Memory & Cognition
2000, 28 (8), 1406-1418

Effects of hearing words, imaging hearing words, and reading on auditory implicit and explicit memory tests

MAURA PILOTTI, DAVID A. GALLO, and HENRY L. ROEDIGER III
Washington University, St. Louis, Missouri

In four experiments, we examined the degree to which imaging written words as spoken by a familiar talker differs from direct perception (hearing words spoken by that talker) and reading words (without imagery) on implicit and explicit tests. Subjects first performed a surface encoding task on spoken, imagined as spoken, or visually presented words, and then were given either an implicit test (perceptual identification or stem completion) or an explicit test (recognition or cued recall) involving auditorily presented words. Auditory presentation at study produced larger priming effects than did imaging or reading. Imaging and reading yielded priming effects of similar magnitude, whereas imaging produced lower performance than reading on the explicit test of cued recall. Voice changes between study and test weakened priming on the implicit tests, but did not affect performance on the explicit tests. Imagined voice changes affected priming only in the implicit task of stem completion. These findings show that the sensitivity of a memory test to perceptual information, either directly perceived or imagined, is an important dimension for dissociating incidental (implicit) and intentional (explicit) retrieval processes.

The distinction between implicit and explicit memory has guided numerous theoretical and empirical studies designed to elucidate the processes that underlie different types of memory tests. In explicit tests, such as recognition and cued recall, subjects consciously attempt to recollect information encountered in an earlier phase of the experiment (the study or encoding phase). In implicit tests, subjects perform tasks that are ostensibly unrelated to the previously presented information. For instance, subjects identify degraded words or complete word stems with the first word that comes to mind. In these tasks, performance is facilitated for previously presented information relative to information encountered only at test, although subjects typically make no conscious attempt at remembering. This facilitation is termed *priming*.

The vast majority of implicit memory studies have used tests drawing on visual perceptual processes (see Roediger & McDermott, 1993). These studies have found decrements in priming when the physical attributes of the study stimuli do not match those of the test stimuli. For example, pictures produce less priming than printed words on the implicit test of word fragment completion, and words yield less priming than pictures on the picture fragment completion test (Weldon & Roediger, 1987). Similarly, spoken words produce less priming than printed words in the visual implicit tests of stem and fragment completion and word identification (see, e.g., Rajaram & Roediger, 1993). Printed words, however, produce less priming than spoken words in the auditory version of these tests (see Jackson & Morton, 1984; Pilotti, Bergman, Gallo, Sommers, & Roediger, 2000). Interestingly, decreases in priming are also observed when more fine grained changes within a modality are introduced between study and test, although these decrements are less robust across tests than those involving modality (visual vs. auditory) or format (picture vs. words) changes. For instance, less priming is observed in the auditory implicit test of perceptual identification when study words are spoken in a different voice at test (voice effects; see Pilotti et al., 2000), or are spoken in the same voice as the test, but with a different emotional intonation, such as happy versus angry (Church & Schacter, 1994). Of course, neither coarse nor fine-grained changes in the physical features of the study and test stimuli completely eliminate priming in these tests, thereby suggesting that implicit memory tests engage not only perceptual operations but also linguistic processes, whose abstract nature makes them less sensitive to such changes.

Interestingly, a recent study by Stuart and Jones (1996) has reported that the physical attributes of words play a role in priming even if these attributes are imagined rather than directly perceived. Stuart and Jones found that imaging the sound of printed words at study (auditory imagery) yielded as much priming as hearing them spoken in the implicit test of identification of words degraded by

This work was supported by Grant F32 DC00342 from the National Institute on Deafness and Other Communication Disorders. We thank the speakers who recorded the stimulus material, Ben Chi and Tim Beyer for their help in conducting the experiments, and Erik T. Bergman and Mitchell S. Sommers for helpful discussions. Correspondence should be addressed to H. L. Roediger III, Department of Psychology, One Brookings, Box 1125, Washington University, St. Louis, MO 63130 (e-mail: roediger@artsci.wustl.edu).

noise. However, reading printed words yielded virtually no priming in the same test. How can a covertly fabricated auditory stimulus produce as much priming on implicit tests as the actual stimulus? At first blush, the findings of Stuart and Jones suggest a twofold answer to this question, which, in our opinion, requires further examination. First, auditory imagery at study replicates with great accuracy the perceptual operations involved in actually hearing words, thereby suggesting that subjects can "create" their own data at study and that these data may serve as well as the actual auditory signals for priming in auditory implicit tests. Second, because there was virtually no priming in the read condition, it appeared as though there was no transfer of either perceptual or linguistic information from study to test when items were studied visually. This suggests that priming in the imagery condition was due solely to the transfer of generated (imaged) information from study to test.

We begin by addressing the issue of whether auditory imagery can replicate with great accuracy the perceptual operations involved in actually hearing words (see Finke, 1980). Particularly germane to this issue are two experiments conducted by Geiselman and Glenny (1977; see also Geiselman & Bjork, 1980; Nairne & Pusen, 1984). In these experiments, subjects were first familiarized with two talkers' voices (a male and a female) and then presented with a list of printed words. For each word, subjects were asked to imagine one of the familiar talkers saying the word and to repeat silently the word in that talker's voice for 10 sec. In a surprise recognition test with words spoken by the two familiar talkers (Experiment 1), recognition scores were higher for words spoken at test in the voice used to imagine them at study than for words spoken in a voice not matching the imagined talker. Surprisingly, when the voices of the test words were different from those the subjects imagined at study (Experiment 2), same-gender voices still produced higher recognition scores than did different-gender voices, although the overall recognition scores were significantly lower than those obtained in Experiment 1. In summary, if auditory imagery produces voice effects on recognition memory, and these effects are similar to those observed with spoken words (Craik & Kirsner, 1974; Palmeri, Goldinger, & Pisoni, 1993), it is reasonable to conclude that auditory imagery can engage the perceptual operations used in actually hearing words.

However, it is unlikely that auditory imagery obtained the same degree of perceptual accuracy in the Stuart and Jones (1996) study as in the other studies just reviewed. In Stuart and Jones, subjects were not familiarized with the test voice before the experimental session, and the auditory imagery instructions required them to simply imagine printed words as being spoken without reference to any voice. Therefore, because the auditory qualities (i.e., voice characteristics) of the images generated at study did not match the auditory qualities of the test words in the Stuart and Jones study, auditory imagery should have at least produced less priming than words spoken in the same voice at study and at test. Yet, Stuart and Jones found no difference in the priming effects produced by auditory imagery and direct auditory perception at study. To explain this outcome, they suggested that noise at test masks fine-grained auditory information such as the voice qualities of the talker, but it preserves coarse perceptual information such as modality of presentation. Therefore, the equivalent priming effects produced by the auditory imagery and the auditory perception study conditions would arise from their sharing the same coarse perceptual information.

