Journal of Experimental Psychology:
Learning, Memory, and Cognition
1994, Vol. 20, No. 3, 521-533

0278-7393/94/$3.00

Perceptual Specificity of Auditory Priming: Implicit Memory for Voice Intonation and Fundamental Frequency

Barbara A. Church and Daniel L. Schacter

Five experiments explored the effect of acoustic changes between study and test on implicit and explicit memory for spoken words. Study–test changes in the speaker's voice, intonation, and fundamental frequency produced significant impairments of auditory priming on implicit tests of auditory identification and stem completion but had little or no effect on explicit recall and recognition tests (Experiments 1–4). However, study–test changes in overall decibel level had no effect on priming on an auditory stem-completion test or on cued-recall performance (Experiment 5). The results are consistent with the idea that fundamental frequency information is represented in a perceptual representation system that plays an important role in auditory priming.

Recent research has been sparked by the demonstrations of dissociations between explicit memory, the conscious recollection of previously studied material, and implicit memory, the facilitation of task performance by previously studied material without intentional recollection (for reviews, see Richardson-Klavehn & Bjork, 1988; Roediger, 1990; Roediger & McDermott, 1993; Schacter, 1987; Schacter, Chiu, & Ochsner, 1993; Shimamura, 1986). The majority of research concerning implicit memory has focused on tasks involving visual priming of words and objects, when prior exposure to target items produces facilitated performance on a subsequent implicit test. In addition, most theories of implicit memory phenomena invoke hypotheses about the characteristics of visually based processes or systems (cf. Jacoby, 1983; Kirsner, Dunn, & Standen, 1989; Masson & Macleod, 1992; Roediger, 1990; Schacter, 1990; Squire, 1987; Tulving & Schacter, 1990).

In contrast to the plethora of relevant research in the visual domain, only a handful of studies have examined implicit memory in the auditory domain. For example, several studies have demonstrated repetition priming effects across delays of at least several minutes on auditory word- and sentence-identification tasks, in which subjects attempt to identify previously studied and nonstudied items that have been masked in white noise (Ellis, 1982; Franks, Plybon, & Auble, 1982; Jackson & Morton, 1984; Kempley & Morton, 1982) or to make ratings of the noise level (Jacoby, Allan, Collins, & Larwill, 1988). Priming has also been found on the auditory stem-completion task, in which subjects hear initial syllables of studied and nonstudied items and are asked to complete them with the first words that come to mind (Bassili, Smith, & MacLeod, 1989; McClelland & Pring, 1991).

Note also that numerous experiments have examined indirect or phonological priming effects involving delays of less than a second between the prime and the target (cf. Goldinger, Luce, & Pisoni, 1989; Jakimik, Cole, & Rudnicky, 1985; Slowiaczek, Nusbaum, & Pisoni, 1987; Slowiaczek & Pisoni, 1986). These effects are extremely short-lived and are probably based on mechanisms different from those that subserve priming effects that persist across retention intervals of several minutes or more.

In a series of experiments, Schacter and Church (1992) attempted to build on these earlier findings and systematically explore characteristics of and ideas about auditory implicit memory. The theoretical motivation for the experiments was molded by the perceptual representation systems (PRS) framework described in several articles (e.g., Schacter, 1990, 1992; Schacter, Cooper, & Delaney, 1990; Schacter, Rapcsak, Rubens, Tharan, & Laguna, 1990; Tulving & Schacter, 1990). The premise of this multiple memory systems framework is that implicit memory effects observed on data-driven tasks such as perceptual identification, stem and fragment completion, and object decision depend primarily on the PRS, a presemantic system composed of a number of subsystems that process and represent information about the physical form and structure, but not the meaning or associative properties, of words, objects, and other stimuli. The PRS is assumed to be a cortically based system that is distinct from a system based in limbic structures (e.g., the hippocampus) that is necessary for explicit, episodic retrieval. The primary function of PRS is believed to be perceptual recognition. The empirical motivation for proposing that PRS mediates data-driven priming was initially provided by some characteristics of visual priming on data-driven implicit tasks, namely that it (a) does not require elaborative processing at study, (b) shows a large degree of modality specificity, and (c) is, at times, sensitive to within-modality changes in surface feature information between study and test (for review, see Kirsner et al., 1989; Richardson-Klavehn & Bjork, 1988; Roediger & Blaxton, 1987; Roediger, Weldon, & Challis, 1989; Schacter, 1990). Independent evidence for the existence of PRS subsystems is provided by neuropsychological literature indicating a dissociation between form and semantic representations in patients with

---

Barbara A. Church and Daniel L. Schacter, Department of Psychology, Harvard University.

This research was supported by National Institute on Aging Grant R01-8441 and by National Institute of Mental Health Grant R01-MH45398-01A3. We thank Peter Gordon, David Pisoni, Ken Stevens, Henry L. Roediger III, Steve Goldinger, and an anonymous reviewer for useful suggestions and discussion.

Correspondence concerning this article should be addressed to Daniel L. Schacter, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, Massachusetts 02138. Electronic mail may be sent to dls@isr.harvard.edu.

reading and object-processing deficits (see Schacter, 1990, 1992).

Although the PRS framework originally focused on visual subsystems, converging evidence points to an analogous auditory PRS. For example, patients characterized by *word-meaning deafness* cannot understand spoken words but exhibit relatively intact repetition and writing to dictation of the same words (e.g., Ellis, 1984; Kohn & Friedman, 1986). Schacter and Church (1992) found that study tasks that involved semantic elaboration increased explicit recognition and cued-recall performance relative to nonsemantic study but had little or no effect on the priming of auditory word-identification and stem-completion tasks. Moreover, Schacter, McGlynn, Milberg, and Church (1993) observed normal levels of priming on auditory word-identification tasks in a patient with word-meaning deafness who exhibited impaired comprehension of the target words. These observations support the idea that auditory priming depends, to a large extent, on a presemantic PRS subsystem.

In addition, several experimenters have found that study–test shifts in modality reduce or eliminate priming on auditory word-identification (Ellis, 1982; Jackson & Morton, 1984) and stem-completion (Bassili et al., 1989; McClelland & Pring, 1991) tasks. However, cross-modal priming does occur, and McClelland and Pring have found that such priming is greatest in study conditions, such as naming, that maximize auditory phonological processing. This observation suggests that the processing of phonological information may play a crucial role in the cross-modal priming effects that have been observed on auditory stem-completion tasks.

Evidence concerning the effects of within modality study–test changes in speaker's voice on auditory priming is less clear-cut. Jackson and Morton (1984) found that study–test changes in the speaker's voice do not affect priming on a word-identification-in-noise task. Schacter and Church (1992) reported similar findings, but they also reported that study–test changes in speaker's voice do reduce priming on a stem-completion test. However, voice-change effects were not observed when the stems were embedded in noise. Schacter and Church concluded that the absence of voice-change effects in their experiments and in Jackson and Morton's (1984) study could be attributed to the presence of white noise on the identification and completion tests (see Goldinger, 1992, for additional data and qualifications).

The fact that priming does occur when there is a study–test change in the speaker's voice indicates that PRS contains an abstract component. Nevertheless, Schacter and Church's (1992) finding of a reduction of priming on a stem-completion test when the speaker's voice changed between study and test suggests that voice information is represented in PRS and can influence priming of word completion. However, little is known about the nature of the voice information that influenced priming in Schacter and Church's experiments. Does it depend on the representation of specific acoustic information within the PRS? If so, what kind of information? Alternatively, it is also possible that the voice-change effects observed in Schacter and Church's experiments were not mediated by acoustical representations but rather by a semantically based "gender code" that was associated with target words during study and then influenced priming. Similarly, the gender of the speaker might have influenced the connotation of the words during study, and consequently, a change in the connotation of the word between study and test could have reduced priming. The experiments described in this article were designed to address these issues.

Experiment 1 examines voice-change effects on an auditory word-identification task in an attempt to generalize Schacter and Church's (1992) findings. Experiment 2 explores whether within-voice variables can affect priming and addresses the possibility that gender coding mediates voice-change effects by varying a single speaker's intonation between study and test. The effects of small study–test changes in the fundamental frequency of a speaker's voice are examined in Experiments 3 and 4 in an attempt to further understand the specificity of the information available to PRS and to address the possibility that semantic-level information mediates voice-change effects. Experiment 5 explores the specificity of priming by manipulating the overall intensity of the words.

