically, words were studied in the clear; during the implicit test, they were low-pass filtered, and the subjects' task was to identify words, guessing if necessary. Immediately after this test, the subjects were given a word recognition task that used words that were the same as those in the identification test but without filtering. Priming on the filter identification test was reduced by voice mismatches, whereas recognition was indifferent to this manipulation. Conversely, a second experiment showed that depth of processing at study affected recognition but resulted in identical priming scores.

A multiple memory systems view (Schacter, 1987; Squire, 1987; Tulving, 1972) accounts for this double dissociation between variables affecting implicit and explicit memory tasks by proposing that implicit and explicit tasks tap functionally and anatomically distinct memory systems. According to Schacter (1992; see also Tulving & Schacter, 1990), explicit memory task performance is mediated by explicit (or episodic) memory, which does not contribute to priming. Verbal implicit memory tasks depend on two presemantic representational systems (PRSSs) that repre-

The research presented in this article was supported by NICHD Grant HD-01994 to Haskins Laboratories, NICHD Grant DC00012 to Indiana University, and an APA Science Directorate Dissertation Research Award. The author thanks Carol Fowler, Jay Rueckl, Donald Shankweiler, Gwen Gustafson, David Pisoni, Geoffrey Loftus, Robert Greene, and an anonymous reviewer for comments and suggestions. Correspondence should be addressed to S. M. Sheffert, who is now at the Department of Psychology, Central Michigan University, Mt. Pleasant, MI 48859.

—Accepted by previous editor, Geoffrey R. Loftus

sent word forms but not meanings. During encoding, two representations of a word form are computed. One is an abstract symbolic phonological representation, stripped of its surface characteristics and stored in the left hemisphere. The second representation is a perceptual representation of the word, devoid of meaning and stored in a right-hemisphere PRS. Each subsystem makes an independent contribution to word priming. The fact that priming is not completely eliminated when the study and test materials are presented in different formats supports the inference that one memory system underlying performance preserves abstract information. In addition, the finding that changes in perceptual format attenuate auditory priming is taken to mean that a format-specific system also underlies priming.

However, other studies show that the magnitude of auditory priming is not attenuated by study-to-test changes in voice. Jackson and Morton (1984) report equivalent repetition priming effects for test words spoken in a voice that was either the same as or different from that in the study condition. A priming phase of their experiment consisted of semantic judgments of words presented in either a male or a female voice. Test words were same-voice or different-voice repetitions of studied words, and implicit memory was assessed by measuring identification accuracy to words presented in white noise. Jackson and Morton found that word repetition effects were independent of the voice in which words were spoken. They interpreted their results as evidence that auditory priming is mediated by a lexical system that is indifferent to voice information.

Schacter and Church (1992) offer a different interpretation of Jackson and Morton's (1984) findings. They suggest that the absence of a voice effect on the white noise task reflects special features of the perceptual identification test. Like Jackson and Morton, Schacter and his colleagues (Schacter & Church, 1992; Schacter, Church, & Treadwell, 1994) have been unable to find evidence of voice-specificity effects on implicit memory tasks that use white noise. Schacter and Church suggest that voice-specific priming is based on a right-hemisphere subsystem and speculate that the presence of white noise eliminates right-hemisphere contributions to auditory word priming, resulting in equivalent priming on the same-voice and different-voice trials.¹

Although the findings reviewed so far are consistent with the notion that voice-specificity effects reflect the workings of an auditory PRS, a study by Goldinger (1992, 1996) calls into question this conclusion. Goldinger used perceptual identification to examine implicit memory for spoken words. The study task was to identify monosyllabic words presented in noise. After a delay period, the subjects heard the same words in a new order; however, at test, half of the words were presented in a voice that was the same as that used during study, and half were presented in a different voice. Goldinger found that identification of words in noise was faster and more accurate when words were repeated in the same voice than when the voice was different. Furthermore, Goldinger found no evidence of a dissociation between implicit and explicit memory performance, in that voice matches improved recognition memory for spoken words (Craig & Kirsner,

1974; Geiselman & Bjork, 1980; Palmeri, Goldinger, & Pisoni, 1993; Sheffert & Fowler, 1995).

