

The marriage of perception and memory: Creating two-way illusions with words and voices

STEPHEN D. GOLDINGER, HEATHER M. KLEIDER, and ERIC SHELLEY
Arizona State University, Tempe, Arizona

By most theories of lexical access, idiosyncratic aspects of speech (such as voice details) are considered noise and are filtered in perception. However, episodic theories suggest that perceptual details are stored in memory and mediate later perception. By this view, perception and memory are intimately linked. The present investigation tested this hypothesis by creating symmetric illusions, using words and voices. In two experiments, listeners gave reduced noise estimates to previously heard words, but only when the original voices were preserved. Conversely, in two recognition memory experiments, listeners gave increased *old* responses to words (or voices) presented in relatively soft background noise. The data suggest that memory can be mistaken for perceptual fluency, and perceptual fluency can be mistaken for memory. The data also underscore the role of detailed episodes in lexical access.

By its most basic definition, speech is a *medium* of communication—a carrier signal for the words, sentences, and ideas that constitute conversation. In theories of perception, the distinction between perceptual objects and their media was forcefully applied by Gibson (1966). When an observer gazes upon an object, it is perceived via reflected light that is uniquely structured by physical characteristics, such as edges and contours. The observer does not perceive the *light*; it is merely an informational medium. In his ecological theory, Gibson suggested that perceptual systems evolved to efficiently use media to acquaint observers with the environment. In this regard, linguistic theories resemble Gibson's view: Speech signals are not considered to be true perceptual objects. Instead, the important objects in speech are phonemes, syllables, words, or phrases (depending on context; McNeill & Lindig, 1973). Although few people would dispute this self-evident description of communication, the speech medium is also, in itself, a potential object of perception. Instead of focusing on a speaker's message, listeners may primarily attend to tone of voice, dialect, and so forth. Thus, spoken words lead "double lives," serving both as perceptual objects (with unique voice characteristics) and as gateways to linguistic representations.

As it happens, numerous investigations have focused on the perceptual domain, primarily through the study of *surface memory* for printed and spoken words (see reviews by Goldinger, 1996; Tenpenny, 1995). Although the literature contains occasional null results (Brown & Carr, 1993; Jack-

son & Morton, 1984; Krulee, Tondo, & Wightman, 1983), an impressive collection of positive findings now exists. With respect to printed words, font information is often (but not always) preserved, whether tested with direct or indirect memory measures (Graf & Ryan, 1990; Hintzman, Block, & Inskeep, 1972; Jacoby & Hayman, 1987; Kirsner, 1973; Manso de Zuniga, Humphreys, & Evett, 1991; Roediger & Blaxton 1987; Whittlesea & Brooks, 1988).

Similar effects arise with spoken words: Voices are reliably stored in long-term memory as a side effect of lexical access, affecting both direct and indirect memory tests (Church & Schacter, 1994; Craik & Kirsner, 1974; Geiselman & Bellezza, 1977; Goldinger, 1996; Goldinger, Pisoni, & Logan, 1991; Palmeri, Goldinger, & Pisoni, 1993; Schacter & Church, 1992). Detailed episodic traces are apparently created in spoken word perception, influencing later perceptual and memorial tests. Therefore, some authors have suggested that the mental lexicon may consist of episodes rather than abstract units (Goldinger, 1996, 1997; Jacoby, 1983a, 1983b). Beyond parsimony, this idea finds empirical support. For example, Goldinger (1998) recently found that single-word shadowing data can be predicted by a pure exemplar model (MINERVA 2; Hintzman, 1986). As another example, Roediger and Blaxton (1987) found larger repetition effects in word fragment completion for fragments studied and later tested in a constant font, relative to a changed font.

However, with respect to the hypothesis of an episodic lexicon (and its supporting data), a difficulty arises. Specifically, very few experimental paradigms can be convincingly classified as either *perceptual* or *memorial* tasks. This difficulty also arises in theory: Virtually all models of perception require access to long-term memory, and all memory judgments require perceptual targets. Thus, a decision that data are relevant to perception as opposed to memory is largely arbitrary. This ambiguity between perceptual and memorial research was noted by Roediger (1996):

This research was supported by Grant R29-DC02629-03 from the National Institute for Deafness and Other Communicative Disorders. We thank Kara Dolan for helping conduct Experiments 3A, 3B, and 4. We also thank Michael Masson, Steve Lindsay, Barbara Church, Tamiko Azuma, and Elliot Hirshman for providing helpful feedback on an earlier manuscript. Please address correspondence to S. D. Goldinger, Department of Psychology, Box 871104, Arizona State University, Tempe, AZ 85287-1104 (e-mail: goldinger@asu.edu).

Several other types of illusion are also difficult to classify as perceptual or as memorial, and both labels likely apply. Part of the problem is that the operational procedures for studying perception and memory are often formally quite similar. In many experiments on perception of words or pictures, the stimuli are presented very briefly and subjects are asked to report what they saw (or heard) as soon as possible. In many memory experiments, words or pictures are shown and the subjects are asked, after some retention interval, to recall what occurred. . . . Obviously, the experimental situations blend into one another and it is impossible to say when perceiving is completed and remembering takes over in subjects' reports. (p. 91)

In this passage, Roediger describes explicit memory research, in which subjects actively try to access memory traces (Richardson-Klavehn & Bjork, 1988). To extend his suggestion, we would argue that implicit memory research has virtually erased the perception–memory boundary. Implicit memory measures typically entail some perceptual benefit conferred on previously experienced items. For example, Jacoby and Dallas (1981; Jacoby, 1983a) found that prior experience with words improved their later identification, regardless of overt recognition memory. As a result, memory and perception appear fundamentally linked—neither process can be examined in isolation from the other (see Johnson, 1983). This may reflect imprecise methods, or it may reflect true interdependence. By our view, for an episodic lexicon (cf. Goldinger, 1996) to be feasible, perception and memory *must* be considered to be different functional expressions of a common system. Therefore, rather than try to dissociate perception and memory, the present research sought to confirm their linkage by examining *two-way illusions*: Following prior research, we tried to create perceptual illusions by manipulating memory (Experiments 1 and 2), and to create memorial illusions by manipulating perception (Experiments 3A, 3B, and 4). To extend prior research, we tried to create illusions that were *voice specific* in nature.

