Specificity of memory representations for spoken words

PAUL A. LUCE and EMILY A. LYONS
State University of New York, Buffalo, New York

Many theories of spoken word recognition assume that lexical items are stored in memory as abstract representations. However, recent research (e.g., Goldinger, 1996) has suggested that representations of spoken words in memory are veridical exemplars that encode specific information, such as characteristics of the talker's voice. If representations are exemplar based, effects of stimulus variation such as that arising from changes in the identity of the talker may have an effect on identification of and memory for spoken words. This prediction was examined for an implicit and explicit task (lexical decision and recognition, respectively). Comparable amounts of repetition priming in lexical decision were found for repeated words, regardless of whether the repetitions were in the same or in different voices. However, reaction times in the recognition task were faster if the repetition was in the same voice. These results suggest a role for both abstract and specific representations in models of spoken word recognition.

Theories of spoken word recognition have traditionally assumed, either implicitly (see, e.g., Luce & Pisoni, 1998; McClelland & Elman, 1986; Norris, 1994) or explicitly (see, e.g., Jackson & Morton, 1984), that lexical items are represented in memory by abstract phonological codes that only preserve information relevant for lexical discrimination. In many current models of word recognition, stimulus variation—arising from factors such as changes in speaking rate and the identity of the talker—is treated as irrelevant information that is discarded early in the encoding process. The extraction of information that is solely relevant for identification is referred to as *normalization*, and it is during the normalization phase that representations of stimuli that vary in physical detail but fall within a given perceptual category are equated by processes that abstract defining information.

For example, feature-based accounts of speech perception (see Klatt, 1989; Pisoni & Luce, 1987) have proposed that speech sounds and words are processed using the elemental features of linguistic description (e.g., [vocalic], [consonantal], [sonorant]). However, spoken words may differ on many physical dimensions not captured by these features. The normalization process is responsible for winnowing the information in the speech signal and extracting only the featural information that is relevant for identification. This process thereby serves a sub-

stantial data reduction function that may ultimately result in considerable economy of process and representation.

It should be noted that this theoretical stance on the nature of perceptually based representations is not unique to research on language perception. Some theories of visual object recognition (see, e.g., Biederman, 1987) assume that most objects can be uniquely defined by a small number of abstract units. In fact, Biederman has argued that the rapidity of visual object identification necessitates the use of a small number of easily extractable units.

Despite the arguments that have been made for abstract lexical representations in memory, recent research (see Goldinger, 1996, for a review) has suggested that putatively irrelevant *surface* details of words—such as information specific to a given talker—are preserved in some form in memory. These findings regarding specificity effects have led to the proposal (see, e.g., Goldinger, 1996) that lexical items are represented in memory by exemplar-based representations that preserve, rather than discard, much of the physical detail of the stimulus.

Research has demonstrated that variation in the surface details of spoken stimuli (usually measured by changes in the identity of the talker, hereafter referred to broadly as changes in *voice*) has implications for both identification and memory. Typically, subjects have more difficulty in identifying (Mullennix, Pisoni, & Martin, 1989), recognizing (Church & Schacter, 1994; Goldinger, 1996; Palmeri, Goldinger, & Pisoni, 1993; Schacter & Church, 1992; Sheffert, 1995, 1998, in press), and recalling (Goldinger, Pisoni, & Logan, 1991; Martin, Mullennix, Pisoni, & Summers, 1989) lists of stimuli composed of words spoken by multiple talkers, as compared with lists composed of stimuli spoken by a single talker (see Palmeri et al., 1993, for one interesting exception). One explanation for these effects is that normalization pro-

708

cesses reduce the resources available for encoding and/or rehearsal.

The effects of changes in the surface details of stimuli between study and test in recognition memory experiments have been of particular interest in the literature. For example, Church and Schacter (1994) and Schacter and Church (1992) investigated the effects of talker variation on implicit and explicit memory.[1] They observed effects of talker variation in *implicit* tasks, such as fragment completion and the identification of low-pass filtered stimuli. Subjects were more likely to complete a fragment of a word if the fragment was repeated in the same voice. Subjects were also more accurate at identifying low-pass filtered words that were repetitions of previously presented items if the repetition preserved surface characteristics of the stimulus. However, these researchers failed to observe effects of stimulus specificity in *explicit* tasks. When subjects performed cued recall or recognition of previously presented items, changing surface characteristics between study and test had no statistically significant effects on performance.