Stuart and Jones's (1996) conclusion that noise at test was responsible for equating the priming effects produced by direct auditory perception and imagery raises two questions. First, was there any direct evidence in the Stuart and Jones study suggesting that their identification task was truly insensitive to fine-grained perceptual information? Pilotti et al. (2000) reported that the sensitivity of an identification task to fine-grained perceptual information depends on the level of noise at test, with high levels eliminating the voice effects that can be observed at lower levels. Stuart and Jones, however, did not report the signal-to-noise (S/N) ratio used in their experiment, nor did they include a condition with study-to-test voice changes, which would have permitted a direct examination of this claim. Second, if noise at test was responsible for equating the priming effects arising from imagery and direct perception at study, would auditory imagery produce less priming than hearing in implicit tasks that do not involve this specific form of stimulus degradation? Interestingly, although this question still awaits experimental testing in the auditory domain, an answer is available in the visual domain: Visual imagery generally yields smaller priming effects than direct visual perception on visual implicit tests. For example, Donnelly (1988) found that forming images of spoken words as they would appear if printed produced less priming than actually studying printed words on the implicit test of word fragment completion (see also Roediger & Blaxton, 1987). McDermott and Roediger (1994) found that seeing pictures and imaging the words that would label the pictures yielded less priming than reading the same words on word fragment completion. By the same token, they also found that reading words and creating visual images of their referents produced less priming than seeing pictures on a picture fragment completion test. Lastly, Jacoby and Witherspoon (1982) reported that hearing and then spelling words (assuming that subjects indeed imagined the written forms of the spoken words to spell them) produced less priming than actually reading the words in the visual implicit task of word identification.

In light of the interesting questions that the findings of Stuart and Jones (1996) raise, our study was specifically designed to examine the role of auditory imagery in different auditory memory tests. In our experiments, subjects were first familiarized with two voices (one male and one female). They were then randomly assigned to one of three study conditions in which they either heard words spoken by the two familiar speakers (hear condition), imagined printed words as spoken by these speakers (image condition), or silently read words printed in two different fonts (read condition). The last condition served as a control against which to measure the effects of hearing and imaging. In all the conditions, subjects focused on the physical attributes of the study words by rating the clarity of enunciation of spoken or imagined words (hear and image conditions) or by rating the ease with which printed words could be read. These tasks were chosen to assure the encoding of both coarse- and fine-grained perceptual information at study so that any null effect of study-to-test perceptual changes could not be attributed to insufficient perceptual processing at study. Each encoding task was followed by one of five memory tests in which all the words were spoken by one of the familiar talkers so that, in the hear and image conditions, half of the studied words were presented in the same voice as at study and half in a different voice. We examined memory performance in three implicit tests (identification of words masked by noise, identification of low-pass filtered words, and stem completion) and in two explicit tests (recognition and cued recall) to assess each memory task's sensitivity to coarse perceptual information (modality of study presentation: visual, auditory, or visual with auditory imagination) and fine-grained perceptual information (a speaker's voice: same or different between study and test). We selected implicit and explicit tests to see whether perceptual dimensions manipulated within the same experimental setting can dissociate priming from explicit memory performance (see Rajaram & Roediger, 1993; Roediger, Weldon, Stadler, & Riegler, 1992).

All the implicit tasks selected for this investigation have been consistently found to be sensitive to study-to-test modality shifts (see Jackson & Morton, 1984; Light, La Voie, Valencia-Laver, Albertson Owens, & Mead, 1992; Pilotti et al., 2000). Therefore, in all our auditory implicit tests, we expected that studying printed words (read condition) would produce less priming than the condition in which words were spoken at study. On the other hand, more fine grained perceptual changes, such as voice shifts, have been found to produce less consistent effects in auditory implicit tests (see Goldinger, 1996; Pilotti et al., 2000; Schacter & Church, 1992). Nonetheless, under specific experimental conditions these tests have yielded reliable voice effects across different experiments. For example, when noise does not excessively mask the test words, perceptual identification produces voice effects, even if the encoding task focuses subjects' attention on the semantic properties of the studied words (see Goldinger, 1996; Pilotti et al., 2000). Similarly, stem completion yields voice effects when the encoding task focuses subjects' attention on the perceptual properties of the study words, such as in the case of a clarity rating task (Church & Schacter, 1994; Schacter & Church, 1992). Therefore, under these conditions, we predicted that hearing study words spoken in one voice at study and then in a different voice at test would reduce priming in all our implicit tests relative to a condition without voice changes between study and test. Differences in priming between the image condition and the other two study conditions were expected to depend on the extent to which the processes engaged by auditory imagery replicate those involved in actually hearing spoken words at study. Similarly, differences in voice effects between the image condition and the hear condition would depend on the extent to which imaged words contain the same voice-specific information as spoken words.

The predictions regarding the effects of auditory imagery on explicit tests were based on earlier findings suggesting that typical explicit tests are more sensitive to encoding manipulations that affect conceptual processing at study than are their implicit counterparts (Blaxton, 1989; Church & Schacter, 1994; Craik, Moscovitch, & McDowd, 1994; Schacter & Church, 1992). Therefore, we predicted that the effects of auditory imagery on auditory explicit tests would depend primarily on the extent to which the auditory imagery instructions enhanced or weakened conceptual processing at study relative to the other study conditions. Specifically, auditory imagery from visual words requires that subjects translate orthographic codes into phonological codes, retrieve the voice qualities of the talker from long-term memory, and then adjust phonological codes to match these voice qualities. Therefore, it is reasonable to assume that imagery demands more processing than simply hearing or reading, and if this additional processing enhances conceptual analyses at study, imaging should facilitate recall and recognition. On the other hand, the opposite may be true if imagery focuses subjects' attention to more surface-like linguistic operations, taking away resources that would otherwise be allocated to operations that are beneficial to long-term retention, such as the activation of conceptual information. In the latter case, imaging should have a detrimental effect on recognition and recall relative to the other study conditions.

Lastly, we predicted that if recognition memory and cued recall engage perceptual processes, these tests should produce modality and voice effects, as do their implicit counterparts. Unfortunately, it is still a matter of debate whether the explicit tasks selected for this investigation are as sensitive as implicit tests to perceptual changes between study and test. Notably, voice effects in recognition memory have been reported by Craik and Kirsner (1974), Goldinger (1996), Palmeri et al. (1993), and Sheffert (1998a), but not by Schacter and Church (1992). Similarly, modality effects, but not voice effects, in cued recall have been reported by Pilotti et al. (2000). In the present study, we examined the sensitivity of explicit memory tests to perceptual information following encoding tasks that enhanced attention to perceptual information, thereby providing the ideal testing ground for assessing whether the retrieval operations engaged by these tests rely on coarse- and fine-grained perceptual information.

GENERAL METHOD

Stimuli and Apparatus

The stimuli for all the experiments were selected from a set of 300 bisyllabic and trisyllabic English words, each with a unique initial stem (see Pilotti et al., 2000). The words were either low or medium frequency ($M = 9.4$, $SD = 14.8$; Kučera & Francis, 1967), and were all highly familiar ($M = 6.7$, $SD = .57$; Nusbaum, Pisoni, & Davis, 1984). Two talkers, one male and one female, recorded all the stimuli. Two other talkers, one male and one female, also recorded the 20 stimuli for the generalization phase (explained below).

The stimuli of the study phase were selected from this set so as to comprise four study blocks of 36 words matched for frequency and familiarity (frequency, $M = 8.7$, $SD = 13.7$; familiarity, $M = 6.7$, $SD = .6$). The 36 stimuli of the voice familiarization phase and the 20 stimuli of the generalization phase were selected from the remaining stimuli. Five versions of each study block were devised: words printed in a regular font (Avian) for the image condition, words spoken in either a male or a female voice for the hear condition, and words printed in two irregular fonts for the read condition (Florentine and Formal fonts). The test block included 10 filler words and 144 randomly presented words (72 studied and 72 nonstudied). A Latin square design was used to counterbalance the pairing of study blocks, the assignment of talkers (male and female) or fonts (Florentine and Formal) to study blocks, and the assignment of talkers (male and female) to the test block. This process created 16 unique combinations of study and test blocks.