Experiment 1

Because voice-specific priming has thus far been observed only on the stem-completion test, questions concerning the generalizability of this effect remain. For example, it would be desirable to determine whether voice-specific priming can be observed on an auditory identification test that does not use white noise (cf. Schacter & Church, 1992). To accomplish this objective, we developed a word-identification task in which words were degraded with a low pass filter. A low pass filter reduces the decibel level of a distribution of frequencies above a chosen cutoff point. This filtering process generally leaves fundamental frequency and prosodic contour information intact while reducing the higher frequency information available in the speech signal. Subjectively, words that have been low pass filtered sound rather muffled and inaudible.

In the experiment, subjects initially heard a list of 24 words spoken by six different speakers (three men and three women) and performed a nonsemantic encoding task in which they rated the clarity of the speaker's enunciation of each target word; significant voice-change effects have been observed previously following this encoding task (Schacter & Church, 1992). After a brief delay during which the subjects engaged in a city-name-generation task as a distractor, they were asked to identify 24 studied and 24 nonstudied words that had been subject to low pass filtering. After the implicit low pass filter identification task, the subjects were asked to make explicit recognition judgments about the words. In both tasks, half the studied items were spoken in the same voice as at study and half the words were spoken in a different voice.

Method

Subjects. Twenty-four Harvard University undergraduates participated in the experiment in exchange for a $5 payment.

Materials. Target materials comprised 48 words that were divided into two subsets of 24 words for the encoding task. The two subsets were matched for frequency, first letter, number of syllables, and length (Graf & Williams, 1987; Kucera & Francis, 1967). Target words were recorded at normal conversational levels by six speakers (three men and three women). Each word was recorded by one man and one woman, so that changes between study and test always included a change in the speaker's gender. The words were recorded on a Macintosh computer using a MacRecorder at a sampling rate of 22 kHz. Each item for the identification task was filtered with the low pass filter function found in the SoundEdit program. The filter extracted a distribution of frequencies above 2 kHz and reduced these by 20 dB during a pass through the filter. It also reduced the decibel levels of a distribution of frequencies between 1 kHz and 2 kHz by a range of 5 to 20 dB, with the highest frequencies being reduced the most in a steeply sloping function. Each item was passed through the filter three times, allowing a maximum reduction of 60 dB for the higher frequencies.

Four tapes were used to record two versions of each study list, two versions of the low pass filter identification task, and two versions of the recognition test that used the same presentation orders as the identification task. Because any item spoken by a man on Version 1 of any of the tapes would be spoken by a woman on its counterpart and vice versa, all of the items and speakers could be completely counterbalanced. Each study list tape included 24 words spoken clearly. The low pass filter identification tapes included 48 filtered words, 24 that had been presented at study and another 24 items that had not been presented. The recognition test tapes included the same 48 words spoken clearly: Half of them had appeared on the study list, and half of them had not appeared on the study list.

The tapes were presented using a cassette deck and headphones. A booklet was provided for the subjects to record all their responses. The first page of the booklet contained a 4-point number scale for the subjects to use in rating how clearly the speaker enunciated each word (well, moderately well, moderately poorly, or poorly) and 24 spaces for the subject's responses to each item. The second page contained 15 letters and spaces for the subjects to write in the city name they generated for each letter. The last two pages contained 48 spaces each in which the subjects could write down the words they heard during the low pass filter identification task and respond by writing yes or no to the items on the recognition test.

Design and procedure. We used a 2 × 2 × 2 within-subject design. The variables were type of test (auditory identification vs. auditory recognition), item type (studied vs. nonstudied), and speaker's voice (same vs. different). The experiment was completely counterbalanced so that each item appeared equally as often in each of the experimental conditions defined by the orthogonal combination of the variables. Also, each item was spoken equally as often by a male and a female voice.

All of the subjects were tested individually. During the encoding task, 24 words were presented auditorily, and subjects were asked to rate the speaker's clarity of enunciation on a 4-point scale. There were 5 s between the items for subjects to make their ratings. Subjects then completed a distractor task, during which they generated the names of 15 cities beginning with the letters given in their booklets. The task generally engaged the subject for 3 or 4 min.

After the distractor task, subjects were given the implicit low pass filter identification task, during which they heard 48 filtered words and were asked to write down the first word that came to mind in response to each item. There were 7 s between the items for the subjects to write down their answers. Once subjects had completed the low pass filter identification task, they were given an explicit auditory recognition test. They heard 48 clearly spoken words and were asked to judge which of those words had been presented during the encoding task. They were warned that many of the words had also been presented in the filter task and that it was important that they only provide a yes response for items they remembered specifically from the encoding task. Also, subjects always received the version of the recognition test that had the items in a different presentation order and with different voices than the filter task. This helped to reduce the number of false alarms the subjects reported. There were 7 s between the items. After the experiment subjects were debriefed.

Table 1

Proportion of Studied and Nonstudied Target Words Reported on the Filter Identification and Recognition Tests as a Function of Speaker's Voice in Experiment 1

Type of test	Studied target words: Same voice	Studied target words: Different voice	Studied target words: *M*	Nonstudied target words
Identification	.80	.72	.76	.59
Recognition	.76	.80	.78	.24
M	.78	.76	.77	.42

Note. On the recognition test, studied items reported represents the proportion of studied items called *old* (hit rate); nonstudied items reported represents the proportion of nonstudied items called *old* (false-alarm rate).

Results

Table 1 presents the proportion of studied items and nonstudied items correctly identified on the low pass filter identification task and the proportion of hits and false alarms recognized during the recognition test as a function of same or different speaker's voice. As indicated in Table 1, there was evidence of priming on the low pass filter identification task. There was also evidence of an effect of voice change on priming, but not on recognition.

A two-tailed t test that compared the proportion of studied and nonstudied items correctly identified on the low pass filter identification task achieved significance, $t(23) = 2.93, p < .01$, thus confirming that priming occurred. A 2 × 2 analysis of variance (ANOVA) was performed on the priming scores (i.e., the proportion of studied items minus nonstudied items correctly identified on the filter task), and recognition scores (i.e., the proportion of hits minus false alarms on the recognition test), with type of test and speaker's voice as the within-subject variables. There was a main effect of type of test, $F(1, 23) = 63.98, p < .0001, MS_e = .053$, indicating higher recognition scores than priming scores. The overall main effect of speaker's voice was not significant, $F(1, 23) = 1.68, p > .20$, $MS_e = .009$, but there was a significant interaction between type of test and speaker's voice, $F(1, 23) = 5.27, p < .05$, $MS_e = .015$. Planned comparisons revealed that on the low pass filter identification task, there was significantly more priming when speaker's voice was the same at study and test than when the voice changed between study and test, $t(23) = 2.42, p < .05$. However, there was no effect of voice changes on the recognition test, $t(23) < 1$. All t tests were two-tailed.

Discussion

Experiment 1 revealed a significant priming effect on an auditory word-identification task when the items were degraded using a low pass filter. More important, it revealed a significant reduction of priming on this task when the voice was changed between study and test, together with no effect of the identical voice-change manipulation on recognition memory. This dissociation between implicit and explicit memory is consistent with the hypothesis that auditory priming is mediated by a PRS that stores information about the acoustical form of the words, whereas explicit memory is mediated by an episodic system that generally relies on conceptually driven processes.

The primary goal of this experiment was to ascertain whether voice changes affect priming on an auditory identification task when words are not embedded in noise. The clear evidence for a voice-change effect indicates that the absence of voice-change effects in previous experiments using identification-in-noise tasks (Jackson & Morton, 1984; Schacter & Church, 1992) was probably not simply a function of using degraded items on an identification task but is more likely attributable to specific features of white noise. The filter task used in Experiment 1 does significantly degrade target items, but the nature of the degradation differs from the degradation produced by noise in at least two ways: (a) lower frequency information, including fundamental frequency and prosodic contour information, is completely preserved in the filter but not the noise task, and (b) information is subtracted from the speech signal in the filter task, whereas it is added in the noise task.

We cannot easily determine which of these differences between the two auditory identification tasks is responsible for the differing effects of changing speaker's voice between study and test on priming. If, however, such acoustic features as fundamental frequency and prosodic contour that are preserved in the filter task are crucial to the observed voice-change effects, then study–test changes to these aspects of the acoustic signal should produce significant decreases in the magnitude of auditory priming.