It is implausible that a disruption in the contribution of a right-hemisphere PRS could be responsible for the null findings on tests of word identification in white noise (Jackson & Morton, 1984; Schacter & Church, 1992), given that Goldinger (1992, 1996) found robust voice effects on this task. An alternative theoretical account to the multiple memory systems proposal postulates a single episodic memory system within which a variety of processes operate. The transfer appropriate processing view (see, e.g., Jacoby & Brooks, 1984; Kolers, 1979; Kolers & Roediger, 1984; Morris, Bransford, & Franks, 1977) assumes that events establish unique episodic traces in memory, including a record of the processing carried out during encoding, and that words are identified by direct similarity to prior episodes. Memory performance is a consequence of the degree to which retrieval conditions reinstate or match processing operations performed at encoding. What qualifies as a processing match between study and test varies, because different memory tests engage different processing procedures and, thus, recover different information. For example, an implicit task such as perceptual identification is considered *data driven* (Roediger, 1990), because it is sensitive to the degree to which the stimulus input at test resembles the stimulus input from study. Repeating a word in the same rather than a different sensory format will increase the ease or *fluency* with which the item is processed (Jacoby, 1983), so that the listener will be more likely to identify a degraded version of the word. In addition, redundant dimensions across episodes may coalesce and emerge as abstract lexical/semantic properties (see Hintzman, 1986) that mediate word identity priming.

A careful comparison of the experimental designs of Goldinger (1992, 1996) and of Schacter and Church (1992) reveals an important methodological difference that may account for the discrepant perceptual identification findings. In Goldinger's experiment, the masking noise was present at study and test, whereas the subjects in Schacter's experiments (see also Jackson & Morton, 1984) studied the words in the clear. According to Goldinger, and consistent with a processing account of format specificity effects, "the masking stimulus [should be] encoded along with all other perceptual and conceptual details. Thus, repetition of the exact stimulus/mask combination during a later test might be expected to yield greater repetition effects" (1992, p. 77). Failure to find voice-specificity effects can be explained by assuming that the relations between processing operations at study and at test were poorly matched. The importance of this design difference is evaluated in the following experiments.

EXPERIMENT 1

The present study was designed to address a discrepancy in the literature concerning the effects of study-to-test changes in voice on the perceptual identification of words in noise. As discussed above, Schacter and Church (1992) and Jackson and Morton (1984) have failed to find evidence

of voice effects on this task, in contrast to the positive findings of Goldinger (1992). In order to determine whether the null effects found by Schacter and Church were due to a disruption in the contribution of the right-hemisphere PRS or, instead, were due to a mismatch of the exact stimulus/mask combination at test, the subjects identified words embedded in noise or identified words spoken in the clear. The words were spoken by six speakers. The subjects then completed one of two implicit perceptual identification tests: identification of words in white noise (Experiment 1) or identification of low-pass filtered words (Experiment 2). In each test, an equal number of studied and filler items was presented, with an equal number of old words repeated in a voice that was the same as or different from that in the study phase. The dependent measure in each task was accuracy.

According to the memory systems view, sensitivity to study-to-test changes in voice should be observed to the extent that an implicit task is able to access the perceptual form of a spoken word in a right-hemisphere PRS subsystem. After repeated failures to obtain voice effects on implicit tasks that used noise (Schacter & Church, 1992; Schacter et al., 1994), Schacter and Church have proposed that noise interferes with the contribution of the right hemisphere to language processing. If this proposal is accurate, voice effects should not be observed when an implicit task uses noise to degrade test items, regardless of the study task.

The process view predicts a different pattern of effects. Because data-driven implicit tests are sensitive to the physical features of a word, study-to-test changes in voice should reduce the benefits of prior study inasmuch as the processing operations by which the physical features of the word are analyzed at study are not completely recapitulated at test. If the traces that support word identification include information about the processing operations required to encode the stimulus, different encoding tasks may produce differences in the magnitude of voice-specific retrieval. In particular, the processing view predicts that voice-specific transfer should be enhanced when both the encoding and the test task focus the subjects' attention on perceptual features of the word—for example, by the use of a noise transformation that requires more extensive bottom-up analysis of the input signal. Therefore, repetition effects should be enhanced on the perceptual identification in noise task when the word was also studied in noise, as compared with when the word was studied in the clear.