The present investigation borrowed methods used by Jacoby and his colleagues to study illusions of perception and memory. In these methods, observers are lured into attributing memorial familiarity to perceptual fluency, or vice versa. This approach has shown wide applicability (see, e.g., Jacoby, Kelley, & Dwyane, 1989; Kelley & Jacoby, 1996; Kelley & Lindsay, 1993), and its corresponding attribution framework accounts for an impressive array of findings (Johnson, Hashtroudi, & Lindsay, 1993; Kelley & Jacoby, 1990). With respect to perceptual illusions, several studies show that experience with a stimulus can change its later subjective perceptual quality. For example, Witherspoon and Allan (1985; see also Mandler, Nakamura, & Van Zandt, 1987) had subjects judge flash durations of *old* and *new* words on a computer screen. Old words led to reliably longer estimates of flash duration: Their context-specific memory traces apparently facilitated perception (Masson & Freedman, 1990; Whittlesea & Brooks, 1988), creating a false impression of

increased bottom-up support. In a conceptually similar study, Jacoby, Allan, Collins, and Larwill (1988) played *old* and *new* sentences to listeners for a perceptual test. The sentences were degraded by varying levels of white noise; the subjects gave subjective noise estimates on a 1–5 scale. Previously heard sentences led to quieter noise estimates, even if listeners did not remember hearing them earlier. Taking a different approach, Jacoby, Woloshyn, and Kelley (1989) induced subjects to give higher fame estimates to previously seen, nonfamous names. Apparently, subjective experience is affected by memory for prior stimuli or situations.

In addition to memory-driven perceptual illusions, Jacoby and colleagues have shown the converse effect—perceptually driven memory illusions. Johnston, Dark, and Jacoby (1985; see also Johnston, Hawley, & Elliott, 1991) found that recognition memory judgments for degraded words were predicted by their speed of perceptual identification. To follow this, Jacoby and Whitehouse (1989) had subjects study a word list. In a later recognition test, old and new words were preceded by subliminal primes (cf. Holender, 1986). When the subliminal primes were related to the target words, the subjects were more likely to respond *old*. This effect vanished when the subjects were aware of the primes. Apparently, when the words were easily perceived (for no obvious reason), they generated an inappropriate feeling of familiarity (cf. Mandler, 1980). The memory illusion was created by a sense of perceptual fluency (see also Lindsay & Kelley, 1996; Whittlesea, Jacoby, & Girard, 1990), further suggesting that memory and perception are intimately linked. The present investigation complemented these prior reports, examining the role of voice information in these symmetric illusions. By the hypothesis of episodic lexical access (Goldinger, 1996, 1997), such illusions should be voice sensitive.

EXPERIMENT 1

Experiments 1 and 2 were modeled after the Jacoby et al. (1988) study. As reviewed, they tested memory for sentences through subjective noise estimates. Memory for any given sentence was indexed by its ability to create an impression of greater perceptual clarity when it was later presented in noise. In their first experiment, Jacoby et al. (1988) examined two conditions. In a *single-task* condition, listeners heard clear sentences during a study session. In a later test, they heard old and new sentences in white noise. The noise was set at three levels (soft, medium, loud), but listeners made subjective noise ratings on a 5-point scale. A *dual-task* condition was similar, but listeners also made old/new sentence judgments. In both conditions, old sentences created an impression of softer noise, suggesting a memory–perception link. Although it was not statistically significant, a trend emerged for dual-task subjects to rate the noise as being louder, suggesting that attention levels also affect subjective ex-

perience. In the dual task, the real and perceived old/new status of sentences affected noise estimates: The noise sounded quieter when old sentences were played, and it also sounded quieter when the subjects made correct old/new decisions. Taken together, the data suggested that perceptual clarity is jointly determined by sentence quality, familiarity, and the availability of attentional resources.

Experiment 1 was modeled after Jacoby et al.'s (1988) first experiment, including a single-task condition (in which listeners rated noise levels while hearing words) and a dual-task condition (in which listeners rated noise levels and made memory judgments). However, we also manipulated study and test voices. Words were heard in two voices during study sessions. In test sessions, half the old words retained their study voices. Prior research (Church & Schacter, 1994; Goldinger, 1996) shows that memory traces for spoken words are quite precise, such that old word + voice combinations show the strongest memory benefits. Following Jacoby et al. (1988) and Goldinger (1996), several predictions arose for Experiment 1. First, we expected old word + voice tokens to create an impression of softer background noise, relative to all the other words. Second, we expected single-task subjects to perceive generally softer noise, relative to dual-task subjects. Third, in the dual task, we expected listeners to perceive softer noise in correct memory trials, relative to incorrect trials.

Method

Subjects. Ninety-six introductory psychology students at Arizona State University participated for partial course credit. All the subjects were native speakers of English who reported normal hearing and either normal or corrected vision. (These participation constraints were applied in all the experiments reported in this article.)