The auditory stimuli were recorded and subsequently digitized at a sampling rate of 20-kHz, and their amplitude levels were digitally equated using a software package specifically designed to modify speech waveforms. The stimuli were presented over headphones via an IBM-compatible computer equipped with a 16-bit digital-to-analog converter with anti-aliasing filters.

Procedure

Each experiment consisted of five phases: a voice familiarization session, a generalization test, a study phase, a 5-min break, and a test phase. The voice familiarization phase involved two blocks of 36 trials, one spoken by a female talker (Mary) and the other spoken by a male talker (Paul). The same 36 words were used in each block so that subjects heard each word spoken by both the female and the male talkers. On each trial, the name of one of the talkers appeared in the center of the computer screen 1 sec before the onset of a spoken word and remained on the screen for 2.5 sec. A 1-sec blank screen separated consecutive trials. Subjects were instructed to become familiar with each speaker's voice so as to prepare themselves for the next phase of the experiment, which involved discriminating between familiar and novel voices. During the generalization test, subjects heard 20 words randomly presented, half produced by the two familiar talkers (Mary and Paul) and half produced by two novel talkers (one male and one female). On any given trial, their task was to indicate whether the voice was familiar or unfamiliar by pressing one of two keys on the computer keyboard, appropriately labeled as "O" (old voice) and "N" (new voice). They received feedback on each trial. In this task, subjects reached an average accuracy of 95% (range = 90–100) before entering the next phase of the experiment.

During study, subjects were presented with two blocks of 36 trials. Each trial consisted of a 500-msec blank screen, a 150-msec tone serving as a warning signal, a visually presented number labeling the trial, a 500-msec blank screen, and the auditory or visual presentation of a stimulus word. A trial ended and the next trial began 8 sec after stimulus presentation. There were three between-subjects study conditions, all involving incidental encoding: hear, image, and read. In the hear condition, the words of one block were spoken by the familiar male talker (Paul), and those of the other block were spoken by the familiar female talker (Mary). On any given trial, subjects were asked to rate how clearly the talker pronounced the word of that trial on a 7-point scale (7 = *very clear*; 1 = *very unclear*). In the image condition, at the beginning of each block, subjects were reminded of either Mary's or Paul's voice by being given a three-word example of that voice. As in the familiarization phase, the name of one of the talkers would first appear on the screen followed by a word spoken by that talker. Subsequently, on any given trial, subjects read a word presented in the center of the computer screen. Depending on the block, they were asked to imagine either Paul or Mary saying the word. Subjects were to rate the clarity of their image of the spoken word on the 7-point scale described above. In the read condition, each block contained words printed in either Florentine or Formal font. On each trial, subjects were asked to rate how easy it was to read a word printed in one of the two fonts on a 7-point scale (7 = *very easy*; 1 = *very difficult*). All the subjects were given a two-page booklet to record their answers.

The test block included 10 filler trials, placed at the beginning of the block, and 144 randomly presented trials. Stimulus presentation was self-paced. Subjects pressed the Enter key to initiate each trial, which consisted of a 500-msec blank screen, a 150-msec tone serving as a warning signal, a visually presented number labeling the trial, a 500-msec blank screen, and the presentation of a speech stimulus. Depending on the test, the speech stimulus was a word degraded by noise or via a low-pass filter (perceptual identification), a word stem (stem completion and cued recall), or a word spoken in the clear (i.e., without degradation; recognition test). Half the subjects heard stimuli spoken by the familiar male talker, and the other half heard stimuli spoken by the familiar female talker. In the implicit tests, subjects were given the task of either identifying degraded words (identification of words masked by noise or low-pass filtered) or completing stems with the first word that came to mind (auditory stem completion). In the explicit tests, subjects were asked to perform a surprise recognition or cued recall task. Additional information regarding these tests is presented in the Method section of each experiment.

EXPERIMENT 1
Identification of Words Masked by Noise

In Experiment 1, we attempted to replicate Stuart and Jones's (1996) finding of equal priming for hearing and imaging in the implicit test of perceptual identification of words masked by noise. In this experiment, after familiarizing subjects with the voices of two talkers, we presented them with one of three study conditions: hear, image, or read. In all the conditions, we asked subjects to focus on the perceptual properties of the study words by rating either the clarity of enunciation of spoken or imaged-to-be-spoken words, or the ease of reading words printed in two irregular fonts. We thought that the instructions to imagine words spoken by a specific talker would enhance the similarity between the image condition and the hear condition, in which words spoken by that talker were actually heard. Furthermore, we thought that a rating task involving surface processing would encourage the encoding of fine-grained auditory information (either directly perceived or imagined). Thus, this design provided a strong test for the effects of auditory imagery on the implicit task of perceptual identification.

At test, studied and nonstudied words, all spoken by one of the familiar talkers, were masked by noise. As in the Stuart and Jones (1996) experiment, subjects were asked to identify these words without reference to the previous study phase. We used two test conditions involving different levels of masking noise: high (−10, e.g., the noise was 10 dB louder than the signal) and low (−5, e.g., the noise was 5 dB louder than the signal). We hypothesized that if noise masks voice information so that it is unusable at test, as argued by Schacter and Church (1992), the task's sensitivity to voice information should vary as a function of the S/N ratio, with the −5 S/N ratio more likely to produce voice effects than the −10 ratio (see Pilotti et al., 2000; Sheffert, 1998b). In contrast to the Stuart and Jones experimental design, in which only same-voice repetitions were available at test, half our studied words matched the voice at study and half presented a voice mismatch. The voice manipulation permitted a direct evaluation of whether subjects encoded fine-grained perceptual information at study rather than only coarse information about modality of presentation.

Method

Subjects. One hundred and twelve undergraduates from Washington University participated in the experiment in partial fulfillment of course requirements. Sixty-four subjects were assigned to the test phase involving a S/N ratio of −10 (hear = 32, image = 16, and read = 16), and 48 were assigned to the test phase involving a S/N ratio of −5 (16 subjects in each study condition).[1]

Design and Procedure. A mixed factorial design was used. Study condition (hear, image, or read) and test type (S/N ratio of −5 or −10) were the between-subjects factors. Item type (studied or nonstudied words) and test voice in the hear and image conditions (same voice as at study or different voice) were the within-subjects factors.

The experiment was presented to the subjects as an investigation of auditory and visual perception consisting of a series of brief tasks. Prior to the memory tests, the experimental procedure was as described in the earlier General Method section. At test, subjects were instructed to identify degraded words. Stimulus degradation was accomplished by presenting words in continuous white noise at a S/N ratio of either −5 or −10 (the noise was 5 or 10 dB louder than the signal). Subjects were to write their answers on a booklet containing 154 numbered blanks.

Results and Discussion

Table 1 displays the proportions of studied and nonstudied words correctly identified at each S/N ratio as a function of study condition and test voice. The second and third columns of Table 1 display the proportion of studied words correctly identified as a function of test voice (same as at study or different), and the fourth column displays their average. Note that because the voice manipulation is not relevant to the read condition, only the proportion correct irrespective of test voice is reported for this condition. The fifth column reports the proportion of nonstudied words correctly identified (baseline). The last column displays the priming effects in each study condition.