Prosodic contour shifts are part of the normal variation within an individual's speech. These contour shifts convey emotional, syntactic, phrasal, and lexical information, and they are a fundamental part of the speech signal (e.g., Blumstein & Goodglass, 1972; Heilman, Bowers, Speedie, & Coslett, 1984; Shapiro & Danly, 1985). However, with the exception of lexical prosody, the information conveyed is independent of the semantic content of an utterance. Emotional contour imparts information about the emotional state of the speaker, whereas syntactic and phrasal prosody impart information about the structure and type of sentence or phrase being expressed. If the prosodic contour differences that exist when different speakers' voices are used contribute to the voice-change effect, then study–test changes of prosodic contour within a single voice should produce similar effects.

Alternatively, it is possible that the specific acoustic features of the speaker's voice are not an important component of the observed voice-change effect on priming. These effects may instead be produced by the encoding and retrieval of a gender code that subjects generate for each speaker. Subjects may encode or tag each voice according to the gender of the speaker, and when that gender code is reinstated on an implicit test priming is higher than when the gender code is not reinstated. Many researchers have found circumstances under which variations in voice information influence recall and recognition (e.g., Craik & Kirsner, 1974; Goldinger, Pisoni, & Logan, 1991; Martin, Mullennix, Pisoni, & Summers, 1989), and it has been suggested that voice information is explicitly retained when that information can be related connotatively to the word (Geiselman & Bellezza, 1976, 1977; Geiselman & Crawley, 1983). Previous research examining explicit memory for voice information has indicated that subjects may interpret information differently depending on the gender of the speaker, thus producing a semantic link between the word and the speaker's gender (Geiselman & Bellezza, 1976). In Experiment 1 and in our previous demonstrations of voice-specific priming (Schacter & Church, 1992), study–test changes in the speaker's voice always included a change in the speaker's gender. Experiment 2 was designed to examine the possibility that a semantic gender code is responsible for voice-specific priming. In addition, it allows us to test the hypothesis that study–test changes in the prosodic contour of a single speaker's voice will reduce priming.

Experiment 2

The basic design of Experiment 2 was similar to Experiment 1 except that it explored the effect of both emotional and phrasal prosodic contour changes within a single voice rather than changes in speakers' gender across multiple voices. The words were recorded by only one speaker, but each word was spoken in either one of two emotional intonations (happy or angry) or one of two phrasal intonations (question or statement). During the filter task and subsequent recognition test, the intonation for half of the studied items was the same as at study (e.g., angry–angry or question–question), and for the other half the intonation was changed between study and test (e.g., angry–happy or question–statement).

Method

Subjects. Forty-eight Harvard University undergraduates were paid $5 for their participation.

Materials. Target materials consisted of 48 familiar words. The words were divided into two sets of 24 that were matched for frequency, concreteness, first letter, and length (Pavio, Yuille, & Madigan, 1968). A single female speaker recorded the words into a Macintosh computer using a MacRecorder with a sampling rate of 22 kHz. Each word was recorded in four different intonations, two emotional intonations (happy and angry), and two linguistic intonations (statement and question). Pilot subjects were asked to judge the intonation of the speaker for both the emotionally intoned words (i.e., happy–angry judgment) and the linguistically intoned words (i.e., question–statement judgment) on a two-item forced-choice test. The subjects judged the intonation correctly 98% of the time in both intonation type conditions. The items were filtered using the same procedure described in Experiment 1.

We recorded four versions of each of the 24 word-study lists using two emotional intonation versions and two phrasal intonation versions. Any given item presented in an angry or questioning intonation on one version would be presented in a happy or statement intonation on the other version. Within a particular study list, half of the target items were spoken in one intonation (e.g., angry or questioning), and half of the items were spoken in the other intonation (e.g., happy or statement). The filter task and recognition test each had four versions of the 48 (24 studied and 24 nonstudied) words recorded, two using emotional intonation and two using phrasal intonation. Each item presented in an angry or questioning intonation on one version was presented in a happy or statement intonation on the other version, and vice versa. As in Experiment 1, all study and recognition items were presented clearly at normal conversational levels. We presented the tapes using a cassette deck and headphones, and the subjects recorded their responses using the same type of booklet described in Experiment 1.

Design and procedure. We used a mixed-factorial design. The between-subjects variable was intonation type (emotional vs. phrasal), and the within-subject variables were item type (studied vs. nonstudied), type of test (low pass filter identification vs. recognition), and intonation change (same vs. different). The experiment was counterbalanced such that each item appeared equally often in each experimental condition. The procedures were identical to those used in Experiment 1. Subjects completed in order the clarity rating encoding task, the city-name-generation distractor task, the implicit filter task, and finally the explicit auditory recognition test.

Results

Table 2 presents the proportion of studied and nonstudied items identified correctly on the low pass filter identification task as a function of type of intonation (emotional vs. phrasal) and intonation change (same vs. different). As indicated in Table 2, there were similar amounts of priming on the low pass filter identification task in both intonation type conditions, and, more important, there was an effect of intonation change in both intonation type conditions. A two-tailed t test that compared the overall proportion of studied and nonstudied items identified correctly was highly significant, $t(47) = 7.94, p < .001$. A 2 × 2 ANOVA was performed on the priming scores (i.e., the proportion of studied items minus nonstudied items identified on the filter task), with type of intonation as the between-subjects variable and intonation change as the within-subject variable. There was a significant main effect of intonation change, $F(1, 46) = 15.92, p < .001, MS_e = .006$, indicating higher identification scores when the item was presented in the same intonation than in a different intonation at study and test. Neither type of intonation nor the interaction between intonation type and intonation change was significant (all Fs < 1).

Table 3 presents the proportion hits and false alarms as a function of intonation type and intonation change. A 2 × 2 ANOVA was performed on the corrected recognition scores (i.e., the proportion of hits minus false alarms), with type of intonation as the between-subjects variable and intonation change as the within-subject variable. There was a nonsignificant main effect of intonation change, $F(1, 46) = 1.80, p > .18$, $MS_e = .01$, and neither the main effect of intonation type nor the interaction between intonation type and intonation change approached significance (all Fs < 1).

Table 2
Proportion of Studied and Nonstudied Words Correctly Identified on the Filter Identification Test as a Function of Intonation Type and Intonation Change in Experiment 2

Type of intonation	Studied words: Same intonation	Studied words: Different intonation	Studied words: *M*	Nonstudied words
Emotional	.73	.62	.67	.49
Phrasal	.72	.64	.68	.51
M	.72	.63	.68	.50

Table 3
Proportion of Hits and False Alarms on the Recognition Test as a Function of Type of Intonation and Intonation Change in Experiment 2

Type of intonation	Studied words (hits): Same intonation	Studied words (hits): Different intonation	Studied words (hits): *M*	Nonstudied words (false alarms)
Emotional	.83	.78	.80	.25
Phrasal	.82	.81	.82	.23
M	.83	.80	.81	.24

A 2 × 2 × 2 ANOVA was performed on the priming and recognition scores to examine the effect of intonation changes between study and test on the two test and intonation types. There was a main effect of test, $F(1, 46) = 242.85, p < .0001$, $MS_e = .031$, indicating that the recognition scores were higher than the priming scores. There was a main effect of intonation change, $F(1, 46) = 23.57, p < .0001, MS_e = .009$, indicating an overall superiority for same intonation items versus changed intonation items. There was a nonsignificant interaction between intonation type and intonation change, $F(1, 46) = 1.84$, $p > .10, MS_e = .009$, and a marginally significant interaction between type of test and intonation change, $F(1, 46) = 2.87$, $p < .10, MS_e = .019$. No other effects approached significance (all Fs < 1).

A series of planned comparisons using two-tailed t tests showed a significant decrease in priming in the different intonation condition compared with the same intonation condition for both intonation types: emotional, $t(23) = 2.21$, $p < .05$; phrasal, $t(23) = 2.26, p < .05$. There were nonsignificant effects of intonation change on recognition performance for both intonation types: emotional, $t(23) = 1.00, p > .1$; phrasal, $t(23) < 1$. However, there was a trend for an intonation change effect on recognition in the emotional intonation condition, and this trend may be responsible for the fact that the interaction between intonation change and test type was only marginally significant.