Method

Materials. The stimulus materials consisted of 250 high-frequency monosyllabic words derived from Kučera and Francis (1967) and the Modified Rhyme Test (House, Williams, Hecker, & Kryter, 1965). Target words were recorded in citation form using a list intonation (the final pitch was flat) by six different speakers (three males and three females) in a sound-proof booth. After the words were recorded, they were digitized at a 10.4-kHz sampling rate and equated for peak amplitude. Each word was recorded by every speaker, and, from these items, 20 experimental tokens (present on both the study and test tapes) and 20 control fillers (present only on the test tape) were randomly selected from each speaker. Of the 20 experimental tokens, 10 were designated as *same-voice* repetitions, and 10 as

different-voice repetitions. Ten additional tokens were randomly selected from among the six speakers (four speakers contributed two items; two speakers contributed one item) to serve as primacy and recency buffers for the study list.

Two study tapes of 130 words (120 target words, five primacy and five recency buffer items) were created, one presenting words in white noise and a second presenting words in the clear. Following Goldinger (1992, 1996), words embedded in white noise had a signal-to-noise ratio of -5 dB. The onset and offset of the noise coincided exactly with the onset and offset of the word. In the study and test tapes, a new word was spoken every 5 sec.

The test tape presented 240 words in a random order. One hundred twenty words were experimental tokens from the study phase, and 120 words were new fillers. Of the 120 studied words, 60 were repeated in the same voice and were the same tokens as those used as study items, and 60 were repeated in a different voice of the same gender. The test presented words embedded in noise at a signal-to-noise ratio of -5 dB.

Subjects and Design. The subjects were 45 undergraduate students at the University of Connecticut. The subjects either volunteered in exchange for partial course credit or were paid \$6.00. All subjects were native speakers of English, with normal hearing and normal or corrected vision at the time of testing.

A 2×2 mixed factorial design was used. The between-subjects variable was study condition (identification of words in noise or in the clear). Fifteen subjects were randomly assigned to each of the study conditions. The within-subjects variable was voice (same or different). Another group of 15 undergraduates completed the noise identification task under identical test conditions. These subjects did not receive any study; their data served as baseline controls.

Procedure. The subjects were tested in groups of four or fewer in a quiet room. The study and test stimuli were presented over headphones, and the subjects recorded their responses on answer sheets. They were told that they were providing normative data for a different experiment. During the study phase, the subjects heard 130 words spoken by six speakers, presented at a rate of 1 word every 5 sec. The first 5 and last 5 words served as primacy and recency buffers and were not included in the data analysis. In the identification tasks, the subjects listened to each word and wrote its spelling on their answer sheet. After the study phase, the subjects engaged in a filler task in which they completed a series of mathematical problems. Subjects worked on the mathematical problems for 4.5 to 5 min.

After the filler test, the subjects heard 240 words, half of which were studied words, with an equal number of words repeated in a voice that was the same as or different from the voice in the study phase. When a voice change occurred across study and test, it was always within gender. The subjects attempted to identify words degraded by noise under an implicit retrieval orientation. They were instructed to listen carefully to each word and to write down what word they heard or, if they were unsure, to write down the first word that came to mind. The experimental session lasted less than 45 min, after which the subjects were debriefed.

Results and Discussion

The analysis of the perceptual identification data was conducted on conditionalized scores. This analysis used a modified conditionalization procedure that differed from the standard conditionalization procedure. The standard conditionalization procedure omits test items that are misidentified at study (and, therefore, are not legitimate primes). Although conditionalizing test performance on study performance is justifiable, many more items are omitted after noise study than after clear study (45% of the noise identification trials and 16% of the clear study trials were misheard, and the corresponding test trials would not

Figure 1. Mean proportion of conditionalized target words and the corresponding baselines reported in the noise identification test for the clear encoding (top) and noise encoding (bottom) conditions of Experiment 1. The data are presented as a function of the speaker's voice.

enter into the test analysis; these values are well within the normal limits of intelligibility for monosyllabic isolated words; see Miller & Nicely, 1955).

The large number of omitted items in this analysis therefore increases the possibility of stimulus selection effects that may differentially affect the same- and different-voice conditions and may create a spurious voice effect. When the standard conditionalization procedure is employed, only tokens heard correctly at study enter into the analysis. Accordingly, items for which the study item is particularly unintelligible but for which the different-voice test item may have been more intelligible will be eliminated from the analysis. In contrast, items for which the study word was particularly easy to hear in noise were included in the

analysis, with the possibility that the corresponding different-voice test trial was less intelligible. If these hard-to-hear test tokens are then misidentified, a voice effect might arise artifactually, because conditionalization eliminates more hard-easy pairs and fewer easy-hard pairs. In contrast, stimulus selection effects should not be a problem for the same-voice pairs, which can only be easy-easy or hard-hard, with the latter often eliminated from the analysis.