There are several points to note from Table 1: First, under difficult listening conditions (S/N ratio of −10), hearing and imaging produced small and virtually equivalent priming effects, whereas reading did not produce any facilitation (+.07, +.05, and −.01, respectively). Second, under less difficult listening conditions (S/N ratio of −5), hearing produced larger priming effects than either imaging or reading (+.15, +.04, and +.06, respectively). Third, hearing words involved the encoding of talker-specific information, as indicated by the voice effects in the −5 S/N ratio condition (same voice priming = +.17 and different voice priming = +.13). Fourth, under the difficult listening conditions produced by the −10 S/N ratio, no talker-specific effects on priming occurred following hearing (same voice priming = + .07 and different voice priming = +.07). Fifth, imaging spoken words did not produce talker-specific effects on priming in either S/N condition.

These observations were supported by a 2 (studied or nonstudied) × 3 (hear, image, or read) mixed factorial analysis of variance (ANOVA) on the proportions of studied and nonstudied words correctly identified at each S/N ratio (all results reported here and in the following experiments are reliable at the .05 level unless otherwise specified). The data at each S/N ratio were analyzed separately because the different baselines generally obtained under these listening conditions make cross-test statistical analyses difficult (see Rajaram & Roediger, 1993). The −10 S/N ratio condition yielded a main effect of item type [$F(1,61) = 12.73$, $MS_e = 0.004$] and a significant interaction between item type and study condition [$F(2,61) = 5.53$, $MS_e = 0.004$; study condition, $F = 1.4$]. The reliable interaction suggested that priming var-

Table 1
Experiment 1, Proportion of Correctly Identified Words as a Function of Study Condition (Hear, Image, or Read), Item Type (Studied or Nonstudied), Test Voice (Same or Different), and S/N Ratio (−10 or −5)

	Voice Studied				
Study Condition	Same	Different	Studied	Nonstudied	Priming (S − NS)
			S/N Ratio −10		
Hear	.28	.28	.28	.21	+.07
Image	.22	.20	.21	.16	+.05
Read			.20	.21	−.01
			S/N Ratio −5		
Hear	.61	.57	.59	.44	+.15
Image	.48	.48	.48	.44	+.04
Read			.50	.44	+.06

Note—S/N, signal-to-noise; S–NS, studied minus nonstudied.

ied as a function of study condition, and the analysis of the proportions of correctly identified nonstudied words indicated that this interaction was not due to differences in baseline performance across the three conditions (F = 1.1). The analysis of the priming scores (studied minus nonstudied) indicated that hearing and imaging yielded virtually equivalent priming effects [$t(46) < 1$], whereas reading produced no priming [$t(15) < 1$] (all the t tests were one-tailed).

The −5 S/N ratio condition yielded a main effect of item type [$F(1,45) = 43.83$, $MS_e = 0.004$] and a significant interaction between item type and study condition [$F(2,45) = 6.90$, $MS_e = 0.004$; study condition, $F < 1$]. Again, the reliable interaction indicated that priming varied as a function of study condition with the identification of nonstudied words not reliably changing between study conditions ($F < 1$). Hearing a word at study produced more priming than imaging or reading it [$t(30) = 3.69$; $t(30) = 2.71$, respectively]. Imaging and reading, however, produced roughly equivalent priming effects [$t(30) < 1$].

Separate analyses of the priming scores of the image and hear conditions were conducted at each S/N ratio to determine whether the priming effects for words heard at test in the same voice as at study were greater than the priming effects for words heard in a different voice. At the S/N ratio of −10, priming did not vary as a function of same and different voice (hear—same voice = +. 07 and different voice = + .07, $t < 1$; image—same voice = +.06 and different voice = +.04, $t < 1$). At the S/N ratio of −5, however, hearing a word in the same voice produced more priming than hearing it in a different voice [+.17 and +.13, respectively; $t(15) = 2.25$], whereas imaging in either same or different voice did not affect performance at test (+.04 and +.04, respectively; $t < 1$).

Taken together, these findings suggest that the encoding of talker-specific information occurred when subjects were perceiving spoken words at study, but the transfer of this information at test was dependent on the level of noise. Interestingly, the findings of the perceptual identification test at −10 S/N ratio closely resemble those reported by Stuart and Jones (1996). That is, we found no reliable difference in priming between hearing and imaging at study and no priming effects following reading. It is reasonable to assume that these findings were the result of a high level of masking noise, which not only made the implicit task insensitive to fine-grained perceptual information, but also lowered the priming effects in the hear and read conditions. Indeed, when noise at test was less intense (S/N ratio of −5), hearing words at study produced relatively more priming than imaging, whereas there was no reliable difference in the priming effects produced by imaging and reading.

Although the level of masking noise at test modulated the task's sensitivity to auditory-specific information and to cross-modality information (priming effects following reading were observed only at −5 S/N ratio), the magnitude of the priming effects following imaging remained unchanged by the noise level at test. These priming effects appeared to be unaffected not only by the level of noise at test but also by study-to-test voice changes. The absence of voice effects in this condition could have been due to the small size of the overall priming scores, which made voice effects more difficult to detect. It is also possible that the stimulus degradation produced by noise in this experiment enhanced the mismatch between the talker-specific information embedded in the test stimuli and that generated at study, thereby making the latter largely unusable at test. Consistent with this notion, the voice effects following auditory imagery instructions observed by Geiselman and his colleagues with a recognition task (Geiselman & Bjork, 1980; Geiselman & Glenny, 1977) were obtained with words spoken in the clear (i.e., without any degradation). Of course, if noise disrupts voice effects in the image study condition, but not in the hear condition at the −5 S/N ratio, generated and directly perceived voice characteristics may be assumed to be qualitatively different. As noted, however, the small size of the priming effects in the image condition makes any conclusion regarding this null finding difficult to interpret. Because the finding that imaging did not yield voice effects in this experiment has important theoretical implications regarding the extent to which auditory imagery replicates speech perception processes, in the next experiment, we reexamined this issue with a different form of stimulus degradation.

EXPERIMENT 2
Identification of Low-Pass Filtered Words and Explicit Recognition

As discussed earlier, the finding that auditory imagery did not yield voice effects in Experiment 1 could be due to the small size of the priming effects observed in the identification task. Alternatively, it is possible that generated voice characteristics did not match those directly perceived at test and the masking noise simply increased this perceptual mismatch, thereby eliminating any benefit that would have arisen from reinstating the talker's voice at test. In Experiment 2, we examined these alternative explanations by degrading the test words of the identification task of Experiment 1 with a low-pass filter (Church & Schacter, 1994). In contrast to noise, which adds irrelevant information to the speech signal, this form of stimulus degradation attenuates specific portions of the signal (i.e., the higher frequencies) while leaving the remaining portions intact. Because the task of identification of low-pass filtered words has been consistently found to be sensitive to voice information in earlier experiments (Church & Schacter, 1994; Pilotti et al., 2000), we thought that this task would provide a strong test for voice effects in the auditory imagery condition.