A 2 × 2 ANOVA was also performed on the overall identification scores to determine whether there were any effects of the particular intonation in which a word was spoken. The between-subjects variable was intonation type (emotional vs. phrasal), and the within-subject variable was the particular intonation contour (e.g., angry vs. happy or statement vs. question) identified. There was a significant effect of intonation type, $F(1, 46) = 6.25, p < .02, MS_e = .047$, indicating that subjects in the emotional intonation condition identified more filtered words than did subjects in the linguistic intonation condition. However, the effect of the particular intonation contour heard by the subjects was not significant, and there was a nonsignificant interaction between intonation type and intonation contour, all $F\text{s}(1, 46) < 1$.

Discussion

Experiment 2 revealed significantly lower levels of priming when voice intonation was changed between study and test than when it was held constant, whereas there were nonsignificant effects of intonation changes on the recognition test. These findings are consistent with the hypothesis that prosodic contour information is retained by the auditory PRS system and that it can contribute to voice-change effects on auditory priming. Also, because we used the same materials in this experiment as Schacter and Church (1992) did in their identification in noise tasks, our findings add further support to the hypothesis that the lack of voice-change effects in the noise tasks is attributable to the presence of the white noise. The trend toward a voice-change effect on recognition in the emotional intonation condition indicates that though linguistic information and general prosodic contour information do not seem to drive explicit retrieval, emotional information may.

It is also clear from the results of Experiment 2 that a simple gender-coding effect cannot explain voice specificity effects on priming because such effects can be produced without a gender change in the speaker's voice. These results are also consistent with Palmeri, Goldinger, and Pisoni's (1993) study, which found that although recognition was more accurate if the word was tested in the same voice than in a different voice, it was impaired equally by within-gender and between-gender changes in voice. They concluded that gender coding could not be responsible for their subjects' reduction in performance when the speaker's voice was changed. Note, however, that Geiselman and Bellezza (1977) concluded that incidental retention of a speaker's voice occurs when the voice adds meaningful information to a sentence. If voice characteristics are retained when they add connotative information, then it is possible that intonation information is retained because it changes the connotation of target words by adding emotional or linguistic information. Although Experiment 2 provides evidence against gender coding as the basis for voice-specific priming, it does not rule out the possibility that voice-change effects depend on the encoding and retention of semantically based connotative information. We addressed this issue in Experiment 3.

Experiment 3

In Experiment 3, we manipulated a specific acoustic feature of a single speaker's voice: fundamental frequency (f_0). We made relatively small changes in f_0 between study and test that did not involve any alteration of emotional or linguistic connotation, and we examined whether priming was lower in a different f_0 condition than in a same f_0 condition.

There are several reasons why f_0 is of particular interest to us. First, f_0 differences are present in our previous experiments in which voice-change effects on auditory priming have been observed. An analysis of 40 fundamental frequency points from two words uttered by each speaker of the male–female pairs in Experiment 1 revealed significant between-voice differences in mean fundamental frequency, $ts(78) = 7.69$, 2.77, and 11.99, for each of the three pairs. A similar analysis of the stimuli used in Experiment 2 revealed a significant difference between the mean f_0 of the happy and angry intonations used in the emotional condition and between the mean f_0 of the statement and question intonations used in the phrasal condition, $ts(78) = 2.42$ and 2.13, respectively.

Second, f_0 is the major determinant of the pitch of a speaker's voice (Lieberman, 1961) and hence would be expected to play a critical role in acoustic representation. Third, it seemed unlikely that small overall changes in f_0 would yield important connotative information. Thus, if f_0 changes do produce voice-specific priming, we can be reasonably confident that the effect is not attributable to semantic-level processes. However, to investigate further the possible role of such processes, we included both semantic (meaning judgment task) and nonsemantic (clarity rating task) encoding manipulations in Experiment 3. The reasoning here was straightforward: If changing f_0 between study and test has similar effects following both semantic and nonsemantic study tasks, then it is unlikely that conceptual factors play much of a role in the observed specificity effect. This manipulation also allowed us to examine the possibility that priming effects on the low pass filter identification tasks are influenced by explicit memory. Previous findings have indicated that semantic versus nonsemantic encoding manipulations have little or no effect on auditory identification-in-noise and stem-completion tests (Schacter & Church, 1992) despite large effects on explicit memory, and we expected a similar finding with the low pass filter identification task.

Method

Subjects. Sixty-four Harvard University undergraduates participated in exchange for a $5 payment.

Materials. Target materials were the same as those used in Experiment 2. The words were originally recorded by a single male speaker into a Gateway 2000 IBM-compatible computer using the Kay CSL 4300 acoustic package at a sampling rate of 10 kHz. For each of the items, the f_0 of the speaker's voice was estimated and then lowered by 10%. A resynthesized duplicate of each item with the lowered f_0 was made using the linear predictive coding algorithms of the Kay ASL software package. The items were then put through a low pass Kaiser filter using the same software. The filter was set to reduce the filtered frequencies by a maximum of 60 dB, and a midpoint of 1.8 kHz was used for the filter's frequency by decibel slope.

Four tapes were used to record two versions of each of the 24-word study lists, two versions of the low pass filter identification task, and two versions of the recognition test that corresponded to the identification task versions. Because any item spoken in a lowered f_0 on the first version would be spoken with the speaker's original f_0 on the second version and vice versa, all of the items and f_0s could be completely counterbalanced. The tapes were made in the same way as the tapes in Experiment 1.

We presented the tapes using a cassette deck and headphones. A booklet similar to the ones used in the previous experiments was provided. The only difference between this booklet and the previous booklet was that the first page of the booklet used for the encoding task contained a 4-point numeric scale for the subjects to use in rating either how clearly the speaker enunciated each word (well, moderately well, moderately poorly, or poorly) or the number of meanings for each word (one, two, three, or four or more meanings).

Design and procedure. We used a 2 × 2 × 2 × 2 mixed factorial design. The between-subjects variable was type of encoding task (meaning vs. clarity). The within-subject variables were type of test (identification vs. recognition), item type (studied vs. nonstudied), and speaker's f_0 (same vs. changed). The experiment was completely counterbalanced so that each item appeared equally often in each of the experimental conditions defined by the orthogonal combination of the variables. Also, each item was spoken equally often in the lowered and the original f_0. The procedures were identical to those in Experiments 1 and 2 except that during encoding the subjects were asked to rate either the speaker's clarity of enunciation or the number of meanings for each word on a 4-point numeric scale.

Results

Table 4 presents the proportion of studied and nonstudied items reported on the low pass filter identification task as a function of type of encoding task (clarity vs. meaning rating) and speaker's f_0 (same vs. changed). As indicated in Table 4, there was clear evidence of priming, the magnitude of priming was larger in the same f_0 than in the different f_0 condition following both encoding tasks, and there was no overall effect of encoding task.

A two-tailed t test that compared the overall proportion of studied and nonstudied items correctly identified on the auditory identification task indicated significant priming, $t(63) = 10.71$, $p < .001$. A 2 × 2 ANOVA was performed on the priming scores, with type of encoding task as the between-subjects variable and speaker's f_0 as the within-subject variable. There was a main effect of speaker's f_0, $F(1, 62) = 13.29$, $p < .001$, $MS_e = .017$, indicating higher identification scores when the item was presented in the same f_0 than in a different f_0 at study and test. There was a nonsignificant main effect of the type of encoding task, $F(1, 62) < 1$, indicating that priming was equivalent following both encoding tasks, and no significant interaction between type of encoding task and speaker's f_0, $F(1, 62) < 1$.

Table 5 presents the proportion of hits and false alarms reported on the recognition test as a function of encoding task and speaker's f_0. As indicated in Table 5, the meaning rating (semantic) encoding task produced greater recognition than the clarity rating (nonsemantic) encoding task, and there were somewhat higher levels of recognition in the different than in the same f_0 condition.

A 2 × 2 ANOVA was also performed on the recognition scores, with type of encoding task as the between-subjects variable and speaker's f_0 as the within-subject variable. There was a main effect of speaker's f_0, $F(1, 62) = 4.59$, $p < .05$, $MS_e = .014$, indicating, somewhat unexpectedly, higher identification scores when the item was presented in a different f_0 than in the same f_0 at study and test. There was a highly significant main effect of the type of encoding task, $F(1, 62) = 16.45$, $p < .0001$, $MS_e = .054$, indicating that recognition was much higher following the meaning rating encoding task than the clarity rating encoding task. There was a nonsignificant interaction between type of encoding task and speaker's f_0, $F(1, 62) < 1$.