Stimulus materials. All the experiments reported in this article used a common pool of 120 English words, each recorded by two male speakers. To provide variety, an equal number of monosyllabic and bisyllabic words were used, each represented equally at three levels of word frequency (Kučera & Francis, 1967). All the words and their frequencies are listed in the Appendix. To prepare the words for use, all were recorded by two male volunteers in a soundproof booth equipped with an IBM computer, a Beyer-dynamics microphone, and a Marantz DAT recorder. Words were shown on the computer in standard DOS font; volunteers were asked to say each twice and to avoid lapsing into a monotone. The tapes were low-pass filtered at 4.8 kHz, digitized at 10 kHz (with a 16-bit A/D processor), and the clearer token of each word was stored in a digital file.

To ensure that the stimuli were appropriate, two pilot tests were conducted: First, 20 volunteers (10 per speaker) identified all tokens $\geq 90\%$. Next, to ensure that the voices were sufficiently different, 10 volunteers tried to discriminate the stimulus voices in a two-alternative forced-choice test, yielding 98% correct performance. Four separate copies were created for each word: one without noise, one in soft noise (+24 dB signal-noise [S/N] ratio), one in medium noise (+17 dB), and one in loud noise (+10 dB). Notably, even the words in *loud* noise are readily identified (above 90%); a 10-dB S/N ratio is generally considered to be a minor distortion in speech research.¹

Design. Experiment 1 comprised two main conditions. In both conditions, the subjects first identified 60 clear words, with 30 presented in each of two male voices (study session). After a delay,

they heard 120 words, including 60 old words and 60 new words (test session). Half the old words were presented in their original voice; half were presented in the opposite voice. The new words were evenly divided between the two voices. In addition to the voice manipulation, the words were presented in three different levels of noise (orthogonally crossed with the old/new and voice manipulations). In the *single-task* condition, listeners merely rated the noise on a 5-point loudness scale. In the *dual-task* condition, listeners rated noise levels and also made recognition memory judgments.

Procedure. All the subjects were tested in groups of 5–8 students. Each sat in a separate booth, equipped with an IBM computer, Sennheiser headphones, and a two-button computer mouse. Both conditions began with a study session that required identification of 60 words, with 30 presented in each voice. This was a simple two-alternative forced-choice task: On each trial, the subjects were shown two response options (e.g., *river* and *onion*), shown side-by-side in 20-point New York font. (All the experiments were programmed using CSRE, by Avaaz Innovations.) After 1 sec, a clear spoken word (e.g., *river*) was played at approximately 65 dB (SPL), and the subjects had 5 sec to click the appropriate word with the left mouse key. Correct words were shown equally often on each side. The study task took about 3 min, and performance was virtually perfect. During a 5-min delay, the subjects completed a questionnaire on their language habits. Afterward, they received test sessions, which varied across conditions.

In the test sessions, the subjects heard 120 words (60 old, 60 new). Among the 60 new words, 30 were presented in each voice (10 per noise level). Among the 60 old words, 30 were presented in old and new voices, respectively (10 per noise level). Across groups, all the factors (study set, voices, and noise levels) were completely counterbalanced. Thus, all the words were equally experienced as *old* or *new*, in each voice and at each noise level. Upon hearing each word, single-task subjects merely rated its subjective noise level: They were shown a horizontal row of boxes labeled 1 through 5 and clicked their selections with the left mouse key. In dual-task test sessions, the procedure was slightly different. Instead of a single row of boxes labeled 1–5, the subjects saw two rows of five boxes, one above the other. The boxes in the upper row were labeled *new 1* through *new 5*, in 14-point New York font. The boxes in the lower row were labeled *old 1* through *old 5*. The subjects selected a row corresponding to their recognition memory judgment and rated the noise within that row (thereby specifying both judgments by one mouse click).

Results and Discussion

Single Versus Dual Task. The first analysis examined mean noise judgments as a function of task (single, dual), word status (new, old/old voice, old/new voice), and noise level (soft, medium, loud). These data (shown in Table 1) were analyzed in a $2 \times 3 \times 3$ analysis of variance (ANOVA), which revealed several major effects. First, a task effect [$F(1,94) = 24.72, MS_e = 2.35$] reflected a tendency for dual-task subjects to consider the noise to be generally louder. (All the results are reported with implicit reference to a $p < .05$ criterion.) Second, a word status effect [$F(2,188) = 241.81, MS_e = 0.95$] reflected the rank ordering shown in Table 1: Noise levels were considered to be equal when the subjects heard either new or old/new-voice words. By contrast, noise that accompanied old/old-voice words was considered to be far quieter. Third, a noise level effect [$F(2,188) = 204.03, MS_e = 0.59$] confirmed that judgments followed the actual noise values. Among interactions, only task \times noise level was

Table 1
Noise Ratings on a 5-Point Scale (and Standard Errors)
in the Single- and Dual-Task Conditions, Experiment 1

Condition	Probe Type					
	Old/Same-Voice		Old/Different-Voice		New Words	
	Rating	SE	Rating	SE	Rating	SE
Single-Task						
Soft noise	1.57	0.2	2.26	0.3	2.27	0.2
Medium noise	2.40	0.3	2.98	0.4	2.94	0.3
Loud noise	3.18	0.3	3.72	0.3	3.73	0.3
Dual-Task						
Soft noise	2.27	0.3	2.88	0.3	3.03	0.2
Medium noise	2.92	0.2	3.57	0.3	3.60	0.3
Loud noise	3.41	0.3	4.02	0.4	4.00	0.4

significant [$F(2,188) = 6.09$, $MS_e = 0.59$], reflecting a slight convergence across groups at the loud noise level.