In Experiment 2, we also addressed the issue of whether the effects of auditory imagery on memory vary with the implicit or explicit nature of the test. To this end, an auditory recognition test with words presented in the clear (without degradation) served as the explicit counterpart of the low-pass filter identification task. In the recognition task of this experiment, subjects were presented with whole word tokens, and they were required to indicate whether or not the test words had been previously studied. In contrast to the identification task, recognition memory is particularly sensitive to encoding manipulations affecting conceptual processing at study (see Schacter & Church, 1992). The sensitivity of recognition memory to perceptual information, however, is still a matter of debate, with some studies finding voice (Goldinger, 1996; Sheffert, 1998a) and modality (Hintzman, Block, & Inskeep, 1972; Kirsner, 1974) effects and others reporting no effects (Church & Schacter, 1994; Craik et al., 1994; Schacter & Church, 1992). In our experiment, prior to the memory test, the voice familiarization phase and the encoding task focused subjects' attention on perceptual details besides abstract linguistic information, thereby ensuring the encoding of both coarse- and fine-grained perceptual details. Therefore, our experimental design provided a propitious testing condition for assessing the extent to which recognition memory relies on perceptual details besides linguistic information. Given this encoding context, finding that recognition memory is not as sensitive to perceptual information as its implicit counterpart would provide further support for the notion that the former is driven by abstract linguistic operations—specifically, those involving conceptual information. Furthermore, finding that the auditory imagery condition either reduces or enhances recognition memory relative to the other study conditions would provide evidence regarding the extent to which imaging spoken words affects conceptual processing at study.

In contrast to the contradictory findings regarding recognition memory and voice effects discussed above, Geiselman (Geiselman & Bjork, 1980; Geiselman & Glenny, 1977) found that recognition memory following auditory imagery was impaired by study-to-test voice changes. This finding supports the notion that recognition memory is sensitive to perceptual information (even if perceptual information is imagined rather than directly perceived at study). In their experiments, however, subjects were asked to form an auditory image of a word spoken by a given talker and then repeat to themselves that word for 10 sec, thereby confounding the effects of imagery and rehearsal on recognition memory. In Experiment 2, we examined the effect of auditory imagery without rehearsal on recognition performance. The question of interest was whether imagery enhances or disrupts recognition memory relative to the other study conditions (read or hear).

Method

Subjects. Ninety-six undergraduates from the same subject pool as that used in Experiment 1 participated in the experiment in partial fulfillment of course requirements (48 were assigned to the identification task and 48 to the recognition task). Two additional subjects who performed the recognition task were excluded from the analyses for displaying a disproportionately high false alarm rate (.58 and .46, respectively).

Design and Procedure. A mixed factorial design was used for this experiment. Study condition (hear, image, or read) and test type (identification of low-pass filtered words or recognition) were the between-subjects factors. Item type (studied or nonstudied words) and test voice in the hear and image conditions (same or different) were the within-subjects factors.

The experiment was presented as an investigation of auditory and visual perception consisting of a series of brief tasks. Prior to the memory test, the procedure was as described earlier. In the implicit test, subjects were given a booklet containing 154 numbered blanks and were asked to identify degraded words. Stimulus degradation was accomplished by low-pass filtering words at 1 kHz. In the explicit test, subjects were asked to recognize the words presented during study among others that had not been presented before. They were to write their answers ("yes" for words they recognized and "no" for words they did not) in a booklet containing 154 numbered blanks. Words were presented in the clear.

Results and Discussion

Table 2 presents the proportions of studied and nonstudied words identified in the low-pass filter task along with the proportions of words correctly or incorrectly recognized as old (hits and false alarms) in the recognition task. This table also displays the priming scores and the

Table 2
Experiment 2, Proportion of Words Identified or Recognized as a Function of Study Condition (Hear, Image, or Read), Item Type (Studied or Nonstudied), and Test Voice (Same or Different)

	Identification Task				
	Voice Studied				
Study Condition	Same	Different	Studied (M)	Nonstudied	Priming (S − NS)*
Hear	.47	.40	.44	.29	+.15
Image	.38	.36	.37	.32	+.05
Read			.39	.31	+.08

	Recognition Task				
	Voice Studied				
Study Condition	Same	Different	Hit Rate (Average)	False Alarm Rate	Corrected Scores (H − FA)†
Hear	.79 (.61)	.77 (.59)	.78	.18	.60
Image	.61 (.46)	.64 (.48)	.63	.15	.48
Read			.77	.15	.62

*Priming scores (studied − nonstudied) in the low-pass filter task are shown in this column. †The proportion of corrected recognition scores (hits minus false alarms) is shown in this column.

corrected recognition scores (hits minus false alarms). There are three points to note from Table 2: First, hearing produced more priming than imaging and reading, and, as in the −5 S/N ratio condition of Experiment 1, imaging did not increase priming over reading (+.15, +.05, and +.08, respectively). Second, voice effects were found in the hear condition but not in the image condition (hear: same voice priming = +.18 and different voice priming = +.11; image: same voice priming = +.06 and different voice priming = +.04). Third, although hearing and reading produced virtually equivalent recognition scores (corrected recognition scores: hits minus false alarms), there was a trend in the direction of an impairment in recognition memory resulting from imaging words at study.

Identification of low-pass filtered words. The data of the identification test were submitted to a 2 (studied or nonstudied) × 3 (hear, image, or read) mixed factorial ANOVA. These analyses revealed a main effect of item type [$F(1,45) = 68.07$, $MS_e = 0.003$] and a reliable interaction between item type and study condition [$F(2,45) = 6.61$, $MS_e = 0.003$] (study condition: $F < 1$). The reliable interaction indicated that priming varied as a function of study condition, and the analysis of the identification scores for nonstudied words indicated that this interaction was not due to different base rates across the three conditions ($F < 1$). As in the −5 S/N condition of Experiment 1, hearing a word at study produced more priming than imaging or reading it [$t(30) = 4.10$; $t(30) = 2.33$, respectively]. The latter study conditions did not yield reliably different priming effects [$t(30) = 1.02$].

Separate analyses of the priming scores of the hear and image conditions were conducted to determine whether there was greater priming in the same voice than in the different voice condition. Consistent with the findings observed at the −5 S/N condition of Experiment 1, hearing a word in the same voice as at study produced more priming than hearing the word in a different voice [$t(15) = 2.48$], whereas imaging a word at study did not produce voice effects [$t(15) < 1$]. As in Experiment 1, the small size of the overall priming observed in this condition, however, might have been responsible for this null result.

Recognition. A one-way ANOVA was performed on the corrected recognition scores (hits minus false alarms) with study condition (hear, image, or read) as factor. Although this analysis did not show a reliable effect of study condition [$F(2,45) = 2.08$], the corrected recognition scores showed a trend in the direction of a decrement in performance in the image condition relative to the other conditions (image = .48; hear = .60; read = .62). There were no reliable voice effects ($ts < 1$); corrected recognition scores (H-FA) are shown in parentheses in the Same and Different Voice columns in Table 2.

Taken together, these findings suggest that the encoding of talker-specific information occurs when subjects are perceiving spoken words at study, but the transfer of this information at test is dependent upon the type of memory task (explicit vs. implicit). Specifically, in the hear condition, voice matches between study and test facilitated performance on the implicit, but not the explicit, task. These results are consistent with those reported by Church and Schacter (1994; Schacter & Church, 1992), who also found that voice changes between study and test were more likely to affect implicit than explicit memory performance in the same tasks.

The finding that reading and hearing produced equivalent recognition memory (for similar findings in visual recognition, see Craik et al., 1994) supports the notion that intentional retrieval operations in this task rely primarily on abstract linguistic information (perhaps conceptual). The differences in performance between the image condition and the other study conditions suggest that auditory imagery might have weakened conceptual processing at study. These differences, however, did not reach statistical significance.