One method to control for stimulus selection effects is to eliminate from the baseline data those trials that the test subjects misheard at study. Comparing test subjects' performance against the conditionalized baseline eliminates the possible contribution of stimulus selection effects and provides a conservative method for the assessment of a voice effect. To this end, the baseline scores were conditionalized on each subject's study performance. If a study item, such as the word *shop*, was misheard by a particular subject, that item was dropped from the mean baseline score that was used to assess that subject's performance. However, if the subject correctly identified the word *shop* at study, their baseline for that item included the proportion of baseline subjects who got that item correct. For example, if 13 of the 15 baseline subjects identified the word *shop* correctly, the baseline proportion for that item would be .867. Thus, there were separate same- and different-voice baselines created for each subject in each of the identification conditions. With the exception of decreases in overall priming and voice-specific priming, the pattern of results for both study groups was identical to what would be obtained using the standard conditionalization procedure.

Figure 1 presents the conditionalized perceptual identification data and the corresponding baseline values, displayed as a function of voice and study condition. To determine whether voice effects were present in each study condition, the subjects' identification performance was compared with the baseline levels. Because each subject now had a baseline tailored to his or her study performance, baseline performance was treated as a within-subjects variable (test or baseline), as was voice (same or different). A .05 level of significance was used to interpret all statistical comparisons, and all *t* tests were two-tailed.

For subjects who identified words in the clear at study, overall performance exceeded baseline [$t(14) = 8.34$, $RMS_e = .012$; RMS_e , root mean squared error], and there was evidence of priming in both voice conditions [$t(14) = 10.22$, $RMS_e = .014$ and $t(14) = 5.8$, $RMS_e = .022$ for same- and different-voice trials, respectively]. However, a *t* test on the same-voice/different-voice difference scores showed that the difference between same- and different-voice trials (3%) in the primed group was not significantly different from baseline (i.e., there was no voice effect).

Among subjects who identified words in noise at study, there was again evidence of overall priming [$t(14) = 9.08$, $RMS_e = .016$], of priming for same-voice trials [$t(14) = 13.04$, $RMS_e = .015$], and of priming for different-voice trials [$t(14) = 4.72$, $RMS_e = .02$]. More importantly, the difference between same- and different-voice trials (14%) among primed subjects over and above baseline performance was significant [$t(14) = 6.41$, $RMS_e = .016$]. This indicates that

there was additional priming as a result of voice preservation across repetitions after noise study.²

Finally, a *t* test comparing the same-voice/different-voice difference scores (over and above baseline) for the noise and clear study conditions confirmed that the study manipulation produced differences in the magnitude of the voice effect on implicit memory [$t(28) = 4.42$, $RMS_e = .022$].

The objective of Experiment 1 was to determine whether voice-specificity effects in implicit memory depend on the presence or absence of noise at test or, alternatively, on the similarity of study and test exemplars. The two competing views were evaluated by requiring subjects to study words either in the clear or in noise.

The most important finding from Experiment 1 is that performance on the perceptual identification of words in noise was significantly affected by study-to-test changes in voice. Although there was a numerical trend, in both study conditions, for same-voice items to be more accurately identified than different-voice items, preservation of voice across word repetitions significantly increased identification accuracy over and above the baseline only when the words were studied in noise and not after the words were studied in the clear.

The data from this experiment confirmed the same-voice advantage found by Goldinger (1992, 1996) when conditions resembled his experimental design (e.g., noise encoding–noise test) and confirmed the negative voice-specificity effects from Schacter and Church (1992) and from Jackson and Morton (1984) when conditions resembled their designs (e.g., clear encoding–noise test). Taken together, the data show that the presence of noise at study and at test enhances repetition priming above that found when the study item matches the test item on features such as modality, voice, and symbol form. The data further suggest that voice-specificity effects depend on the similarity of the study and test episodes, rather than on the memory system a task is assumed to tap.

In order to assess the generalizability of these results, a second experiment was conducted with a different identification test. The test presented words that had been low-pass filtered. Unlike the noise stimuli, which degraded the words by adding an additional signal, the low-pass filtered words were degraded by selectively removing aspects of the signal. Regardless of the signal processing manipulations, both transformations presumably encourage subjects to attend to perceptual aspects of the signal and foster data-driven processing.