Dual Task Only. The second analysis examined only the dual-task data, shown in Figure 1. Mean noise judgments were examined as a function of word status (new, old/new voice, old/old voice), accuracy (correct, incorrect), and noise level (soft, medium, loud). Correct responses included both hits and correct rejections; incorrect responses included misses and false alarms.² The data were remarkably uniform, leading to statistical reliability in all the comparisons. Considering, first, main effects, a word status effect [$F(2,94) = 168.16$, $MS_e = 0.88$] reflected the same rank ordering as that seen in Table 1. New and old/new-voice words created equivalent noise judgments, which exceeded noise judgments to old/old-voice words. An accuracy effect [$F(1,47) = 36.42$, $MS_e = 0.90$] was also observed; the subjects considered the noise levels to be quieter when their recognition judgments were correct. As before, a noise level effect [$F(2,94) = 21.70$, $MS_e = 0.49$] confirmed that judgments followed the true noise levels.

Despite the relatively small discontinuities in the data, all the interactions were reliable. A word status \times accuracy interaction [$F(2,94) = 4.20$, $MS_e = 0.50$] reflected a slightly larger accuracy effect in new-word trials. A word status \times noise level interaction [$F(4,188) = 83.11$, $MS_e = 0.35$] reflected the slightly asymmetric functions (i.e., squares, circles, and triangles in Figure 1) across noise levels. An accuracy \times noise level interaction [$F(4,188) = 90.86$, $MS_e = 0.28$] reflected the asymmetric accuracy functions (i.e., left and right panels in Figure 1) across noise levels. Finally, a word status \times accuracy \times noise level interaction [$F(4,188) = 108.25$, $MS_e = 0.30$] reflected these combined asymmetries in the data.

Experiment 1 replicated and extended Jacoby et al.'s (1988) first experiment. In terms of replication, we observed two effects with spoken words that Jacoby et al. (1988) observed with sentences. First, noise estimates were modified by omnibus task difficulty, so that equivalent noises seemed to be louder to dual-task subjects. In certain regards, this observation is analogous to interactions of word frequency and cognitive workload. For ex-

ample, Becker (1976) found larger frequency effects in lexical decision when subjects engaged in a difficult secondary task, relative to an easy one (see also Herdman, 1992; Herdman & Dobbs, 1989). A related observation was the accuracy effect in the dual task; when the subjects made correct memory judgments, the background noise seemed to be quieter. This suggests that incorrect recognition trials usurp extra processing resources, relative to correct trials.

Second, we observed a memory effect similar to Jacoby et al.'s (1988) result—old words created an illusion of softer background noise. Notably, as is shown in Figure 1, this illusion was not limited to correct recognition trials. Even when listeners missed old words, they still gave them lower noise estimates, relative to new words. This finding rules out a simple response bias, in which listeners uniformly reduce noise estimates whenever they respond *old* (see Jacoby et al., 1988). However, this effect was limited to old-voice repetitions of old words. New stimuli appeared to be functionally equivalent, whether they were novel along the conceptual (new-word) or perceptual (new-voice) dimension.

EXPERIMENT 2

Following earlier claims (Goldinger, 1996; Jacoby, 1983a, 1983b), Experiment 1 suggested that detailed memory traces are created in word perception and affect later perception. To reinforce this hypothesis, Experiment 2 was conducted. Rather than using separate study and test sessions, Experiment 2 involved a simple precue-postcue comparison. This was modeled after Jacoby et al.'s (1988) second experiment, in which listeners heard sets of three words in noise (rather than full sentences) per trial and estimated noise levels. However, the noise-embedded words were presented in a context of clear words, in four possible conditions. In *precue-match* trials, the clear and degraded words were identical, and the clear words were heard first. In *postcue-match* trials, the words were identical, but the clear words were heard after the degraded words. These trials were compared to

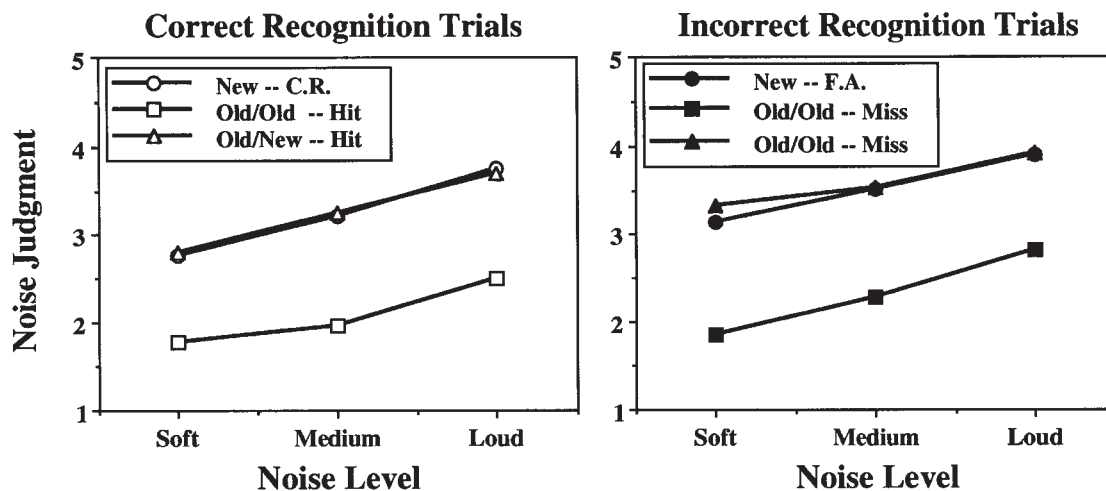


Figure 1. Mean noise ratings from the dual-task condition of Experiment 1, shown as a function of each word's true status (new words, old words/old voices, and old words/new voices) and as a function of the subjects' memory judgments.

precue- and *postcue-mismatch* trials, in which the clear and degraded words differed. Their data were easily summarized: In the *precue-match* trials, the listeners considered the noise to be relatively quiet. In all other trials, noise levels were rated as being equal and relatively loud.