EXPERIMENT 3
Stem Completion and Cued Recall

In Experiment 3, we further examined the effects of imagery in explicit memory with a cued recall task. We reasoned that the intact test stimuli of the recognition task of Experiment 2 represented strong retrieval cues, which might have made any effect of imagery on explicit memory more difficult to detect. If this is true, cued recall would provide a clearer pattern of results than the recognition task of Experiment 2 because subjects would be presented with data-limited cues (test stimuli in which only the initial portions of whole words serve as retrieval cues). Of course, like the recognition task of the earlier experiment, cued recall is especially sensitive to encoding manipulations affecting conceptual processing at study (Schacter & Church, 1992), making this task an ideal testing ground for clarifying the effects of auditory imagery on explicit memory.

In Experiment 3, we also reexamined the role of auditory imagery in a different implicit test, stem completion. In stem completion, subjects are presented with the initial portion of a spoken word (stem), and their task is to generate the first word that comes to mind. Although stem completion is known to be sensitive to coarse perceptual information (Pilotti et al., 2000), recent findings have demonstrated that its sensitivity to voice information depends on the encoding task. For example, stem completion has yielded voice effects across different experiments following a clarity rating encoding task (Church & Schacter, 1994; Schacter & Church, 1992), whereas it has produced mixed findings following a semantically oriented encoding task (Pilotti et al., 2000; Schacter & Church, 1992). Because Schacter and Church reported voice effects in this test following a clarity rating task, we expected that stem completion would also be sensitive to voice information in this experiment.

Method

Subjects. One hundred and sixty undergraduates from the same subject pool participated in the experiment in partial fulfillment of course requirements. Eighty were assigned to the stem completion task (hear = 32, image = 32, and read = 16) and 80 to the cued recall task (hear = 32, image = 32, and read = 16).

Design and Procedure. A mixed factorial design was used. Study condition (hear, image, or read) and test type (stem completion or cued recall) were the between-subjects factors. Item type (studied or nonstudied word) and test voice in the hear and image conditions (same or different) were the within-subjects factors.

The experimental phases were as described earlier with the exception of the memory tests. The test stimuli were word stems obtained by isolating the initial portion of each word that made that portion distinguishable from the others (basically, the first syllable). In the implicit test, subjects were told that word stems would be spoken over the headphones and that their task was to respond to each stem with the first word that came to mind. They were to report their answers in a booklet with 154 numbered blanks. In the explicit test, subjects were given the same booklet to fill out, but with different instructions. Subjects were asked to remember the words presented during the earlier study phase. If a stem was the beginning of a word encountered at study, their task was to write the word in the booklet. Alternatively, if they thought that a stem was not the beginning of a studied word, their task was to write an X in the booklet.

Results and Discussion

The proportions of words completed and recalled in the various study conditions are shown in Table 3. In this table, priming and corrected recall scores (studied minus intrusions) are reported in the last column. The data were analyzed in the same manner as those of Experiment 2. There are four points to note from Table 3: First, hearing produced more priming than imaging and reading (+.23, +.12, +.11, respectively). Second, same-voice percepts produced a numerically larger priming effect than did different-voice percepts in both the hear and image conditions. Third, hearing and reading produced better recall than imaging, indicating that the auditory imagery task impaired explicit memory performance. Fourth, cued recall performance was not affected by study-to-test voice changes.

Stem completion. The effect of item type was significant [$F(1,77) = 212.51$, $MS_e = 0.004$], as well as the effect of study condition [$F(2,77) = 7.37$, $MS_e = 0.017$]. The interaction between item type and study condition was also significant [$F(2,77) = 15.47$, $MS_e = 0.004$]. The reliable interaction indicated that priming varied as a function of study condition, and as in the previous experiments, this interaction was not due to differences in baseline performance across the three study conditions ($F = 1.26$). As in Experiment 2, hearing a word at study produced considerably more priming than imaging or reading it [$t(62) = 4.69$; $t(46) = 4.56$, respectively], whereas there was no difference between these latter two conditions [$t(46) < 1$]. Interestingly, changing voice between study and test reliably decreased the priming effects obtained in both the hear and the image conditions [hear, $t(31) = 1.90$; image, $t(31) = 1.91$].

Cued recall. The analysis of the corrected recall scores yielded a significant effect of study condition [$F(2,77) = 12.43$, $MS_e = 0.020$]. Although hearing and reading produced equivalent recall [$t(46) = 1.41$], they induced much better recall than did imaging [$t(62) = 4.69$; $t(46) = 3.66$]. No reliable evidence of voice effects was observed in the hear and image conditions (ts < 1.0).

Taken together, the findings of this experiment indicate that sensitivity to voice information and auditory imagery can both dissociate explicit and implicit auditory memory tests. Whereas stem completion yielded voice effects both in the hear and image conditions, cued recall did not yield voice effects in either condition. Auditory imagery at study had a detrimental effect on cued recall performance relative to reading, but did not reduce priming below reading in stem completion. This replicates the pattern of results observed in the recognition test of Experiment 2 and further corroborates the claim that our imagery task weakened conceptual processing at study, affecting performance on tests that rely heavily (if not exclusively) on abstract linguistic information.

Interestingly, as in the identification tasks of the earlier experiments, imaging and reading produced virtually equivalent priming effects (albeit smaller than hearing) in

Table 3
Experiment 3, Proportion of Words Completed or Recalled as a Function of Study Condition (Hear, Image, or Read), Item Type (Studied or Nonstudied), and Test Voice (Same or Different)

	Stem Completion Task				
	Voice Studied				
Study Condition	Same	Different	Studied (M)	Nonstudied	Priming (S − NS)*
Hear	.54	.51	.53	.30	+.23
Image	.43	.39	.41	.29	+.12
Read			.37	.26	+.11
	Cued Recall Task				
	Voice Studied				
Study Condition	Same	Different	Studied (M)	Intrusions	Corrected Scores (S − I)†
Hear	.54 (.38)	.52 (.37)	.53	.16	.37
Image	.37 (.20)	.37 (.20)	.37	.17	.20
Read			.45	.15	.30

*Priming scores (studied − nonstudied) in the completion task are shown in this column. † The proportion of corrected recall scores (studied − intrusions) is displayed in this column.

the stem completion task. A reasonable explanation for this finding is that the priming effects arising from imaging and reading at study were due solely to abstract linguistic information. However, the finding that auditory imagery produced voice effects in stem completion, which were similar in magnitude to those observed in the hear condition, suggests that generated auditory information did influence priming in this task. This pattern of results is puzzling. If generated auditory information adds to the linguistic information encoded at study, one would expect the overall priming effects in the image condition to be greater than those in the read condition. One tentative interpretation is that imaging words spoken by a familiar talker did not produce more priming than reading because it weakened conceptual processing at study. Specifically, the possibility exists that auditory imagery focused subjects' attention on surface-like linguistic information (orthographic and phonological word codes) at the expense of conceptual processing. Indeed, perceptual changes reduced but did not eliminate priming in our implicit tasks, suggesting that the priming effects observed in these tasks may not be entirely insensitive to shifts in the processing of different forms of linguistic information.