Table 4
Proportion of Studied and Nonstudied Words Correctly Identified on the Filter Identification Test as a Function of Type of Encoding and Speaker's Fundamental Frequency (f_0) in Experiment 3

	Studied words			
Type of encoding	Same speaker's f_0	Different speaker's f_0	*M*	Nonstudied words
Clarity	.33	.23	.28	.17
Meaning	.33	.25	.29	.17
M	.33	.24	.28	.17

Table 5
Proportion of Hits and False Alarms on the Recognition Test as a Function of Type of Encoding and Speaker's Fundamental Frequency (f_0) in Experiment 3

	Studied words (hits)			
Type of encoding	Same speaker's f_0	Different speaker's f_0	*M*	Nonstudied words (false alarms)
Clarity	.62	.65	.64	.27
Meaning	.76	.82	.79	.26
M	.69	.74	.72	.26

To examine the relation between auditory identification and recognition performance more directly, a combined ANOVA was performed on the priming and recognition scores; type of encoding task was the between-subjects variable and type of test and speaker's f_0 were the within-subject variables. The two critical outcomes of this analysis were a significant interaction between type of encoding task and type of test, $F(1, 62) = 11.57$, $p < .01$, $MS_e = .034$, indicating that the encoding manipulation influenced recognition but not priming, and an interaction between type of test and speaker's f_0, $F(1, 62) = 14.53$, $p < .001$, $MS_e = .018$, indicating greater priming in the same than in the different f_0 condition together with an opposite effect on recognition performance.

Discussion

There were two key outcomes in Experiment 3. First, priming was entirely unaffected by the semantic versus nonsemantic encoding manipulation even though it produced large effects on recognition performance. This finding replicates earlier work using auditory identification-in-noise and stem-completion tasks (Schacter & Church, 1992). The other key outcome was that priming, in contrast to recognition memory, was significantly higher when the items were spoken in the same fundamental frequency at study and test than when the f_0 was changed. The anomalous finding that subjects performed better on the recognition test when the f_0 had changed is addressed in detail later in this discussion.

The finding that small changes in the f_0 of a single speaker's voice can produce specificity effects on an auditory priming task provides evidence against the possibility that previously observed voice-change effects can be explained by some variant of the connotative hypothesis (Geiselman & Bellezza, 1976, 1977; Geiselman & Crawley, 1983; see Palmeri et al., 1993, for converging evidence). Although it is conceivable that differences in meaning are assigned to words with slightly different f_0 values independent of changes in intonation contour or speaker's voice, the possibility seems highly unlikely. It is made even less likely by the finding that f_0 changes had just as large an effect in the nonsemantic encoding condition as in the semantic encoding condition. Instead, our data suggest that auditory priming can be based on highly specific acoustic representations of the studied items, whereas recognition performance does not appear to depend on the same sort of acoustic information.

There was, however, an apparent anomaly in the data; recognition performance was higher when the f_0 had changed between study and test than when it was the same. This finding is attributable to the effects of the prior identification task on recognition performance. Because we used a within-subject manipulation of type of test, with the recognition test always following the auditory identification task and always being composed of items with the opposite f_0s of those on the identification task, the increased identification of same f_0 items may subsequently have produced increased recognition of items in the changed f_0 condition. However, the observed pattern of results raises the possibility that a potentially significant advantage for the same versus different f_0 items on the recognition test was masked by the effect of the prior identification.

In our previous research (Schacter & Church, 1992), we failed to find significant voice-change effects on recognition in an experiment in which we manipulated type of test with a between-subjects design. To evaluate the issue with respect to f_0 changes, we tested an additional 16 subjects on the recognition test, 8 following the meaning encoding task and 8 following the clarity encoding task. All aspects of the procedure, materials, and design were identical to Experiment 3, except that the recognition test was given immediately after the filler task. Analysis of the corrected recognition scores revealed no effects of f_0 change in either the meaning condition (same f_0 = .85, changed f_0 = .87) or the clarity condition (same f_0 = .47, changed f_0 = .46). A 2 × 2 ANOVA was performed on the recognition scores, with encoding task as the between-subjects variable and speaker's f_0 as the within-subject variable. It revealed a significant main effect of encoding task, $F(1, 14) = 28.59$, $p < .001$, $MS_e = .044$; a nonsignificant main effect of speaker's f_0, $F(1, 14) < 1$; and no significant interaction, $F(1, 14) < 1$. These results clearly indicate that our use of a within-subject design in Experiment 3 did not mask significant f_0 change effects on the recognition test.

The foregoing analyses help to establish more securely that Experiment 3 provided a double dissociation between priming and explicit memory. Priming was affected by f_0 change but not by type of encoding task, whereas recognition was affected by type of encoding task but not by f_0 change. This double dissociation indicates that the types of information important to implicit and explicit retrieval are fundamentally different. Explicit recognition is enhanced by semantic–conceptual encoding of the items but does not depend on detailed acoustical information. By contrast, implicit memory, as indexed by priming, depends on detailed acoustical information but does not rely on semantic–conceptual information.

One question left unanswered by Experiment 3 concerns the generality of the observed effect of f_0 change in auditory priming. Because the low pass filter that was used for the identification task degrades higher frequency information while leaving f_0 information intact, it is possible that the effects of f_0 changes on auditory priming are restricted to the low pass filter task. That is, because f_0 was the primary information preserved in the identification task, it may have thus played an unusually large role in identification performance. Moreover, because the recognition test comprised words that were clearly presented, the apparent dissociation provided by f_0 changes may be a function of the different physical cues used in the identification and recognition tests rather than the implicit versus explicit retrieval processes that were tapped by the two tasks. Perhaps study–test f_0 changes do not significantly affect the amount of auditory priming on an implicit test other than the low pass filter identification task. Experiment 4 addressed this concern.

Experiment 4

Our primary goal in Experiment 4 was to assess the effects of f_0 changes with an implicit task in which the items are not degraded by a low pass filter and in which the explicit test is more directly comparable with the implicit task. To accomplish this goal, we used an auditory stem-completion task and cued-recall test in which the nondegraded first syllables of words served as cues on both tasks. On the (implicit) stem-completion task, subjects provided the first word that came to mind; on the (explicit) cued-recall test, subjects attempted to recollect the study list target. Previous research has found voice-change effects on an auditory stem-completion test but not on an auditory cued-recall test (Schacter & Church, 1992).

We also used a slightly different type of f_0 manipulation than in Experiment 3. The natural voice of the male speaker who recorded the stimuli in Experiment 3 was unusually low. It is possible that f_0 change effects were present because of the novelty of the extremely low voice produced when we further lowered his f_0 rather than a general sensitivity to any f_0 changes. Therefore, it seemed prudent to examine the effect of raising his f_0 by 10% as well as lowering it by 10%, as we did in Experiment 3.

Method

Subjects. Thirty-two Harvard University undergraduates participated in the experiment for a $5 payment.

Materials. Target materials comprised 48 words that were divided into two subsets of 24 words. The two subsets were matched for frequency, first letter, number of syllables, and length, and all of the words had first syllables that allowed at least three English word completions (Graf & Williams, 1987; Kucera & Francis, 1967). The words were originally recorded by a single male speaker into a Gateway 2000 IBM-compatible computer using the Kay CSL 4300 acoustic package at a sampling rate of 10 kHz. Auditory stems were created by using the computer to edit each word so that only the first syllable was preserved. We made two resynthesized duplicates of each stem using the software package described in Experiment 3. For one of the resynthesized duplicates, the f_0 of the speaker's voice was lowered by 10% and for the other it was raised by 10%. Multisyllabic words were required for this task, and it proved too difficult to change the f_0 for the entire word without degrading the sound quality of the stimuli. Therefore, only the stems were resynthesized.

Three tapes were used to record one version of each of the 24-word study lists and four versions of the auditory stem-completion task and cued-recall test. Each study list tape included 24 words spoken clearly. The auditory stem-completion and cued-recall tapes were identical; they included 48 word stems, 24 that were the first syllables of words that had been presented at study and 24 that were the first syllables of words from the nonstudied list. Of the 24 studied items, half were presented in the same (original) f_0 and half were presented in a changed (resynthesized) f_0; of the 12 items in the changed condition, half were higher and half were lower then the study list f_0. The same number of original and resynthesized f_0 stems were present among the nonstudied stems. The tapes were presented using a cassette deck and headphones. The same type of booklet used in Experiments 1 and 2 was provided for the subjects to record all of their responses.