In Experiment 2, we slightly modified Jacoby et al.'s (1988) method. Rather than word triplets, we presented word pairs as test and context stimuli. More importantly, our definition of *matching* was based only on voice matches, rather than word identities. In Experiment 1, it appeared that *new* and *old/new-voice* words were functionally identical, leading to equivalent noise ratings. In Experiment 2, we sought to replicate this finding in a slightly different paradigm. We presented four types of trials: *precue/same voice*, *precue/different voice*, *postcue/same voice*, and *postcue/different-voice*. Following Jacoby et al. (1988), the *precue-postcue* manipulation was used to rule out a simple response bias (i.e., to reduce noise estimates whenever clear, same-voice tokens are used as context). If we observe a true perceptual effect, noise ratings should be reduced only in *precue* trials. In all the trials, the same word pair was presented twice, but the voice was either identical or changed across pairs (same vs. different voice). Target pairs were presented in soft or loud noise, and they either followed or preceded clear context pairs (*precue* vs. *postcue*).

Method

Subjects. Forty introductory psychology students participated for partial course credit.

Stimulus materials. The words from Experiment 1 were used in Experiment 2 but were now grouped into pairs. No test words were presented at the medium noise level; only soft and loud noise was used.

Design. Experiment 2 used a $2 \times 2 \times 2$ design, crossing *precue* versus *postcue*, *same voice* versus *different voice*, and *soft noise* versus *loud noise*. Clear context words were presented either before or

after the degraded target words. The context words either matched or mismatched the target words along the voice dimension, and the target words were presented in either soft or loud noise. All the factors were orthogonally crossed.

Procedure. The subjects were tested with the apparatus used in Experiment 1. On each trial, they saw a ready signal (***) for 500 msec, followed by two spoken words. Trials were randomly drawn from four within-subjects conditions: Depending on condition, the first two words were presented either in the clear or in noise (soft = +24, loud = +10 dB S/N ratios). Words in a pair were separated by 750 msec. After a 1,000-msec interval, the next word pair was presented. The subjects waited until all four words were complete, then rated the noise level for the degraded words on a 1–5 scale, as in Experiment 1. The subjects had 5 sec to respond; any response initiated a 1,500-msec intertrial interval, followed by the next word pair. The entire session consisted of 4 practice trials, followed by 60 experimental trials (15 per design cell).

Results and Discussion

Figure 2 displays mean noise judgments for all conditions. These data were analyzed in a $2 \times 2 \times 2$ ANOVA, examining noise level (soft, loud), cue order (pre, post), and voice (same, different). A main effect of noise level [$F(1,39) = 154.50$, $MS_e = 0.77$] verified that ratings followed the actual noise levels. A voice effect [$F(1,39) = 15.95$, $MS_e = 0.07$] reflected the lower noise ratings in *same-voice* trials. A marginal cue order effect [$F(1,39) = 4.01$, $MS_e = 0.18$; $p = .052$] reflected a trend for *precue* trials to yield lower noise ratings. No two-way interactions were significant, but a noise level \times cue order \times voice interaction [$F(1,39) = 6.20$, $MS_e = 0.07$] was observed. As is shown in Figure 2, the predicted advantage for *precue/same-voice* trials was observed, but it was stronger when words were presented in louder noise.

In general, the Experiment 2 results resemble those from Jacoby et al.'s (1988) second experiment, although they observed a *precue/match* advantage at all noise levels. Given their data, Jacoby et al. (1988) wrote:

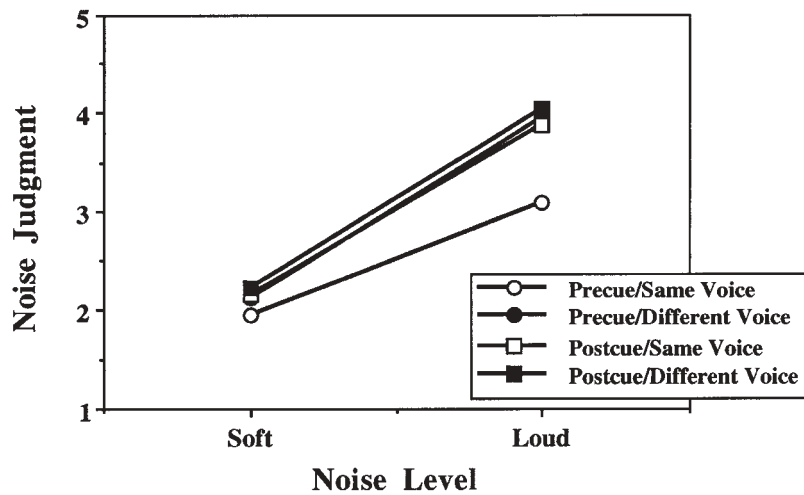


Figure 2. Mean noise ratings (1–5 scale) from Experiment 2, shown as a function of cue order and voice.