EXPERIMENT 4

Experiment 3 was the only experiment in which imaging produced voice effects in an implicit memory task. It was also the only experiment in which the size of the priming scores following imaging was sufficiently large to make the detection of these effects possible. The presence of voice effects in this condition suggests that generating spoken words can mimic to some degree the perceptual operations used when subjects are perceiving spoken words. Of course, both the stem completion task and the auditory imagery instructions required subjects to generate information, and this may have made the completion task more sensitive to the voice characteristics that subjects generated at study (a transfer-appropriate processing account; Blaxton, 1989; Roediger, Weldon, & Challis, 1989). Given the theoretical importance of this finding, we administered this task to another group of subjects in Experiment 4 to assess whether we could replicate the voice effects observed in Experiment 3.

Method

Forty-eight undergraduates from the same subject pool described above participated in this experiment in partial fulfillment of course requirements. Sixteen subjects were assigned to each study condition.

Results and Discussion

The proportions of words completed in the various study conditions are shown in Table 4. The findings of this experiment replicate those observed in Experiment 3. Specifically, as in Experiment 3, hearing produced more priming than imaging and reading (+.19, +.12, +.11, respectively), and study-to-test voice changes were observed in both the hear and the image conditions.

These observations were supported by a 2 (studied or nonstudied) × 3 (hear, image, or read) mixed factorial ANOVA. This analysis yielded a main effect of item type [$F(1,45) = 121.79$, MS_e = both 0.004] and a significant interaction between item type and study condition [$F(2,45) = 3.66$, $MS_e = 0.004$] (study condition $F < 1$). The reliable interaction indicated that priming varied as a function of study condition (note that this interaction was not due to differences in baseline performance across the three study conditions, $F < 1$). As in Experiment 3, hearing a word at study again produced considerably more priming than imaging or reading it [$t(30) = 2.16$; $t(30) = 2.34$], whereas imaging and reading did not yield reliably different priming effects [$t(30) < 1$]. Changing voice between study and test reliably decreased the priming effects obtained in both the hear and the image conditions [hear, $t(15) = 2.37$; image, $t(15) = 2.21$].

As in Experiment 3, stem completion yielded voice effects in both the hear and the image study conditions, bolstering the claim that imaging printed words as spoken by a familiar talker mimics to some degree the perceptual processes engaged by spoken words. Although imaging may reproduce speech perception operations, it does not substantially improve priming over reading, thereby supporting the earlier conclusion that imaging is not an exact replica of hearing.

GENERAL DISCUSSION

The findings of these four experiments can be summarized in four main points. First, hearing words at study

Table 4
Experiment 4, Proportion of Words Correctly Completed in the Stem Completion Task as a Function of Study Condition (Hear, Image, or Read), Item Type (Studied or Nonstudied), and Test Voice (Same or Different Voice)

	Voice Studied				
Study Condition	Same	Different	Studied (M)	Nonstudied	Priming (S − NS)*
Hear	.48	.41	.45	.26	+.19
Image	.43	.38	.41	.29	+.12
Read			.38	.27	+.11

*Priming scores (studied − nonstudied) in the completion task are displayed in this column.

consistently produced larger priming effects than imaging spoken words on three auditory implicit tests (identification in noise, identification of low-pass filtered signals, and stem completion). Second, hearing words produced voice effects (same voice > different voice) on all the implicit tests, whereas imaging produced voice effects only on the stem completion test. The priming effects were relatively larger in stem completion than in the other tests, so it was arguably more sensitive to voice changes for this reason. The fact that both imaging and completing stems required subjects to generate information might have also promoted voice effects in this task. Third, study-to-test voice changes and auditory imagery instructions differentially affected implicit and explicit tests. Specifically, study-to-test perceptual changes affected priming, but not episodic recognition or cued recall. On the other hand, auditory imagery reduced cued recall performance relative to both hearing and reading, whereas it yielded as much priming as reading in our implicit tasks. These variables thereby created dissociations between performance on implicit and explicit tests and showed that the implicit tests were not reflecting conscious recollection.

The results of our experiments have important implications for understanding the role of perceptual information in memory. Specifically, the finding of voice and modality effects in the implicit memory tests selected for this investigation supports the notion that subjects encode perceptual attributes at study and that these attributes affect processing on implicit tests. This finding is consistent with models of memory processes in which perceptual details and linguistic information are assumed to be combined into unified (see Goldinger, 1998; Hintzman, 1986; Hintzman et al., 1972) or separate (see MacKay, 1992; Schacter & Church, 1992) memory representations. However, priming even with changes in voice and modality was still robust on all the implicit tests, indicating that abstract linguistic information contributes to priming in addition to perceptual attributes. This conclusion fits well with the literature on visual implicit tests, where similar contributions from abstract and perceptually specific processes have been shown to promote priming (see Roediger & McDermott, 1993). Furthermore, it is supported by recent neuroimaging data illustrating that visual and auditory within-modality priming effects reflect reduced activity in similar regions of the extrastriate cortex (Badgaiyan, Schacter, & Alpert, 1999). On the other hand, neither cued recall nor recognition memory displayed the same sensitivity to perceptual changes, suggesting that these tasks (at least in our experiments) rely primarily on abstract linguistic information.

The finding that imaging and hearing produced roughly equivalent voice effects in the stem completion task of Experiments 3 and 4 supports the notion that imagined and directly perceived events are similar, at least at the perceptual level (for a similar argument in the visual domain, see Finke, 1980; Kosslyn, 1981; McDermott & Roediger, 1994). Of course, the claim that imagery creates some material, akin to the actual stimulus in perception, is not meant to suggest that imaging spoken words is a perfect replica of hearing them. Indeed, we consistently found that hearing produced more priming than imaging, which would be expected if it is assumed that differences exist between the processes engaged by actual spoken words and those involved in imaging. The only exception to this finding of greater priming for spoken than imagined words was observed in the −10 S/N ratio condition of Experiment 1. Under these difficult listening conditions, hearing and imaging yielded equivalent priming effects in perceptual identification. This was the only condition in which we were able to replicate the Stuart and Jones (1996) finding of equivalent priming effects for hearing and imaging. Under these especially harsh listening conditions, however, the priming effects produced by hearing words at study were considerably smaller than in the other experiments (i.e., there was a floor effect), and, not surprisingly, there was no evidence of voice effects. Even though we do not know what level of masking noise Stuart and Jones used in their experiment, we are inclined to conclude from our series of experiments that imaging and hearing yield equivalent priming effects only when identification is disrupted by a high level of masking noise.

Although imaging and reading share visual-orthographic and abstract phonological information, imaging also involves generating voice-specific information, and the latter should make imaging different from reading in its effects on priming. Interestingly, we found that although imaging did not improve priming above reading, it impaired cued recall performance relative to both reading and hearing (a trend in the same direction was also observed in recognition memory). Because our explicit tasks were shown not to be sensitive to perceptual information, if generation of perceptual information were the only dimension that differentiated imaging from the other conditions, we should have observed equivalent levels of explicit performance across all our study conditions. However, we found that imaging words yielded lower cued recall performance than did reading or hearing, and thus it is reasonable to assume that some other factor was at work here. Imaging words spoken by a familiar talker probably requires that processing resources be allocated to more surface-like processing than in the case of hearing or reading words at study. Indeed, subjects must translate orthographic codes into phonological word codes, retrieve the voice qualities of a familiar talker, and then adjust phonological codes to match these voice qualities. These latter operations require resources that may otherwise be allocated to other linguistic operations such as conceptual analyses. Accordingly, the condition requiring auditory imagery at study may have been functionally equivalent to a divided attention manipulation, affecting explicit memory more than priming. In fact, divided attention manipulations have large effects on explicit tests, but weaker or no effects on priming in implicit tests (Hawley & Johnston, 1991; Jacoby & Dallas, 1981; Mulligan & Hartman, 1996).