Goldinger (1996) performed a series of experiments in which the effects of voice on memory for spoken words were examined. In one of his experiments, Goldinger presented words in explicit (recognition) and implicit (perceptual identification in noise) tasks with varying delays between study and test. He found significant effects of voice in both recognition and identification, demonstrating that voice effects are not, in fact, restricted to implicit tasks. However, Goldinger (1996) found that the effects of voice were reduced more by delay between study and test in the explicit task than in the implicit task, a result that provides some support for the explicit/implicit distinction. In another experiment, Goldinger manipulated levels of processing and voice in the study–test implicit– explicit format. His results demonstrated that the effects of voice varied with level of processing, such that the strongest effects of stimulus specificity were observed in the shallower processing conditions, especially for recognition memory.

Although somewhat varied, the overall results of studies in which the effects of voice on identification and memory were examined are consistent with exemplar-based models of memory (see, e.g., Hintzman, 1986). According to these models, a new representation of a stimulus item is stored in memory each time it is encountered, and it is hypothesized that these representations preserve surface information about the stimulus. One advantage of exemplar-based models is that they have the potential for solving the long-standing problem of perceptual normalization in speech perception by dispelling the notion that the ultimate goal of the perceptual process is to map acoustic–phonetic information onto *abstract* form-based representations of words in memory. In exemplar-based models, the representational currency of the perceptual encoding process is more or less true to the details of the stimulus itself. In an application of this general theoretical approach to spoken word recognition, Goldinger

(1996, 1998) has proposed an exemplar-based lexicon in which the individual memory traces themselves may encode both abstract and surface information, with the degree of stimulus specificity depending crucially on attentional factors during encoding.

To date, a variety of researchers have found evidence for the existence of specificity effects in tasks involving memory for spoken words. However, the implications of these results for models of word identification are as yet unclear. Our research was therefore aimed at further investigating the effects of stimulus-specific information— in our case, voice—on spoken word identification or, more specifically, on on-line processing assessed by the auditory lexical decision task. Although previous results have been at times conflicting, there appears to be some consensus that surface details of words are stored in some form in the memory. The primary impetus for the present investigation was to attempt to establish more clearly the domain in which surface characteristics of nondegraded spoken stimuli play a role in the identification process itself. Specifically, we were interested in determining whether voice information about spoken words is always implicated in the identification process or whether there are circumstances in which voice information is less relevant, such as in immediate, on-line processing. If we are able to identify situations in which voice information appears to play a minor or inconsequential role in the word identification process, this may serve as evidence that existing abstractionist models of spoken word identification are not necessarily incorrect in restricting their focus to processes that are independent of specific voice effects.

We examined the effects of changing voice on stimulus repetition in both auditory lexical decision and recognition. Our research was based on previous work by Biederman and Cooper (1992) and by Cooper, Schacter, Ballesteros, and Moore (1992), who examined the effects of stimulus variation in implicit and explicit tasks. We first presented subjects with a list of stimuli spoken by two talkers in a lexical decision task. We then followed this first block of lexical decision trials with either (1) another block of lexical decision trials or (2) a block of recognition (old/new) trials. The stimuli in the second block of the experiment were repeated in the same voice or in a new voice or were new items that had not appeared in the first block. We used the lexical decision task in the second block to evaluate the role of specific information in on-line word identification when explicit recall or recognition of a prior item was not required. For purposes of comparison, we also used a recognition task in the second block (with a different group of subjects) in order to further examine the role of stimulus-specific information in a task requiring explicit recollection of previously presented items.

We hypothesized that, if memory representations of spoken words are abstract, repetition effects for words repeated in the same voice should be equivalent to repe-

titions of words repeated in a different voice. That is, changing the voice of the speaker for an item between blocks should have no effect on performance in the second block. On the other hand, if the representations of spoken words used in on-line processing are exemplar based, we hypothesized that words repeated in the same voice would show greater advantages in terms of processing speed and accuracy than would words repeated in a different voice.