The pattern of results is exactly as would be predicted if memory for the cue set in the precue/match condition served to aid interpretation of words presented in the target set, and those effects were mistakenly attributed to a reduction in the loudness of the noise. (p. 245)

This account seems to be appropriate to the present data. Apparently, in precue/same-voice trials, listeners could not distinguish the benefits of memory from the natural variations in noise levels. Given this interpretation, our observed (but unpredicted) interaction makes sense. By definition, loud background noise deprives listeners of bottom-up information to a greater degree than does soft noise. Therefore, listeners should apply top-down knowledge more vigorously when words are presented in louder noise. Only the precue/same-voice trials provided complete top-down support, leading to the interaction.

EXPERIMENTS 3A AND 3B

Taken together, Experiments 1 and 2 support earlier suggestions (Jacoby, 1983a, 1983b) that episodic memory traces are created in word perception and mediate later perception. Listeners' percepts were affected by memories of previously heard words, systematically reducing noise estimates. This effect was limited to words repeated in their original voices; purely conceptual repetition did not produce the illusion. Apparently, episodic traces preserve surface details with considerable precision (Goldinger, 1996). Our general hypothesis is that memory and perception form a continuous system. Experiments 1 and 2 showed that perceptual experience is shaped by episodic memory. The remaining experiments assessed whether memorial experience can be shaped by perceptual manipulations.

Experiments 3A, 3B, and 4 were loosely modeled after a study by Whittlesea et al. (1990) that examined the same basic hypothesis with printed words. In their first

experiment, Whittlesea et al. examined the effect of printed word clarity on subjective impressions of familiarity. On each trial, the subjects saw seven clear words in rapid succession, immediately followed by a probe word for a recognition memory judgment. Half the probes were old (i.e., one of the previous seven words); half were new. The key manipulation was *visual clarity* of the probe words, which were presented either in light (20% occlusion) or in heavy (40% occlusion) visual noise. When the probes were presented in lighter noise, the subjects were biased toward *old* judgments, increasing both hits and false alarms. Whittlesea et al. suggested that perception and memory are interdependent; so, perceptual clarity can be mistaken for familiarity.

Experiment 3A was a conceptual replication of Whittlesea et al.'s (1990) experiment, using auditory words and noise. The subjects heard a series of eight words (four per voice used previously), followed by a probe word for old/new classification. No voice manipulation was included; old words were always presented in their original voices. However, clarity was manipulated; probe words were presented in soft or loud noise, using the S/N ratios from Experiment 2. Following Whittlesea et al., we expected listeners to increase their *old* responses to probe words in softer noise. Experiment 3B was similar, but a voice manipulation was added to the old-word trials. Following prior research (Goldinger, 1996; Palmeri et al., 1993), we expected same-voice probes to generate more hits than did different-voice probes. Among new words, the prediction was identical to that in Experiment 3A; listeners should favor the *old* response to probes presented in softer noise.

Method

Subjects. Experiments 3A and 3B each included 20 introductory psychology students, who received partial course credit.

Table 2
Percentages of Old-Word Responses
(and Standard Errors) in Experiment 3A

Noise Level	Word Status			
	Old Words		New Words	
	%	SE	%	SE
Soft	91.88	0.72	18.88	0.17
Loud	88.50	0.78	8.13	0.25

Stimulus Materials. The previously used words were used again. However, they were now organized into sets of eight clear words, followed by degraded probes.

Procedure. The subjects were tested with the apparatus used previously. Most procedures were identical in Experiments 3A and 3B: Each trial began with a 500-msec cue (**), followed by a series of eight clear spoken words. The words were spoken in two male voices (alternating from one word to the next), with 500-msec intervals between words. After the eighth word, a 400-Hz tone was played for 200 msec, followed 500 msec later by an old or new probe word in soft or loud noise. In Experiment 3B, old probes were presented in either old or new voices (which the subjects were instructed to ignore). In both experiments, old probes were sampled equally from all eight serial positions of the preceding sequences. After each probe, the subjects had 5 sec to select a box labeled *old* or *new* with the left mouse key. Both experiments presented 4 practice trials, followed by 32 test trials.

Results and Discussion

Experiment 3A. Table 2 shows the mean percentages of *old-word* judgments per condition. These data were analyzed in a 2×2 ANOVA, examining word status (old, new) and noise level (soft, loud). A robust word status effect [$F(1,19) = 285.00$, $MS_e = 156.60$] reflected the subjects' generally accurate performance. However, a noise level effect [$F(1,19) = 17.02$, $MS_e = 129.04$] reflected the inflated *old* responses to soft-noise probes. Also, a noise level \times word status interaction [$F(1,19) = 7.44$, $MS_e = 171.18$] reflected an asymmetric noise effect: Soft noise created a bias to respond *old*, but primarily to new probes.

Experiment 3B. Table 3 shows the mean percentages of *old-word* judgments in each condition. These data were analyzed in separate ANOVAs. First, false alarm rates to new words were compared across noise levels, confirming that the 11.63% difference was reliable [$F(1,19) = 15.67$, $MS_e = 86.02$]. Next, hit rates to old words were compared in a 2×2 ANOVA, examining noise level (soft, loud) and voice (same, different). A noise level effect [$F(1,19) = 8.07$, $MS_e = 52.99$] reflected increased hits in soft-noise trials, and a voice effect [$F(1,19) = 32.20$, $MS_e = 55.61$] reflected increased hits in same-voice trials. Also, a noise level \times voice interaction [$F(1,19) = 19.09$, $MS_e = 61.72$] reflected the stronger voice effect in soft-noise trials.