One puzzle raised by the results of the imagery manipulation in Experiments 3 and 4 is why imaging spoken words yielded no greater priming than did reading words, while at the same time it produced reliable voice effects. That is, imaging words spoken in the same voice as at test created more priming than when the imagined voice and test voice did not match. There is no easy explanation for this seeming paradox. In the study of visual imagery in implicit tests, enhanced priming has been obtained (McDermott & Roediger, 1994), so some effect on auditory tests would not be surprising. However, visual imagery has produced enhanced priming primarily in experimental conditions that promote conceptual processing such as imaging the referent of a printed word or generating the name of a visually presented object. Of course, these forms of imagery are quite different from generating spoken words with predetermined voice qualities because the latter promotes fine-grained perceptual processes rather than conceptual operations. Clearly, the main finding we report—that priming from heard voices is greater than that from imagined voices—is secure. The resolution to the puzzle of obtaining voice effects from generated spoken words while failing to obtain enhanced priming from imaging must await future research.

Other researchers have reported a generally similar pattern of results regarding the sensitivity to perceptual information of implicit tasks (see Church, 2000; Craik et al., 1994; Pilotti et al., 2000; Schacter & Church, 1992). However, a handful of studies have also reported voice effects in recognition memory (Craik & Kirsner, 1974; Palmeri et al., 1993; Pilotti et al., 2000; Sheffert, 1998a). The absence of voice effects in the recognition task of our experiment may be considered surprising in view of the fact that we included a familiarization phase and a study task that assured the encoding of voice information. However, because subjects encountered the test voice not only at study but also during the familiarization and the generalization phases of the experiment, it is possible that voice information in this experimental setting became less useful as a cue for recognizing studied words, due to cue overload. If this is true, why did Geiselman and Glenny (1977) find voice effects in recognition memory with a similar experimental procedure? As discussed earlier, subjects in their study formed auditory images of words spoken by familiar talkers and then repeated to themselves these words for 10 sec. In our study, subjects simply rated auditory images. Therefore, it is reasonable to assume that in the Geiselman and Glenny study rehearsal strengthened the association between study words and voice characteristics, thereby making voice a more useful retrieval cue for recognition memory.

In summary, our results support the notion that although perceptual information may be encoded during study, its use at test depends on the nature of the operations (implicit or explicit) engaged by the memory task. Interestingly, this perceptual information may arise not only from stimuli that subjects have actually perceived, but also from stimuli that subjects have internally generated (see McDermott & Roediger, 1994). The findings of this investigation shed light on the role of perceptual information in memory tasks and show that test orientation (implicit or explicit), type of cue (word stem or other degraded token), and level of noise can all affect the sensitivity of a test in revealing aftereffects of perceptual experiences.

REFERENCES

Badgaiyan, R. D., Schacter, D. L., & Alpert, N. M. (1999). Auditory priming within and across modalities: Evidence from positron emission tomography. *Journal of Cognitive Neuroscience*, **11**, 337-348.

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 657-668.

Church, B. A. (2000). *Implicit memory for formant frequency information: Perceptual specificity effects in auditory priming did not depend on the explicit detection of acoustic changes.* Unpublished manuscript.

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. M., & Kirsner, K. (1974). The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology*, **26**, 274-284.

Craik, F. I. M., 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.

Donnelly, R. E. (1988). *Priming across modality in implicit memory: Facilitation from auditory presentation to visual test of word-fragment completion.* Unpublished doctoral dissertation, University of Toronto, Toronto.

Finke, R. A. (1980). Levels of equivalence in imagery and perception. *Psychological Review*, **87**, 113-132.

Geiselman, R. E., & Bjork, R. A. (1980). Primary versus secondary rehearsal in imagined voices: Differential effects on recognition. *Cognitive Psychology*, **12**, 188-205.

Geiselman, R. E., & Glenny, J. (1977). Effects of imagining speakers' voices on the retention of words presented visually. *Memory & Cognition*, **5**, 499-504.

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.

Hawley, K. J., & Johnston, W. A. (1991). Long-term perceptual memory for briefly exposed words as a function of awareness and attention. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 807-815.

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.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306-340.

Jacoby, L. L., & Witherspoon, D. (1982). Remembering without awareness. *Canadian Journal of Psychology*, **36**, 300-324.

Kirsner, K. (1974). Modality differences in recognition memory for words and their attributes. *Journal of Experimental Psychology*, **102**, 579-584.

KOSSLYN, S. M. (1981). The medium and the message in mental imagery: A theory. *Psychological Review*, **88**, 46-66.

KUČERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

LIGHT, L. L., LA VOIE, D., VALENCIA-LAVER, D., ALBERTSON OWENS, S. A., & MEAD, G. (1992). Direct and indirect measures of memory for modality in young and older adults. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1284-1297.

MACKAY, D. G. (1992). Constraints on theories of inner speech. In D. Reisberg (Ed.), *Auditory imagery* (pp. 121-149). Hillsdale, NJ: Erlbaum.

MCDERMOTT, K. B., & ROEDIGER, H. L., III (1994). Effects of imagery on perceptual implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1379-1390.

MULLIGAN, N. W., & HARTMAN, M. (1996). Divided attention and indirect memory tests. *Memory & Cognition*, **24**, 453-465.

NAIRNE, J. S., & PUSEN, C. (1984). Serial recall of imagined voices. *Journal of Verbal Learning & Verbal Behavior*, **23**, 331-342.

NUSBAUM, H. C., PISONI, D. B., & DAVIS, C. K. (1984). *Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words* (Research on Speech Perception, Progress Report No. 10, pp. 357-376). Bloomington: Indiana University 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.

PILOTTI, M., BERGMAN, E. T., GALLO, D. A., SOMMERS, M., & ROEDIGER, H. L., III (2000). Direct comparison of auditory implicit memory tests. *Psychonomic Bulletin & Review*, **7**, 347-353.

RAJARAM, S., & ROEDIGER, H. L., III (1993). Direct comparison of four implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 765-776.

ROEDIGER, H. L., III, & BLAXTON, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word fragment completion. *Memory & Cognition*, **15**, 379-388.

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.

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. 3-42). Hillsdale, NJ: Erlbaum.

ROEDIGER, H. L., III, WELDON, M. S., STADLER, M. L., & RIEGLER, G. L. (1992). Direct comparison of two implicit memory tests: Word fragment completion and word stem completion. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1251-1269.

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.

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.

STUART, G. P., & JONES, D. M. (1996). From auditory image to auditory percept: Facilitation through common processes? *Memory & Cognition*, **24**, 296-304.

WELDON, M. S., & ROEDIGER, H. L., III (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition*, **15**, 269-280.

NOTES

1. In Experiments 1 and 3, not all the between-subjects conditions involved the same number of subjects. The additional data provided by these subjects were intended to increase the power to detect an effect in these conditions. It is important to note, however, that although these data enhanced the power of the statistical tests conducted in these conditions, they did not change the overall pattern of results obtained with 16 subjects.

(Manuscript received December 6, 1999;
revision accepted for publication April 15, 2000.)