We also attempted to extend previous research by using nondegraded stimuli and measuring response latencies (as opposed to accuracy, the typical dependent measure in previous studies; see, however, Goldinger, 1996). Earlier demonstrations of sensitivity to stimulus specificity may have been amplified by the use of degraded stimuli (in the form of noise-masked, low-pass filtered, or gated stimuli), which may have encouraged subjects to interrogate memory representations of previously encountered items, when they might not have normally done so, because of difficulties in processing imposed by degradation.

To summarize, modulation of the magnitude of repetition effects as a function of changes in voice in the explicit recognition task would be further evidence that stimulus-specific information is available in memory. And similar effects for the implicit lexical decision task would support the claim that representations of spoken words are exemplar based. However, diminished effects of voice on the magnitude of the repetition effect in lexical decision may indicate that more abstract representations of spoken words that subserve on-line spoken word identification co-exist with stimulus-specific exemplars.

EXPERIMENT 1

Method

Subjects. Sixty members of the State University of New York at Buffalo community participated in the experiment. Subjects either received credit for an introductory psychology class or were paid $4 for their participation. All the subjects were native English speakers and reported no history of speech or hearing disorders.

Materials. The stimuli consisted of monosyllabic consonant–vowel–consonant words and nonwords, each recorded by a male and a female talker. Nonword stimuli were created by changing one phoneme from an actual word. All nonwords were phonotactically legal. The complete list of stimuli is given in the Appendix.

The stimuli were presented in two successive blocks consisting of 60 stimuli each. The first block consisted of 15 words and 15 nonwords spoken by the male talker and 15 words and 15 nonwords spoken by the female talker. No stimuli were repeated within a block. The second block also contained 60 stimuli and consisted of (1) 10 words and 10 nonwords that were identical repetitions of stimuli heard in the first block (the *same* voice condition), (2) 10 words and 10 nonwords that were repetitions from the first block but spoken by the other talker (the *different* voice condition), and (3) 10 words and 10 nonwords that had not been presented in the first block (the *new* condition).

Three separate stimulus sets were created to ensure that every stimulus item participated in each condition (*same*, *different*, and *new*) in order to control for the possibility of unintended differences among items. Results were averaged over the three stimulus sets.

Recording took place in a sound-attenuated room, and all words were spoken in list format. The stimuli were low-pass filtered at 4.8 kHz and digitized at a sampling rate of 10 kHz with a 12-bit analog-to-digital converter. All words were edited into individual files and stored on a computer disk.

Procedure. The experiment consisted of two blocks of 60 trials each. The subjects performed an auditory lexical decision task in both blocks, in which they were presented with a stimulus and had to decide if the item was a word or a nonword. At the beginning of the experiment, the subjects were *not* instructed that they would be presented with two blocks of trials. Twenty subjects were randomly assigned to each of the three different stimulus sets.

The subjects were tested in groups of 5 or fewer. The stimuli were presented over headphones at a comfortable listening level. The subjects responded by pressing the appropriate button on a response box in front of them. Word responses were made with the right hand by pressing a button on the right labeled *Word*. Nonword responses were made with the left hand by pressing a button on the left labeled *Nonword*. All the subjects were right-handed (defined as the hand with which the subject wrote). Word responses were thus made with the dominant hand.

A PDP-11/34 computer controlled stimulus presentation and response collection. Stimulus presentation within each block was randomized by group. The subjects were instructed to respond as quickly and as accurately as possible. Prior to the first block of experimental trials, the subjects were given a block of 10 practice trials to familiarize them with the task.

A typical trial proceeded as follows: A cue light at the top of the response box was illuminated for 500 msec, indicating that a stimulus was about to be presented. A stimulus was presented 250 msec after the offset of the cuelight. Reaction times were measured from the onset of the stimulus to the onset of the response. Once the subject responded, an intertrial interval of 500 msec was initiated. If the subject failed to respond within 2,200 msec, an incorrect response was recorded, and the next trial began.