Experiments 3A and 3B both produced their anticipated effects. In Experiment 3A, *old* responses increased when probes were presented in softer noise, primarily among new words. As Whittlesea et al. (1990) suggested, when probes are easily perceived, listeners can misinterpret

clarity as being familiarity. In Experiment 3B, the illusion from Experiment 3A (increased false alarms to soft-noise trials) was replicated, along with a same-voice advantage among old words. Presumably, when listeners hear same-voice probes, they have a bona fide experience of increased familiarity, relative to their perception of new-voice probes. Thus, Experiment 3B seemingly produced both appropriate and inappropriate memory effects. By our view, the noise and voice manipulations both modify the probes' perceptual clarity, which may explain why the voice effect increased in soft-noise trials. If both manipulations affect clarity, their joint effect should be additive, creating a double bonus in same-voice/soft-noise trials.

EXPERIMENT 4

Variations in perceptual clarity can influence standard recognition memory judgments, in which observers discriminate old from new words. In addition, they can influence recognition judgments about words' surface attributes, such as presentation modalities (Kelley, Jacoby, & Hollingshead, 1989). In their second experiment, Whittlesea et al. (1990) examined recognition of *case repetitions*. As in their first experiment, the subjects saw a word series, followed by a probe in light or heavy visual noise. However, the study words were shown in both upper- and lowercase letters. The probes were always old, but their original cases were switched in half the trials. The subjects tried to identify case repetitions. The results were similar to their prior data: When probes were shown in light visual noise, the subjects were biased to respond *old case*. Experiment 4 was conceptually similar to Whittlesea et al.'s experiment, but we tested memory for voice repetitions rather than for case repetitions. The study materials from Experiment 3B were used again, with four clear words in each of two voices. Probe words were always old and were presented in soft or loud noise. The subjects tried to discriminate old from new voices.³

Method

Subjects. Forty introductory psychology students participated for partial course credit.

Stimulus materials. The materials from Experiment 3B were used in Experiment 4.

Procedure. The procedures were mostly identical to those of Experiment 3B, except that all probes were old words. The probes were evenly divided between same- and different-voice trials, orthogonally crossed with soft and loud background noise. The subjects selected either an old-voice or a new-voice response with the left mouse key.

Table 3
Percentages of Old-Word Responses
(and Standard Errors) in Experiment 3b

Noise Level	Probe Type					
	Old/Same-Voice		Old/Different-Voice		New Words	
	%	SE	%	SE	%	SE
Soft	91.88	0.65	79.38	0.74	19.08	0.34
Loud	83.67	0.60	78.13	0.43	7.25	0.20

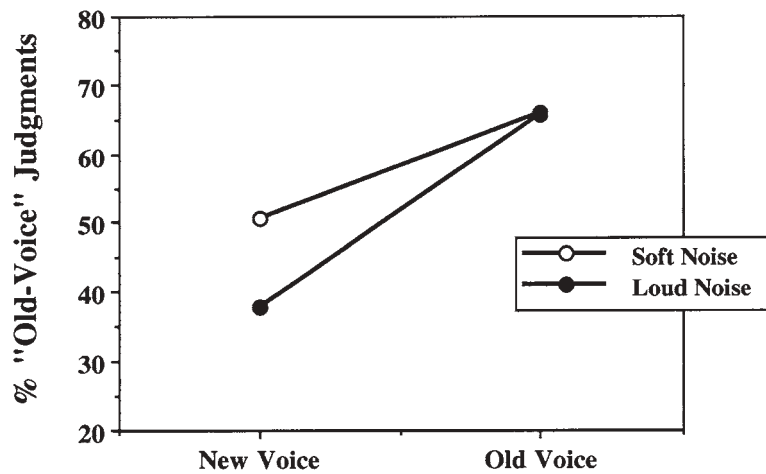


Figure 3. Mean percentages of *old-voice* judgments from Experiment 4, shown as a function of noise level for test words.

Results and Discussion

Figure 3 shows mean percentages of old-voice judgments in each cell of the design. As shown, hit rates were unaffected by the noise manipulation, but false alarms were strongly affected. These data were analyzed in a 2×2 (noise level \times voice) ANOVA. A robust voice effect [$F(1,39) = 60.30$, $MS_e = 308.39$] reflected the generally accurate performance. A marginal noise level effect [$F(1,39) = 3.95$, $MS_e = 326.50$, $p = .054$] reflected a small (6.8%) increase in old-voice judgments in soft-noise trials. However, a noise level \times voice interaction was also observed [$F(1,39) = 7.81$, $MS_e = 294.52$], reflecting the increased false alarms in the soft-noise trials, relative to loud-noise trials (13.4% difference). Experiment 4 conceptually replicated Whittlesea et al.'s (1990) second experiment: Voice judgments to spoken words were influenced by perceptual clarity. As before, listeners were more likely to commit false alarms in response to clearer tokens.

GENERAL DISCUSSION

This investigation complemented two prior studies: Jacoby et al. (1988) showed that memorial familiarity can be mistaken for perceptual fluency. Experiments 1 and 2 extended their study, showing *voice-specific* perceptual illusions: Listeners gave lower noise estimates to previously heard words, but only when the original presentation voices were preserved. Purely conceptual (name) repetitions did not induce this effect. Considering the opposite direction of processing, Whittlesea et al. (1990) showed that variations in perceptual fluency can be mistaken for memorial familiarity: Observers were more likely to guess either *old word* or *old typecase* to words shown in lighter visual noise. Experiments 3A, 3B, and 4 extended their study, showing analogous illusions in the auditory domain: Listeners were more likely to guess either *old word* or *old voice* when words were played at more

favorable S/N ratios. Experiment 3B also replicated the same-voice recognition advantage reported previously (Goldinger, 1996).