EXPERIMENT 2

The design of Experiment 2 was identical to that of Experiment 1. During the study phase, the listeners identified words in noise or in the clear. Priming was assessed by asking the subjects to identify an equal number of studied and new words that were low-pass filtered. Half of the studied words were repeated in a voice that was the same as or different from that in the study phase.

Both the memory systems view and the processing view expect voice effects on the low-pass filter identification

test. The systems view expects voice effects on this task because it is an implicit task that does not use noise and should therefore receive output from the perceptual PRS, whereas the processing view's prediction is based on the classification of perceptual identification tasks as largely data-driven and on the importance of the processing match between study and test.

However, the theories differ in their predictions about the effects of the two study tasks (clear vs. noise encoding). From a systems perspective, the null effects of voice change in the noise identification test occurred because noise may have "interfered with processing those components of the acoustic waveform that provide access to voice information" (Schacter, 1992, p. 251). This statement implies that noise may prevent voice information from being processed at encoding and represented in the PRS, presumably because the noise interferes with right-hemisphere language processing. Therefore, the memory systems view would not expect to find voice effects on the filter identification test after noise encoding. However, voice effects should be present after encoding in the clear (Church & Schacter, 1994).

The processing view predicts the opposite pattern. If it is the match between study and test exemplars that is most critical to the presence of voice effects, voice effects on the filter identification test should be largest after noise encoding. This is because encoding words in noise requires extensive data-driven processing, which should transfer appropriately to a data-driven test task. Encoding clearly spoken words should result in more conceptual processing relative to noise encoding, which would then transfer less well to a data-driven test task.

Method

Materials. The stimulus materials were identical to those in Experiment 1, with the exception that the test words were low-pass filtered. The low-pass filter had a single transmission band extending from zero frequency up to a cutoff frequency of 2.5 kHz. The filter attenuated most of the high-frequency acoustic information above 2500 Hz, while allowing lower frequency sounds to pass through the filter unaltered. This manipulation produced words that sounded as if the speaker had produced them from behind a closed door.

As in Experiment 1, the study tapes presented 130 words (120 target words, plus five primacy and five recency buffer items) in noise (signal-to-noise ratio = -5 dB) or in the clear, at a rate of 1 word every 5 sec.

Subjects and Design. The subjects were 45 undergraduate students at the University of Connecticut. The subjects either volunteered in exchange for partial course credit or were paid \$6.00. All of the subjects were native speakers of English with normal hearing and normal or corrected vision at the time of test.

As in Experiment 1, a 2×2 mixed factorial design was used. The between-subjects variable was study condition (identification of words in noise or in the clear). Fifteen subjects were randomly assigned to each study condition. The within-subjects variable was voice (same or different). Another group of 15 unprimed subjects completed the filter identification test under identical test conditions, and their data served as baseline controls.

Procedure. The procedure was identical to the procedure in Experiment 1. The subjects identified 130 words in noise or in the clear at study. After the study phase, the subjects engaged in a filler task for 4.5 to 5 min. During the test phase, the subjects heard 240 low-pass filtered words, half of which were studied words, with an equal number of words repeated in a voice that was the same as or different from the voice in the study phase. The subjects were instructed to lis-

Figure 2. Mean proportion of conditionalized target words and the corresponding baselines reported in the low-pass filter identification test for the clear encoding (top) and noise encoding (bottom) conditions of Experiment 2. The data are presented as a function of the speaker's voice.

ten carefully to each word and to write down what word they heard, guessing if necessary. The experimental session lasted less than 45 min, after which the subjects were debriefed.

Results and Discussion

As in Experiment 1, the analysis of the perceptual identification data was conducted on conditionalized scores. Figure 2 presents the conditionalized identification data and conditionalized baseline values as a function of voice and study condition.

For subjects who identified words in the clear at study, overall performance exceeded baseline [$t(14) = 3.18$, $RMS_e = .016$]. There was priming in both voice conditions [$t(14) = 2.57$, $RMS_e = .02$ and $t(14) = 3.16$, $RMS_e = .051$ for same- and different-voice trials, respectively] but no

additional priming as a result of voice preservation across repetitions, relative to baseline.