Results

Mean reaction times and percentages correct for the *same*, *different*, and *new* items in Blocks 1 and 2 are shown in Table 1. Recall that Block 1 refers to the first presentation of a stimulus and Block 2 to the second presentation (if, in fact, the stimulus was repeated). Although all items in Block 1 were new, the label *new* in reference to Block 1 refers to filler stimuli that were not repeated in Block 2. The nonword stimuli were included for purposes of the task only and were not responded to with the dominant hand.[2] Thus, only correct responses to the word stimuli will be reported.

We performed 2 (block: first vs. second) $\times$ 3 (repetition: *same* vs. *different* vs. *new*) repeated measures analyses of variance (ANOVAs) on reaction times and percentages correct for both subjects (F_s) and items (F_i). We adopted a .05 level of significance, unless otherwise noted.

Table 1
Mean Reaction Times (RT) (in Milliseconds) and Percentages Correct (PC) for All Conditions and Both Blocks for Experiment 1

Block	Same		Different		New	
	RT	PC	RT	PC	RT	PC
1	859	96	852	97	853	94
2	799	97	812	98	879	95

Reaction times were 60 msec faster in the second block than in the first block for words repeated in the *same* voice, 40 msec faster in the second block than in the first block for words repeated in the *different* voice, and 26 msec slower in the first block than in the second block for new words. This data pattern resulted in a significant block × repetition interaction [$F_s(2,118) = 22.69$, $MS_e = 2,733.12$; $F_i(2,58) = 7.239$, $MS_e = 4,411.02$].

Comparison of the differences in reaction times between Blocks 1 and 2 for the *same* and *different* conditions combined (-50 msec) and the *new* condition ($+26$ msec) revealed a significant effect of repetition [$F_s(1,118) = 42.82$; $F_i(1,58) = 13.85$]. However, when analyzed separately, there was no difference in the magnitude of the repetition effect for the *same* (60 msec) and *different* (40 msec) conditions [$F_s(1,118) = 2.57$; $F_i < 1$; 95% CI: $-6.73 < 20 < 46.73$].[3] In addition, there was no significant difference between *same* (799 msec) and *different* (812 msec) reaction times in block 2 [$F_s(1,59) = 2.18$; $F_i < 1.0$; 95% CI: $-5.89 < 13 < 31.89$].

There was the tendency for percent correct to increase from Block 1 to Block 2, although this was only significant by subjects [$F_s(2,118) = 8.79$, $MS_e = .003$; $F_i(2,58) = 1.92$, $MS_e = .006$].

Discussion

The results of Experiment 1 revealed that the subjects responded significantly more quickly to repeated words than to new words, a replication of the now well-documented repetition effect (D. L. Scarborough, Cortese, & H. S. Scarborough, 1977). Of particular interest, however, was the finding that changing the voice in which the stimulus was produced between first presentation and repetition did *not* have statistically reliable effects on the magnitude of the repetition effect (see also Brown, Fowler, & Rueckl, 1996). That is, same and different repetitions produced roughly equivalent amounts of repetition priming, although there was a small but nonsignificant trend in favor of repetitions occurring in the same voice, as compared with those that changed voices. Thus, although there may be small, albeit statistically undetectable, effects of voice in the present experiment, it appears that changing the surface details of the stimulus in this particular paradigm does not induce a marked diminution of the repetition effect. It should be noted, however, that the absence of a significant effect of voice on the *magnitude* of the priming effect was accompanied by a robust effect of repetition for both the same and different voice conditions, demonstrating that even substantial surface variation between prime and target did not eliminate repetition priming in this experiment.

These results support a number of hypotheses regarding the role of stimulus specificity in the on-line recognition of spoken words. First, our results may indicate that abstract lexical representations are used in spoken word identification. This hypothesis is consistent with a theory of memory that incorporates multiple memory systems consisting of representations of various degrees of abstraction (i.e., Schacter, 1987; Tulving, 1984). The postulation of at least a functional abstract representation is also consistent with strictly exemplar-based memory models. For example, in his MINERVA-2 model, Hintzman (1986) argues that abstract information may emerge through the activation of multiple similar traces in memory and that specific information may be more prone to forgetting than information regarding the central tendency of a category. Thus, the lag between repetitions in the present experiment may have been sufficiently long for specific stimulus information to decay, leaving more abstract, albeit still exemplar-based, information to dominate the repetition effect, as evidenced by the robust effects of repetition even for the different voice condition (see Goldinger, 1996, for results pertaining to effects of delay on implicit memory).