Design and procedure. We used a 2 × 2 × 2 within-subject design. The variables were type of test (auditory stem completion vs. auditory cued recall), item type (studied vs. nonstudied), and speaker's f_0 (same vs. changed).

All of the subjects were tested individually. The subjects completed the same encoding and distractor tasks described in Experiments 1 and 2.

After the distractor task, the subjects were given the auditory stem-completion task during which they heard 48 first syllables and were asked to write down the first word beginning with that syllable that came to mind. There were 7 s between the items for the subjects to write down their answers. Once the subjects had completed the auditory stem-completion task, they were given an auditory cued-recall test. They heard 48 first syllables and were asked to complete the appropriate syllables with the words that had been presented during the encoding task. They were warned that many of the stems were distractors that had been on the stem-completion task and that it was important that they complete stems only with words they remembered specifically from the encoding task. Also, the subjects always received the version of the cued-recall test that had the items in a different presentation order and different f_0 than the stem-completion test. This procedure helped to reduce the number of intrusions from the completion test the subjects reported. There were 7 s between the items.

Results

Because the same f_0 condition used the speaker's natural voice and the changed f_0 condition used the resynthesis of the speaker's voice, careful analyses were done to ensure that any reduction of priming in the changed f_0 condition is not attributable to a degradation of the acoustic signal occurring during resynthesis that renders these stems more difficult to perceive. The mean baseline performance for the natural f_0 (.21) and resynthesized f_0 (.20) stems were virtually equivalent, indicating that this was not a problem. Also, the priming and recall scores were compared with their corresponding baselines (original vs. resynthesized f_0), and we computed all the proportions by dividing subjects' scores by the number of syllables that they correctly perceived in each condition (misperception of a stem is indicated when a subject provides a completion, e.g., *parlor,* that is phonologically incompatible with the stem, e.g., *gar*). For example, if a subject had misperceived two stems in the changed f_0 condition, his or her number of correct completions in that condition would be divided by 10 instead of by the total 12 items presented. This procedure adjusted for any differences between conditions attributable to differences in correctly perceiving the stems. Subjects misperceived an average 15% of the stems with slightly more misperceptions of the unstudied stems (17%) than of the studied stems (14%).

Table 6 presents the proportion of studied and nonstudied items reported on the auditory stem-completion task and auditory cued-recall test as a function of test type (stem completion vs. cued recall) and speaker's f_0 (same vs. changed). As indicated in Table 6, there was evidence of priming on the auditory stem-completion task. There was also evidence that priming was reduced by study–test changes in f_0 but that cued recall was not.

Table 6

Proportion of Studied and Nonstudied Target Words Produced on the Stem-Completion and Cued-Recall Tests as a Function of Speaker's Fundamental Frequency (f_0) in Experiment 4

	Studied target words			
Type of test	Same speaker's f_0	Different speaker's f_0	*M*	Nonstudied target words
Completion	.47	.38	.42	.20
Cued recall	.41	.41	.41	.06
M	.44	.39	.42	.13

A two-tailed *t* test that compared the overall proportion of studied and nonstudied items correctly identified on the auditory stem-completion task showed significant priming, $t(31) = 8.30, p < .001$. A 2 × 2 ANOVA was performed on the priming and recall scores, with type of test and speaker's f_0 as the within-subject variables. There was a main effect of test, $F(1, 31) = 17.55, p < .001, MS_e = .033$, indicating higher recall scores than priming scores, and a nonsignificant main effect of speaker's f_0, $F(1, 31) = 1.31, p > .20, MS_e = .015$. However, there was a significant interaction between type of test and speaker's f_0, $F(1, 31) = 7.43, p < .01, MS_e = .016$. Planned comparisons revealed significantly higher levels of priming in the same f_0 condition than in the different f_0 condition on the stem-completion test, $t(31) = 2.12, p < .05$, but no effect of f_0 changes on the cued-recall test, $t(23) < 1$. All the *t* tests were two-tailed.

A further analysis was also conducted to determine whether there were any differences between subjects' performance when f_0 was resynthesized to a higher versus a lower f_0. The within-subject variables were type of f_0 change (higher vs. lower) and type of test (stem completion vs. cued recall). The proportional means for stem completion were .39 for the higher f_0 and .38 for the lower f_0. The proportional means for cued recall were .41 for both the higher and lower f_0s. Neither the type of f_0 change nor the Type of f_0 Change × Type of Test interaction approached significance (both *F*s < 1). Thus, the effect of changing f_0 was robust across both directions of change.

Discussion

Experiment 4 revealed a significant reduction of priming on the auditory stem-completion task when the f_0 was changed between study and test and no effect of the same f_0 changes on cued recall. Also, there were no differences between changes produced by raising versus lowering the speaker's f_0.

The finding that small changes in the speaker's f_0 reduces priming on a stem-completion task provides evidence against the hypothesis that the f_0 change effects in Experiment 3 were attributable to the particular filter task that was used. Although the low pass filter might have contributed to the f_0

change effect by specifically preserving f_0 information, it is clear that the effect is present even when all of the acoustic information is available. Also, the finding that raising the speaker's f_0 by 10% produces the same f_0 change effect as lowering it by the same amount indicates that the previous results were not attributable to the novelty of the speaker's extremely low voice when the f_0 was lowered.

So far, all of our own and others' experiments that have shown voice-change effects on auditory priming tasks (cf. Goldinger, 1992; Schacter & Church, 1992) have included a change in the speaker's f_0. It is thus possible that f_0 changes are the critical factor underlying both the voice-change and intonation-change effects of previous experiments. However, it is also possible that the observed voice-change effects are not specific to f_0 (or even pitch) and that changes in any type of acoustic information would produce a similar specificity effect. Therefore, it is critical to manipulate an acoustic variable that does not involve change in f_0 to determine whether f_0 change is critical to voice-specific priming.

One acoustic variable that can be manipulated without any changes in the f_0 of the speaker's voice is the intensity or overall amplitude of an item. Sommers, Nygaard, and Pisoni (1992) reported that variability in the overall amplitude of degraded words did not reduce subjects' ability to correctly perceive the words when compared with words presented in a single overall amplitude. By contrast, similar experiments examining talker variability and variable speech rates did yield a reduction in performance with variability (Sommers et al., 1992). Although the auditory priming effects that we examined may not be mediated by the same processes as variability effects, the Sommers et al. findings suggested the possibility that overall amplitude information may not produce specificity effects in auditory priming. We constructed Experiment 5 to test this hypothesis.

Experiment 5

The basic design of Experiment 5 was similar to that of Experiment 4 except that it explored the effect of overall amplitude changes in the stimuli rather than changes in speaker's f_0. The words were recorded by the same male speaker, but only the speaker's natural f_0 was used. The items were presented at two different decibel levels (60 and 75 dB). The decibel levels were chosen because they allowed for a highly salient amplitude difference, 15 dB, while still falling within normal conversational volumes. During the stem-completion and subsequent cued-recall tests, the decibel level for half of the studied items was the same as at study (e.g., 60–60 and 75–75) and for the other half the decibel level was changed between study and test (e.g., 60–75 and 75–60).

Method

Subjects. Twenty-four Harvard University undergraduates participated in the experiment in exchange for $5.

Materials. Target materials comprised the same 48 words described in Experiment 4, and the same male speaker recorded the items. Auditory stems were created by using the computer to edit each word so that only the first syllable was preserved. The decibel levels were manipulated by playing the tapes through an audiometer during the experiment and adjusting the decibel level for each item.

Table 7

Proportion of Studied and Nonstudied Target Words Produced on the Stem-Completion and Cued-Recall Tests as a Function of Decibel Level in Experiment 5

	Studied target words			
Type of test	Same decibel level	Different decibel level	*M*	Nonstudied target words
Completion	.44	.44	.44	.17
Cued recall	.37	.37	.37	.08
M	.40	.40	.40	.12

Three tapes were used to record two versions of each of the 24-word study lists and two versions of the auditory stem-completion task and cued-recall test. Each study list tape included 24 words spoken clearly. Half of the words were presented at 60 dB and half were presented at 75 dB. The auditory stem-completion and cued-recall tapes were identical; they included 48 word stems, 24 that were the first syllables of words that had been presented at study and 24 that were the first syllables of words from the nonstudied list. Of the 24 studied items, half were presented at the same decibel level as study and half were presented at a changed decibel level. The experiment was counterbalanced such that each item appeared equally as often in each experimental condition. The tapes were presented using a cassette deck and headphones. The same type of booklet used in the previous experiments was provided for the subjects to record all of their responses.