A different pattern was found for subjects who identified words in noise at study. The overall identification performance of primed subjects was only slightly higher than that for baseline subjects (.78 vs. .74 proportion correct for primed and unprimed subjects, respectively). The amount of priming was not large enough to reach statistical significance. Priming was absent in the noise encoding condition because of differences in the way in which the noise and filtering transformations reduce word intelligibility. Specifically, the particular consonants and vowels that are most vulnerable to masking effects by noise may not be the same consonants and vowels that are most vulnerable to the effects of low-pass filtering (this will be true regardless of the speaker that produced the word). Therefore, even a highly primed item may be misidentified at test simply because it sounds different.

There was evidence of priming on same-voice trials [$t(14) = 2.95$, $RMS_e = .025$] but not on different-voice trials. In addition, the difference between same- and different-voice trials among primed subjects was significant [$t(14) = 3.45$, $RMS_e = .023$], indicating a voice effect on the low-pass filter test after noise identification at study. Moreover, the magnitude of the voice effect on identification performance (over and above baseline) was greater after noise encoding than after clear encoding [$t(28) = 2.12$, $RMS_e = .028$]. These findings parallel the pattern of voice effects found in the perceptual identification test that used noise—namely, that voice effects were only present when the study exemplars were perceptually degraded.

The most important finding from Experiment 2 is that significant study-to-test changes in voice were obtained on the identification test after the noise study. In this respect, the results from the low-pass filter identification test are identical to those from the noise identification test reported in Experiment 1. However, the lack of a voice effect after the clear study in Experiment 2 is at odds with the significant effect of voice change reported by Church and Schacter (1994). One major difference between Experiment 2 and their study is that, in the present experiment, the subjects identified words, whereas Church and Schacter required subjects to make semantic or nonsemantic judgments to each word.

To ascertain the effects of the different encoding tasks on the retrieval of voice information, two additional control experiments were conducted. The experiments used materials and design that were the same as those in Experiment 2 but replaced the noise and clear identification tasks with a semantic and nonsemantic encoding task from Church and Schacter (1994). In the first control experiment, 15 new subjects rated the clarity of the enunciation of each word on a four-point scale (nonsemantic processing condition). In the second experiment, 15 different subjects gave the number of meanings for each word (semantic processing condition), also on a four-point scale.

The results showed that overall performance levels on primed trials were equivalent across the enunciation and meaning groups and comparable to the performance lev-

els reported by Church and Schacter (1994). The findings also showed no effect of study-to-test changes in voice after either study task. This finding disagrees with the data reported in Church and Schacter. For example, they report that, in same-voice trials, identification was 8% more accurate than in different-voice trials in their filter task after enunciation study, whereas the analogous condition in the present experiment revealed only a numerical trend (3%) toward better identification on same-voice trials.

It is also possible that the null voice effect after clear study in Experiment 2, as compared with the experiment by Church and Schacter (1994) is due to a difference in the potency of the voice manipulation. In particular, in the present experiments (and the two control experiments), when a voice changed, it was always within gender, whereas in the experiment reported by Church and Schacter, voice change was confounded with gender. The fact that voice effects tend to be smaller when the perceptual similarity between study and test voices is high (see Goldinger, 1992, 1996) raises the possibility that the within-gender voice change was too subtle a manipulation to produce robust voice-specificity effects under the present experimental conditions.³

In summary, the fact that effects of study-to-test changes in voice on the filter identification task were obtained after noise study indicates that noise does not prevent voice information from being encoded. The present finding is buttressed by the data from the noise identification test reported in Experiment 1, which found voice effects on a test that used noise. Taken together, the data indicate that noise does not disrupt either the encoding or the retrieval of voice information under these task conditions. An alternative hypothesis is that the voice effects after the noise study condition arose because encoding a word in noise, as compared with words in the clear, placed greater demands on perceptual processing, which transferred appropriately to a test that also challenged perceptual processing by presenting words degraded by filtering.

CONCLUSIONS

The purpose of the experiments reported here was to resolve a conflicting report concerning the effects of study-to-test changes in voice on perceptual identification by determining which study conditions lead to the occurrence of voice-specificity effects with this task.

The major finding was that effects of study-to-test changes in voice were found on the perceptual identification of words in noise and on the perceptual identification of words that had been subjected to low-pass filtering, but only when the encoding task required extensive processing of perceptual information. In contrast, when subjects identified words in the clear, presumably the amount of data-driven processing decreased, and voice differences had no effect on memory performance.