Our results may also be interpreted in terms of Goldinger's (1996) conception of an exemplar-based lexicon in which representations are stored as more or less specific or abstract, depending on the nature of the encoding task. Tasks that emphasize specific over abstract information at encoding are predicted to result in more specific representations in memory, and vice versa. Thus, the lexical decision task in the first block of the experiment may have encouraged storage of the items in a more abstract format, thus producing small and nonsignificant effects of voice changes on repetition.

Finally, the results from Experiment 1 may simply indicate that voice information is relatively slow in having its effects on the recognition process. The speed with which lexical identification is accomplished may be so rapid that, under normal circumstances (e.g., when the stimulus is not purposefully made difficult to perceive), lexical discrimination and recognition are accomplished before the effects of stimulus specificity have an opportunity to exert their influences on the perceptual process. If this is the case, traditional models of spoken word recognition that focus on the abstract nature of form-based lexical representation may be fundamentally correct, although perhaps incomplete in their characterization of the recognition process in situations in which recognition is not as immediate as it is under ideal listening conditions.

Before considering these theoretical accounts in more detail, however, it is incumbent on us to demonstrate that we are able—using our existing stimuli and only minor modifications of our experimental paradigm—to obtain effects of changing voice in a task previously shown to be sensitive to the surface characteristics of stimuli. To this end, we reran Experiment 1, using exactly the same stimuli and the same task (lexical decision) in the first block. The only change that we implemented was to substitute the lexical decision task in the second block with a recognition (old/new) task (see Palmeri et al., 1993, for a demonstration of effects of voice in a recognition task).

As was previously mentioned, prior research on visual object recognition (Biederman & Cooper, 1992; Cooper et al., 1992) has demonstrated that changes in the surface physical characteristics of stimuli fail to show effects in implicit tasks, such as that employed in Experiment 1, but have demonstrable effects in explicit tasks, such as recognition. Although Church and Schacter (1994; Schacter & Church, 1992) found an opposite pattern of results for spoken stimuli—with implicit tasks showing effects of changes in voice and explicit tasks showing no effects— we believed that there were sufficient differences in our study to warrant comparison of the lexical decision task with performance in the recognition memory paradigm.

We thus performed a second experiment that differed from Experiment 1 only in the task employed in the second block. As in Experiment 1, the subjects were not instructed that a second block of trials would be presented. In Experiment 2, after the first block of trials, the subjects were told that another set of stimuli would be presented and that they were to ignore possible changes in voice and to decide as quickly and as accurately as possible whether each word did or did not occur in the previous block by pressing buttons labeled either *old* or *new*.

Experiment 2 was performed in order to examine two issues. First, we were interested in determining whether changes in voice have an effect in a task that explicitly requires subjects to access information about previously occurring items. Second, we were interested in attempting to determine whether the reason for our failure to obtain voice-specific effects in lexical decision was because voice information was simply no longer available when the stimuli were presented in the second block.

EXPERIMENT 2

Method

Subjects. Sixty members of the State University of New York at Buffalo community participated in the experiment. Subjects either received credit for an introductory psychology class or were paid $4 for their participation. All the subjects were native English speakers and reported no history of speech or hearing disorders. None of the subjects in Experiment 2 had participated in Experiment 1.

Materials. The stimuli were the same as those used in Experiment 1.

Procedure. The procedure was identical to that in Experiment 1, except that, prior to the second block of trials, the subjects were instructed to respond to each stimulus as quickly but as accurately as possible by pressing a button labeled *new* if the stimulus had not occurred in the first block and *old* if the stimulus was a repetition. The subjects were *not* instructed prior to block 1 that there would be a recognition test. The subjects were also told to ignore variations in voice in making their recognition judgments.