Design and procedure. We used a 2 × 2 × 2 within-subject design. The variables were type of test (auditory stem completion vs. auditory cued recall), item type (studied vs. nonstudied), and decibel level (same vs. changed). The procedure was identical to those in Experiment 4.

Results

Table 7 presents the proportion of studied and nonstudied items reported on the auditory stem-completion task and the auditory cued-recall test as a function of test type (stem completion vs. cued recall) and decibel level (same vs. changed). As indicated in Table 7, there was evidence of priming on the auditory stem-completion task, and there was no effect of decibel level changes on either priming or cued recall.

A two-tailed t test that compared the overall proportion of studied and nonstudied items correctly identified on the auditory stem-completion task showed significant priming, $t(23) = 9.19, p < .001$. A 2 × 2 ANOVA was performed on the priming scores and corrected recall scores, with type of test and decibel level as the within-subject variables. There was a nonsignificant main effect of test, $F(1, 23) < 1$, and a nonsignificant main effect of decibel level, $F(1, 23) < 1$. The interaction between type of test and decibel level was also nonsignificant, $F(1, 23) < 1$, indicating that the changes in decibel level between study and test had no significant effect on the stem-completion or the cued-recall test.

General Discussion

The present experiments have yielded a number of new findings concerning voice-specific effects in auditory priming. Experiment 1 demonstrated that study–test changes in the speaker's voice reduced priming on a low pass filter identification test. Considered in relation to previous failures to find voice-change effects using identification in noise tasks (Jackson & Morton, 1984; Schacter & Church, 1992), this finding shows that voice-change effects can be found on a degraded identification task. Experiment 2 indicated a similar reduction in priming when the emotional or linguistic intonation of a single speaker's voice was changed between study and test, thereby extending the domain of voice-change effects to include manipulations within a single voice. Experiments 3 and 4 indicated that relatively small study–test changes in the f_0 of the speaker's voice reduced priming on both low pass filter identification and stem-completion tests. These findings, along with the findings of Palmeri et al. (1993), provide evidence against the connotative meaning hypothesis (e.g., Geiselman & Crawley, 1983) as a possible explanation for voice-change effects and provide evidence for the idea that pitch information plays a critical role in the observed effects. Consistent with this idea, Experiment 5 demonstrated that study–test changes in overall amplitude did not affect priming on an auditory stem-completion test.

All of the manipulations in our experiments that yielded voice-change effects included significant changes in f_0. Therefore, it is possible that voice-change effects depend exclusively on the representation of f_0. The f_0 is an extremely important component of auditory information. It is, for example, the primary acoustic correlate of vocal pitch (Lieberman, 1961). Use of f_0 contour and range are important acoustic correlates of emotional and linguistic intonation (Lieberman, 1961; Lieberman & Michaels, 1962; Williams & Stevens, 1972), and it has been shown that f_0 is important in auditory grouping and auditory stream segmentation (Bregman, Liao, & Levitan, 1990). Recent research examining patients with brain damage has indicated that f_0 information may be processed primarily by a cortical system in the right hemisphere (Robin, Tranel, & Damasio, 1990; Zatorre, 1988). There is also some evidence of a left ear advantage for processing relatively low frequencies such as f_0 in normal populations (Ivry & Lebby, 1993), though the findings with normal subjects are somewhat mixed (Wolf, 1977).

In view of these considerations, it is reasonable to hypothesize that specific representations of f_0 information are stored in an auditory PRS, perhaps involving the right hemisphere, and that voice-specificity effects are mediated by these f_0 representations. Not only is this hypothesis consistent with the results of the present experiments but also it fits with previous suggestions by Schacter and Church (1992). We had speculated that voice-specificity effects might be mediated by a right hemisphere subsystem: Neuropsychological evidence indicated that right hemisphere processing of language is especially disrupted by the presence of noise (Zaidel, 1978), and our own data indicated that voice-specificity effects are not present when the test items were degraded with white noise (see Goldinger, 1992, for further qualifications). Also, some recent dichotic listening experiments conducted in our laboratory suggest that voice-change effects are more robust when the test items are presented to the left ear than to the right ear (Schacter, Aminoff, & Church, 1992). However, these results are only preliminary, and no strong conclusions can be drawn at this time (for further discussion, see Schacter, in press; Schacter & Church, 1992).

Although the available data are consistent with the hypothesis that representations of f_0 mediate the voice-specificity effects found in auditory priming, the results of our experiments do not rule out other alternative hypotheses. It is possible that voice-specificity effects are not mediated solely by f_0 representations per se but rather by auditory representations that include information about f_0 among other acoustic attributes. Although Experiment 5 seems to indicate that overall intensity is not an important feature of the representation mediating voice-specificity effects, it might still include other features of the acoustic signal such as timbre and speech rate or even more dramatic changes in overall intensity. The effects on priming of changing such spectral and temporal features of the speech signal have not been explored, and it is possible that they would produce voice-specificity effects similar to those observed in the present experiments. We also must point out that the f_0 manipulation used in Experiments 3 and 4 changes the f_0 without changing the timbre. Therefore, it is not possible to determine whether changes in the quantitative value of f_0 are responsible for the observed effect on priming or whether the effect is produced by the associated changes in the relationship between f_0 and its harmonics.

Nevertheless, because all of the experiments that have found voice-specificity effects have included significant changes in f_0, the most parsimonious explanation is that the effect depends on f_0 representations. Further research is necessary to distinguish between this hypothesis and the various alternative possibilities.

It is also important to note that the data from both Schacter and Church (1992) and the present experiments indicate that auditory priming contains a substantial abstract component. The magnitude of the various voice-specificity effects were relatively modest, and subjects consistently showed evidence of significant priming in the changed voice conditions. We have speculated previously that there may be two PRS subsystems contributing to auditory priming, one that represents abstract phonological information and another that represents specific acoustic information (Schacter & Church, 1992). The present results are consistent with such speculations.

These suggestions are also worth considering in light of recent research that we have conducted in patients with amnesia. Schacter, Church, and Treadwell (1994) found that patients with amnesia exhibited entirely normal priming effects on an auditory identification test that used white noise to degrade target items; patients with amnesia, like controls, showed nearly identical amounts of priming in same and different voice conditions. By contrast, when we used the low pass filter identification task from the present experiments, control subjects showed significantly more priming in the same-voice condition than in the different-voice condition, whereas patients with amnesia showed nonsignificantly less priming in the same-voice condition than in the different-voice condition (Schacter, Church, & Bolton, 1994). Although these results are preliminary and require additional investigation, the failure of patients with amnesia to exhibit voice-specific priming raises the possibility that the voice-specificity effects observed in the present experiments depend in some way on the episodic system that is damaged in patients with amnesia. For instance, although phonological word forms and voice information may be represented in different PRS subsystems, as we have suggested, the episodic system might be needed to bind together, at the time of study, the target word form and a particular speaker's voice. Thus, voice-specific priming may depend, to some extent, on an interaction between PRS and episodic memory (for discussion, see Schacter, in press).

Consistent with the findings of Schacter and Church (1992), we did not find any significant voice-specificity effects on explicit recognition and recall tests. However, it is important to point out that this does not mean that specific voice information is unavailable for explicit recall. Many experiments have shown that subjects can explicitly remember information about a speaker's voice when asked to (e.g., Geiselman & Bellezza, 1976, 1977; Geiselman & Crawley, 1983) and that there are some circumstances in which variations in voice information reduce explicit recall and recognition (e.g., Craik & Kirsner, 1974; Goldinger et al., 1991; Martin et al., 1989). Although it has not been directly demonstrated, it is possible that some within-voice acoustic information, such as f_0, is also available to explicit recollection. However, our data indicate that changing these acoustic features between study and test has little effect on recognition and recall. We assume that this is because subjects tend to rely on conceptually driven processes when engaging in explicit retrieval (e.g., Roediger et al., 1989), with the consequence that acoustic information plays a minimal role as a cue for explicit retrieval. Further research that explores the kinds of acoustic information that influence priming, and the factors that govern the availability of that information for conscious recollection, is critical to understanding both implicit and explicit memory for auditory information.