The finding that subjects were more accurate at identifying a word embedded in noise, when both the voice and the noise were preserved across repeated presentations of a word, confirms the effects reported earlier by Goldinger (1992, 1996). This result suggests that stimulus-specific

details about the voice that produced a word are encoded and represented in memory and that voice information is available to influence performance on subsequent identification tests that use noise.

The memory systems view made exactly the wrong prediction about the effects of voice variation on the perceptual identification of words in noise. According to Schacter (1992), auditory priming is mediated by at least two systems—a phonological system that represents abstract information and a perceptual system that preserves highly specific representations of acoustic form (such as fundamental frequency information). If auditory word priming depends importantly on the contribution of a right-hemisphere PRS system, and if noise nullifies that contribution, as Schacter and Church (1992) have argued, voice effects should not have been found in this task.

The patterning of findings is instructive because it implicates the match between exemplars at study and test, rather than the memory system the task is assumed to tap, as the source of voice-specificity effects. Effects in auditory priming tasks appear to depend on the type of processing the subject initially performed during the study phase, so that, the more data-driven the processing, the greater the retrieval of perceptual information. Of the two encoding tasks, identifying words in noise would force the most extensive processing of stimulus information. In this task, subjects are confronted with a hard-to-hear word and must engage in substantially more data-driven processing than would ordinarily be required for words presented in the clear. The resulting memory trace presumably contains distinctive information about the perceptual form of a word and about the processing operations used to perceive and encode that word in memory (Kolers, 1979; Kolers & Roediger, 1984) that transfers well to another data-driven task. The results of the present study are consistent with this hypothesis.

Similar findings were obtained in the second implicit test, which required subjects to identify words that had been low-pass filtered. Among the two study conditions, only noise encoding produced robust voice effects on this test. A reason the noise study may have transferred well to the filter test is because, in both situations, the word is degraded, albeit differently. The noise manipulation makes a word difficult to hear by adding information to the signal that effectively masks some of the important information-bearing elements of speech, whereas the low-pass filter subtracts information above a certain frequency from the signal. Nevertheless, both transformations reduce the intelligibility of the items and both presumably challenge perceptual processing. The finding that identifying words in noise at encoding transfers to filter identification is consistent with a processing account.

In summary, the implicit memory performance pattern observed in the present experiments demonstrates that the extent to which a talker's voice influences auditory word identification depends on the similarity between the original episode and the test requirements. Spoken word identification tasks rely heavily on the perceptual processing involved in the bottom-up analysis of stimuli; consequently, the similarity of physical conditions between study and

test is the major determinant of data-driven retrieval. The findings also support the inference that the memory traces subserving spoken word perception include detailed talker-specific and context-specific information. As a whole, the data are most parsimoniously accounted for within a transfer appropriate processing framework.