Table 2
Mean Reaction Times (RT) (in Milliseconds) and Percentages Correct (PC) for Both Conditions for Block 2 for Experiment 2

Condition	RT	PC
Same	1,098	74
Different	1,167	70

Results

The results for Experiment 2 are shown in Table 2. We performed one-way (*same* vs. *different*) repeated measures ANOVAs on both subjects and items. In this experiment, we were only interested in the second block of trials, in which a recognition response was required. Furthermore, we were only interested in *old* responses to word stimuli. *New* responses were made with the nondominant hand and were thus not comparable to *old* responses either in handedness of the response or in the nature of the decision. Only reaction times to correct word responses were analyzed. Although nonwords were included in Experiment 2 in order to ensure that the stimulus context was identical with that of Experiment 1, responses to nonwords were again not of primary interest and will thus not be reported.[4]

Words presented in the *same* voice were responded to significantly more quickly than were words presented in the *different* voice [$F_s(1,59) = 13.93$, $MS_e = 10,606.51$; $F_i(1,29) = 7.88$, $MS_e = 10,570.38$]. There was no significant difference in accuracy between *same* and *different* voice [$F_s(1,59) = 1.59$, $MS_e = 0.023$; $F_i(1,29) = 1.88$, $MS_e = 0.010$].

In order to perform a direct statistical comparison of Experiments 1 and 2, we computed difference scores within each experiment for each subject. For Experiment 1, we subtracted mean reaction times and percentages correct for *same* voice stimuli from those for *different* voice stimuli for each subject for the second block of trials. We computed these same difference scores for the recognition data of Experiment 2. We then entered these difference scores into ANOVAs by subjects and items. There was a main effect of task (lexical decision vs. recognition) for the difference scores for reaction times [$F_s(1,118) = 7.63$, $MS_e = 12,361.23$; $F_i(1,29) = 23.57$, $MS_e = 11,500.12$], but no effect for accuracy [$F_s(1,118) = 1.69$, $MS_e = .026$; $F_i(1,29) = 3.91$, $MS_e = 8,844.30$]. Thus, the differences in reaction times for *same* and *different* stimuli were significantly greater in Experiment 2 (recognition) than in Experiment 1 (lexical decision).

Discussion

The results of the recognition experiment show clear effects of voice: Words repeated in the same voice were responded to more quickly than were words repeated in the different voice. However, there was no effect of voice on percent correct recognition. Thus, whereas these results initially seem at odds with Schacter and Church (1992) and Church and Schacter (1994), they actually replicate their findings of no effects of voice on accuracy. However, the response time measure used in Experiment 2 apparently proved to be more sensitive to effects of voice than did accuracy alone, revealing significant differences between the same and different voice conditions. It should also be noted that the longer reaction times in Experiment 2 are indicative of a more difficult task than that employed in Experiment 1. Task difficulty,

in slowing reaction times to make a recognition response, may be partially responsible for the emergence of a voice effect in this experiment.

Our results are in keeping with those of Palmeri et al. (1993), who also demonstrated effects of voice on recognition. Moreover, our results demonstrate once again that the surface characteristics of spoken stimuli have demonstrable effects on *explicit* memory (see Goldinger, 1996). Thus, the results from Experiment 2 more closely resemble the findings of Biederman and Cooper (1992) in demonstrating that the surface characteristics of stimuli are stored in memory and may play a role in the ease of old/new recognition.

Finally, the results of Experiment 2 appear to disconfirm the hypothesis that the failure to observe voice effects in Experiment 1 was due to the forgetting of specific information. The results of the present experiment clearly demonstrate that such information is, in fact, available.

GENERAL DISCUSSION

In two experiments, we examined the effects of changing voice on the processing of spoken words. In Experiment 1, we demonstrated that repetition priming for spoken words may not always be sensitive to changes in the surface characteristics of the stimuli. When the subjects were required simply to make lexical decisions to spoken words in two blocks of trials, responses times to repetitions in the same voice were not statistically different from response times to repetitions in the different voice, although overall effects of repetition priming were robust. Experiment 2 was identical to Experiment 1, except that we required that the subjects perform a surprise recognition task in the second block of trials (instead of lexical decision). In this experiment, we obtained significant effects of voice: The subjects responded *old* more quickly to words repeated in the same voice than to words repeated in the different voice. Thus, we observed diminished effects of voice on response times and accuracy in the implicit task using lexical decision but a large effect on recognition times in the explicit task.