References

Bassili, J. N., Smith, M. C., & MacLeod, C. M. (1989). Auditory and visual word-stem completion: Separating data-driven and conceptually driven processes. *Quarterly Journal of Experimental Psychology, 41A,* 439–453.

Blumstein, S., & Goodglass, H. (1972). The perception of stress as a semantic cue in aphasia. *Journal of Speech and Hearing, 15,* 800–806.

Bregman, A. S., Liao, C., & Levitan, R. (1990). Auditory grouping based on fundamental frequency and formant peak frequency. *Canadian Journal of Psychology, 44,* 400–413.

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

Ellis, A. (1982). Modality-specific repetition priming of auditory word recognition. *Current Psychological Research, 2,* 123–128.

Ellis, A. (1984). Introduction to Byrom Bramwell's (1887) case of word meaning deafness. *Cognitive Neuropsychology, 1,* 245–248.

Franks, J. J., Plybon, C. J., & Auble, P. M. (1982). Units of episodic memory in perceptual recognition. *Memory & Cognition, 10,* 62–68.

Geiselman, R. E., & Bellezza, F. S. (1976). Long-term memory for speaker's voice and source location. *Memory & Cognition, 4,* 483–489.

Geiselman, R. E., & Bellezza, F. S. (1977). Incidental retention of speaker's voice. *Memory & Cognition, 5,* 658–665.

Geiselman, R. E., & Crawley, J. M. (1983). Incidental processing of speaker characteristics: Voice as connotative information. *Journal of Verbal Learning and Verbal Behavior, 22,* 15–23.

Goldinger, S. D. (1992). *Words and voices: Implicit and explicit memory for spoken words.* Unpublished doctoral dissertation, Indiana University, Bloomington.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language, 28,* 501–518.

Goldinger, S. D., Pisoni, D. B., & Logan, J. S. (1991). On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 152–162.

Graf, P., & Williams, D. (1987). Completion norms for 40 three-letter word stems. *Behavior Research Methods, Instruments, & Computers, 19,* 422–445.

Heilman, K., Bowers, D., Speedie, L., & Coslett, H. (1984). Comprehension of affective and non-affective prosody. *Neurology, 34,* 917–921.

Ivry, R. B., & Lebby, P. C. (1993). Hemispheric differences in auditory perception are similar to those found in visual perception. *Psychological Science, 4,* 41–45.

Jackson, A., & Morton, J. (1984). Facilitation of auditory word recognition. *Memory & Cognition, 12,* 568–574.

Jacoby, L. L. (1983). Perceptual enhancement: Persistent effects of experience. *Memory & Cognition, 9,* 21–38.

Jacoby, L. L., Allan, L. G., Collins, J. C., & Larwill, L. K. (1988). Memory influences subjective experience: Noise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 240–247.

Jakimik, J., Cole, R. A., & Rudnicky, A. I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory and Language, 24,* 165–178.

Kempley, T., & Morton, J. (1982). The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology, 73,* 441–454.

Kirsner, K., Dunn, J. C., & Standen, P. (1989). Domain-specific resources in word recognition. In S. Lewandowsky, J. Dunn, & K. Kirsner (Eds.), *Implicit memory: Theoretical issues* (pp. 99–122). Hillsdale, NJ: Erlbaum.

Kohn, S. E., & Friedman, R. B. (1986). Word-meaning deafness: A phonological-semantic dissociation. *Cognitive Neuropsychology, 3,* 291–308.

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

Lieberman, P. (1961). Perturbations in vocal pitch. *Journal of the Acoustical Society of America, 33,* 597–603.

Lieberman, P., & Michaels, S. B. (1962). Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. *Journal of the Acoustical Society of America, 34,* 922–927.

Martin, C. S., Mullennix, J. W., Pisoni, D. B., & Summers, W. V. (1989). Effects of talker variability on recall spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 676–684.

Masson, M. E. J., & Macleod, C. M. (1992). Reenacting the route to interpretation: Context dependency in encoding and retrieval. *Journal of Experimental Psychology: General, 121,* 145–176.

McClelland, A. G. R., & Pring, L. (1991). An investigation of cross-modality effects in implicit and explicit memory. *Quarterly Journal of Experimental Psychology, 43A,* 19–33.

Palmeri, T., Goldinger, S., & Pisoni, D. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 309–328.

Pavio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness: Values for 925 nouns. *Journal of Experimental Psychology Monographs, 76*(1, Pt. 2).

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 36,* 475–543.

Robin, D. A., Tranel, D., & Damasio, H. (1990). Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain and Language, 39,* 539–555.

Roediger, H. L., III. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45,* 1043–1056.

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, S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3–41) Hillsdale, NJ: Erlbaum.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 501–518.

Schacter, D. L. (1990). Perceptual representation systems and implicit memory: Toward a resolution of the multiple memory systems debate. In A. Diamond (Ed.), *Annals of the New York Academy of Sciences: Vol. 608. Development and neural bases of higher cortical functions* (pp. 543–571). New York: New York Academy of Sciences.

Schacter, D. L. (1992). Understanding implicit memory: A cognitive neuroscience approach. *American Psychologist, 47,* 559–569.

Schacter, D. L. (in press). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994.* Cambridge, MA: MIT Press.

Schacter, D. L., Aminoff, A., & Church, B. A. (1992). [A dichotic listening study of voice-specific priming in auditory stem completion]. Unpublished raw data.

Schacter, D. L., Chiu, C. Y. P., & Ochsner, K. N. (1993). Implicit memory: A selective review. *Annual Review of Neuroscience, 16,* 159–182.

Schacter, D. L., & Church, B. A. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 915–930.

Schacter, D. L., Church, B. A., & Bolton, E. (1994). *Implicit memory in amnesic patients: Impairment of voice-specific priming.* Manuscript submitted for publication.

Schacter, D. L., Church, B. A., & Treadwell, J. (1994). Implicit memory in amnesic patients: Evidence for spared auditory priming. *Psychological Science, 5,* 20–25.

Schacter, D. L., Cooper, L. A., & Delaney, S. (1990). Implicit memory for unfamiliar objects depends on access to structural descriptions. *Journal of Experimental Psychology: General, 119,* 5–24.

Schacter, D. L., McGlynn, S. M., Milberg, W., & Church, B. A. (1993). Spared priming despite impaired comprehension: Implicit memory in a case of word meaning deafness. *Neuropsychology, 7,* 107–118.

Schacter, D. L., Rapcsak, S. Z., Rubens, A. B., Tharan, M., & Laguna, J. M. (1990). Priming effects in a letter-by-letter reader depend upon access to the word form system. *Neuropsychologia, 28,* 1079–1094.

Shapiro, B., & Danly, M. (1985). The role of the right hemisphere in the control of speech prosody in propositional and affective contexts. *Brain and Language, 25,* 19–36.

Shimamura, A. P. (1986). Priming effects in amnesia: A dissociable memory function. *Quarterly Journal of Experimental Psychology, 38A,* 619–644.

Slowiaczek, L. M., Nusbaum, H. C., & Pisoni, D. B. (1987). Phonological priming in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 64–75.

Slowiaczek, L. M., & Pisoni, D. B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory & Cognition, 14,* 230–237.

Sommers, M. S., Nygaard, L. C., Pisoni, D. B. (1992). Stimulus variability and the perception of spoken words: Effects of variations in speaking rate and overall amplitude. In J. J. Ohala, T. M. Nearey, B. L. Derwing, M. M. Hodge, & G. E. Wiebe (Eds.), *ICSLP 92 Proceedings: 1992 International Conference on Spoken Language Processing* (Vol. 1, pp. 217–220). Edmonton, Alberta, Canada: Priority Printing.

Squire, L. R. (1987). *Memory and brain.* New York: Oxford University Press.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247,* 301–306.

Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustical correlates. *Journal of the Acoustical Society of America, 52,* 1238–1250.

Wolf, C. G. (1977). The processing of fundamental frequency in a dichotic matching task. *Brain and Language, 4,* 70–77.

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. J. (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America, 84,* 566–572.

Received March 26, 1993
Revision received July 8, 1993
Accepted July 12, 1993 ■