REFERENCES

- CHURCH, B., & SCHACTER, D. (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. M., & KIRSNER, K. (1974). The effects of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology*, **26**, 274-284.
- GEISELMAN, R. E., & BJORK, R. A. (1980). Primary versus secondary rehearsal in imagined voices: Differential effects on recognition. *Cognitive Psychology*, **12**, 188-205.
- GOLDINGER, S. D. (1992). *Words and voices: Implicit and explicit memory for spoken words* (Tech. Rep. No. 7). Bloomington: Indiana University, Department of Psychology, Speech Research Laboratory.
- GOLDINGER, S. D. (1996). Words and voices: Implicit and explicit memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 1166-1183.
- HINTZMAN, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, **93**, 411-428.
- HOUSE, A. S., WILLIAMS, C. E., HECKER, M. H. L., & KRYTER, K. D. (1965). Articulation-testing methods: Consonantal differentiation with a closed-response set. *Journal of the Acoustical Society of America*, **37**, 158-166.
- JACKSON, A., & MORTON, J. (1984). Facilitation of auditory word recognition. *Memory & Cognition*, **12**, 568-574.
- JACOBY, L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning & Verbal Behavior*, **22**, 485-508.
- JACOBY, L., & BROOKS, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*, (Vol. 18, pp. 1-47). New York: Academic Press.
- KOLERS, P. A. (1979). A pattern-analyzing basis of recognition. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing and human memory* (pp. 363-384). Hillsdale, NJ: Erlbaum.
- KOLERS, P. A., & ROEDIGER, H. L., III (1984). Procedures of the mind. *Journal of Verbal Learning & Verbal Behavior*, **23**, 425-449.
- KUČERA, F., & FRANCIS, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- MARSOLEK, C. J., KOSSLYN, S. M., & SQUIRE, L. R. (1992). Form-specific visual priming in the right cerebral hemisphere. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 492-508.
- MILLER, G. A., & NICELY, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, **27**, 338-352.
- MORRIS, C. D., BRANSFORD, J. P., & FRANKS, J. J. (1977). Levels of processing versus transfer-appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, **16**, 519-533.
- MOSCOVITCH, M., GOSHEN-GOTTSTEIN, Y., & VRIEZEN, E. (1994). Memory without conscious recollection: A tutorial review from a neuropsychological perspective. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 619-660). Cambridge, MA: MIT Press.
- PALMERI, T. J., GOLDINGER, S. D., & PISONI, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 309-328.
- ROEDIGER, H. L., III (1990). Implicit Memory: Retention without remembering. *American Psychologist*, **45**, 1043-1056.
- ROEDIGER, H. L., III, & McDERMOTT, K. B. (1993). Implicit memory in normal human subjects. In H. Spinnler & F. Boller (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 63-131). Amsterdam: Elsevier.
- SCHACTER, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 501-518.
- SCHACTER, D. L. (1992). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. *Journal of Cognitive Neuroscience*, **4**, 244-256.
- SCHACTER, D. L., & CHURCH, B. A. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 915-930.
- SCHACTER, D. L., CHURCH, B. A., & TREADWELL, J. (1994). Implicit memory in amnesic patients: Evidence for spared auditory priming. *Psychological Science*, **5**, 20-25.
- SHEFFERT, S. M., & FOWLER, C. A. (1995). The effects of voice and visible speaker change on memory for spoken words. *Journal of Memory & Language*, **34**, 665-685.
- SQUIRE, L. R. (1987). *Memory and brain*. New York: Oxford University Press.
- TULVING, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-403). New York: Academic Press.
- TULVING, E., & SCHACTER, D. L. (1990). Priming and human memory systems. *Science*, **247**, 301-306.
- ZAIDEL, E. (1978). Lexical organization in the right hemisphere. In P. Buser & A. Rougeul-Buser (Eds.), *Cerebral correlates of conscious experience* (pp. 177-197). Amsterdam: Elsevier.
- ZATORRE, R., EVANS, A., MEYER, E., & GJEDDE, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, **256**, 846-849.

NOTES

1. Schacter and Church (1992) cite Zaidel (1978) as providing evidence that "auditory processing abilities of the right-hemisphere are especially impaired by noise" (p. 927), which, when coupled with the notion that the right-hemisphere is more proficient at representing format specific information (cf. Marsolek, Kosslyn, & Squire, 1992), is offered as an explanation of the null effects of voice change when words were presented in white noise. It is important to note that Zaidel himself does not claim that the presence of white noise disrupts the acoustic processing of the right-hemisphere. In his experiments, he used conversational (speech) noise and found that speech sounds interfere with right-hemisphere language processing more than with left-hemisphere language processing. Moreover, it is not obvious how white noise could selectively block the use of voice information but not phonetic information, as Schacter and Church suggest, given that noise, voice (pitch), and phonetic processing appear to occur in separate cortical regions (Zatorre, Evans, Meyer, & Gjedde, 1992).

2. To determine whether these results were simply due to trial-by-trial item effects, an items analysis was conducted with the same factors as those that were used in the subjects analysis. For this experimental condition, as well as for the analogous condition in Experiment 2, the pattern of results from the items analysis were identical to the results from the subject analysis.

3. A reviewer suggested replicating Church and Schacter (1994) to ensure the reliability of voice effects under their experimental conditions. To this end, I required subjects to make nonsemantic ($n = 17$) or semantic judgments ($n = 15$) about 24 multisyllabic low frequency words. The subjects then identified 48 low-pass filtered words, followed by a word recognition memory test. Test words were repeated in the same voice or in a different voice. I found that voice variation affected priming [$t(31) = 2.16$, $RMS_e = .002$; 6% voice effect] but not recognition. Conversely, depth of processing affected recognition [$t(31) = 5.21$, $RMS_e = .04$] but not priming. These findings replicate the dissociation reported by Church and Schacter.