The results from Experiment 2 are consistent with a number of previous demonstrations that voice matters in recognition (see, e.g., Goldinger, 1996; Palmeri et al., 1993). However, our results appear to conflict with those of Church and Schacter (1994) and Schacter and Church (1992), who failed to find effects of stimulus specificity on recognition performance. This conflict may be more apparent than real: We, too, failed to observe effects of voice on accuracy, the dependent measure reported by Church and Schacter (1994) and Schacter and Church (1992). Robust voice effects emerged only for our processing time data, which, as previously mentioned, may provide a more sensitive measure of voice effects in recognition of words presented in the clear.

The results from Experiment 1 are consistent with previous work on visual object identification by Biederman and Cooper (1992), who also failed to find effects of stim-

ulus specificity in implicit tasks. However, our lexical decision results are again in contrast to the findings of Church and Schacter (1994; Schacter & Church, 1992), who observed effects of stimulus specificity in certain implicit tasks. We believe that there are two crucial differences between our study and those of Church and Schacter: First, in Church and Schacter's work, when voice effects were observed, they involved degraded (in particular, filtered) stimuli, which may have introduced processing difficulties that may have amplified the effects of voice, either by slowing processing or by encouraging the activation of specific previous memory traces to aid in identification. Second, responses in the lexical decision task may be so rapid as to precede what may be the slower acting effects of stimulus specificity in processing. Despite the rapidity of the response, however, we nevertheless obtained significant effects of repetition priming, suggesting that, although specific effects of voice failed to strongly develop in the time window within which responses were made, effects of activation of lexical form-based representations were in evidence.

Our results also appear to contrast somewhat with Goldinger's (1996) predictions that encoding during study (i.e., the first block) should determine the presence or absence of specificity effects. In our experiments, in which we held the encoding task constant (lexical decision in both Experiments 1 and 2), we observed different patterns of results for the implicit and explicit tasks. Again, the differences between our results and those of Goldinger may be due to the rapidity of the identification response in lexical decision (see Hintzman & Caulton, 1997; Hintzman & Curran, 1997). Goldinger (1996) reports one of the only other studies that has examined response latencies in an implicit task in which voice was manipulated (Experiment 3). Response latencies to classify stimuli in his fastest condition were almost 100 msec longer than the latencies we observed for words in the same and different conditions in the second block. Thus, it may be that, if subjects are capable of making an identification decision quickly enough, the effects of stimulus specificity are small. Conversely, when responses are slower, as in our recognition experiment or in Goldinger's classification studies, effects of voice emerge.

We believe that our results are in keeping with the growing literature on stimulus specificity effects. In particular, the results from our recognition experiment support the notion that surface details are preserved in some form in memory and have demonstrable effects on performance. However, we have demonstrated that, under certain task conditions, stimulus specificity (at least as measured here, as changes in the identity of the talker) may have negligible effects on on-line processing.

This latter finding is consistent with, although certainly not identical to, claims made by Hintzman (1986). In Hintzman's model, abstract information is generated by mechanisms that make use of activation of multiple instances in memory. Each stimulus activates an aggregate of memory traces, referred to as an *echo*. Because

the echo is composed of a weighted average of multiple traces, the influence of features that *differ* among the stored exemplars will be minimized, whereas features that are *shared* will be strengthened. With the decay of individual features over time, information regarding the central tendency of a category will tend to be more robust. Such a memory system will give the appearance of abstraction, although the representations from which this abstraction emerge are themselves quite specific. Thus, even within strictly exemplar-based models of memory, both abstract and specific information may coexist (at least functionally). Note that our previous hypothesis that the decay of information over time accounts for the lack of voice effects in lexical decision appears to have been disconfirmed by the results of Experiment 2. However, this finding does not refute the claim that abstract information may still emerge from a strictly exemplar-based memory system.

In short, our results suggest that existing models of spoken word identification may be correct in their focus on the immediate processing of spoken words as abstract entities. Nonetheless, our results join a growing body of literature suggesting that stimulus-specific information is indeed preserved in memory and may play a role in off-line processing or in less than ideal listening situations in which the identification of spoken words is difficult.

REFERENCES

BIEDERMAN, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, **94**, 115-147.

BIEDERMAN, I., & COOPER, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 122-133.

BROWN, J. M., FOWLER, C. A., & RUECKL, J. G. (1996, June). *Voice effects in implicit memory tasks*. Poster presented at the 131st Meeting of the Acoustical Society of America, Indianapolis.

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.

COOPER, L. A., SCHACTER, D. L., BALLESTEROS, S. , & MOORE, C. (1992). Priming and recognition of transformed three-dimensional objects: Effects of size and reflection. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 43-57.

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.

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, & Cognition*, **17**, 152-162.

HINTZMAN, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, **93**, 411-428.

HINTZMAN, D. L., & CAULTON, D. A. (1997). Recognition memory and modality judgments: A comparison of retrieval dynamics. *Journal of Memory & Language*, **37**, 1-23.

HINTZMAN, D. L., & CURRAN, T. (1997). Comparing retrieval dynamics in recognition memory and lexical decision. *Journal of Experimental Psychology: General*, **126**, 228.

JACKSON, A., & MORTON, J. (1984). Facilitation of auditory recognition. *Memory & Cognition*, **12**, 568-574.

KLATT, D. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 169-226). Cambridge, MA: MIT Press, Bradford Books.

LUCE, P. A., & PISONI, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, **19**, 1-36.

MARTIN, C. S., MULLENNIX, J. W., PISONI, D. B., & SUMMERS, W. (1989). Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 676-684.

MCCLELLAND, J. L., & ELMAN, J. L. (1986). The Trace model of speech perception. *Cognitive Psychology*, **18**, 1-86.

MULLENNIX, J. W., PISONI, D. B., & MARTIN, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, **85**, 365-378.

NORRIS, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, **52**, 189-234.

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.

PISONI, D. B., & LUCE, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition*, **25**, 21-52.

SCARBOROUGH, D. L., CORTESE, C., & SCARBOROUGH, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 1-17.

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

SCHACTER, D. L., & CHURCH, B. A. (1992). Auditory priming: Implicit and explicit memory for words and voice. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 915-930.

SHEFFERT, S. M. (1995). *Implicit and explicit memory for words and voices*. Unpublished doctoral dissertation, University of Connecticut, Storrs.

SHEFFERT, S. M. (1998). Voice-specificity effects on auditory word priming. *Memory & Cognition*, **26**, 591-598.

SHEFFERT, S. M. (in press). Contributions of surface and conceptual information on spoken word and voice recognition. *Perception & Psychophysics*.

TULVING, E. (1984). How many memory systems are there? *American Psychologist*, **40**, 385-398.

NOTES

1. Schacter (1987) defines implicit memory as improved performance on a task after previous experience, when the task itself does not depend on conscious recall or recollection of the previous experience. In contrast, explicit memory is defined as performance of a task requiring conscious recollection of a previous event.

2. Analyses of the nonword reaction times and accuracy scores revealed only one effect that was significant by both subjects and items: Reaction times in the first block were significantly faster than reaction times in the second block.

3. Confidence intervals of 95% are provided for the crucial null effects as an index of the power of the statistical test.

4. Analyses of nonword reaction times and accuracy scores revealed no effects that were significant by both subjects and items.

APPENDIX

Words	Nonwords
not	/sɛp/
bed	/lɛs/
big	/dɛs/
good	/mɪm/
hot	/lɛl/
live	/bæl/
serve	/pɛs/
tell	/pɪm/
sure	/sɪɡ/
year	/vɪn/
south	/ɡæn/
lot	/ɡɪs/
more	/wɪs/
leave	/nɪs/
type	/vɪs/
have	/sɪz/
run	/pɛl/
give	/sɛm/
game	/hɪn/
full	/lɪn/
date	/sɑn/
care	/sɛk/
use	/mæl/
put	/mɑn/
five	/sɛs/
take	/kæk/
far	/sɪm/
rate	/lɪs/
head	/sɛl/
food	/dɪt/
came	/tɪs/
size	/mɪn/
feel	/sɑl/
peace	/fɪs/
cause	/sæn/
long	/kɑs/
bad	/bɪs/
rise	/lɪs/
hole	/kæs/
seen	/lɪn/