With respect to the illusions demonstrated by Jacoby and his colleagues, Roediger (1996) wrote: "The illusions arising from perceptual fluency [. . .] document the intimate link between perceiving and remembering (p. 90)." By creating similar illusions, the present study also suggests that perception and memory are continuous (or identical) systems. Moreover, the voice effects reinforce the hypothesis (Goldinger, 1996, 1997; Masson & Freedman, 1990) that memory contains detailed episodes. Notably, printed and spoken words seem to *underestimate* such episodic effects, relative to other stimuli. Words are extremely high frequency stimuli and are encountered in various forms and contexts. Thus, even a pure exemplar model (Hintzman, 1986) often predicts *prototype* effects in lexical access, as myriad traces contribute to every perceptual event (see Goldinger, 1998).

In comparison with words, less mainstream stimuli can generate robust, high-fidelity memory traces. This is shown empirically by long-lasting repetition (priming) effects. For example, Kolers and Ostry (1974) had students read passages of *rotated text* in several fonts. After 32 days, reading times were shorter for old passages in their original rotated fonts, relative to new fonts. As another example, Salasoo, Shiffrin, and Feustel (1985) examined word and nonword repetition effects in a threshold estimation procedure. Initially, words showed an identification advantage over nonwords. But, after about five repetitions, nonwords were identified equally to words. One year later, the volunteers returned to the laboratory. The procedure was repeated with old words, new words, old nonwords, and new nonwords. Among the words, no effect of prior experience was observed—old and new words were identified equally. However, the old nonwords showed a clear priming effect, relative to new nonwords. More recently, Cave (1997) found robust picture priming,

resulting from single exposures, that lasted 48 weeks without appreciable decay.

These wide-ranging effects all portray memory as a collection of detailed episodes, and they underscore the hypothesized memory–perception link. Moreover, they suggest that occasional null font or voice specificity effects (see Tenpenny, 1995) are likely to be due to processes that engage many traces, thereby obscuring their individual contributions to performance. In this regard, the present results contain one finding that appears to be inconsistent with prior reports. However, on deeper consideration, we believe that our data fit nicely in a larger framework, complementing other results in the literature.

Experiments 1 and 2 demonstrated clear effects of perceptual fluency but no effects of conceptual fluency. That is, noise judgments were reduced by old words, but only when old voices were also repeated. By contrast, Whittlesea (1993) conducted similar tests and found robust conceptual fluency effects. For example, one experiment used semantic contexts to manipulate fluency, rather than visual noise. In each trial, the subjects saw a list of seven words (as in Whittlesea et al., 1990), followed by a target word for a recognition judgment (old/new, with respect to the first seven words). However, instead of showing target words in light or heavy visual noise, Whittlesea presented them in sentence contexts that either strongly predicted the final (target) word (e.g., *The stormy seas tossed the BOAT*) or were fairly neutral (e.g., *He saved his money to buy a BOAT*). In this manner, the fluency of target word perception was increased by conceptual knowledge, rather than perceptual clarity, leading to a powerful illusion: Subjects were more likely to guess *old* to targets in predictive contexts.

Whittlesea's (1993) study shows that conceptual fluency effects are possible, given suitable, discourse-level processing of the target words. This result is consistent with theories of *transfer-appropriate processing*: Episodic memory effects are strongest when the same stimulus dimensions are emphasized during both study and test (Graf & Ryan, 1990; Roediger, Weldon, & Challis, 1989). For example, voice and font effects are generally enhanced when attention is focused on surface attributes during study (Goldinger, 1996; Meehan & Pilotti, 1996) or when particularly salient attributes are used (Jacoby & Hayman, 1987). This suggests that episodic traces are not perceptual analogues, totally defined by stimulus properties. Instead, they seem to be complex *perceptual–cognitive* objects, jointly specified by perceptual forms and their attended cognitive functions (Van Orden & Goldinger, 1994; see Goldinger, 1998). Indeed, Masson & Freedman (1990) proposed that repetition effects are based on context-specific episodes, jointly specified by stimulus details and encoding processes (see also Blaxton, 1989; Whittlesea & Cantwell, 1987).

Jacoby (1983a) noted that “there is a great deal of unexploited similarity between theories of episodic memory and theories of perception. . . . The difference . . . is

largely removed if it is assumed both types of task involve parallel access to a large population of memories for prior episodes” (pp. 35–36). Although we only tested *perceptual* dimensions in this research, the logic of episodic representation easily extends to deeper, conceptual dimensions. Presumably, the focus of attention in learning largely specifies words' episodic content: Voice details will characterize memory traces to the degree that they *mattered* in original processing (Whittlesea, 1987; Whittlesea & Brooks, 1988; Whittlesea & Cantwell, 1987). For example, attended word pairs are apparently stored as single episodes (Boronat & Logan, 1997; Logan & Etherton, 1994; Logan, Taylor, & Etherton, 1996). By extension, paying attention at the level of discourse will encourage the creation of discourse-sized episodes. The *episodic lexicon* may be more than a word collection; it may be a linguistic history, reflecting words in various contexts, nuances, fonts, and voices.

REFERENCES

- BECKER, C. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 556-566.
- BLAXTON, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 657-668.
- BORONAT, C. B., & LOGAN, G. D. (1997). The role of attention in automatization: Does attention operate at encoding, or retrieval, or both? *Memory & Cognition*, **25**, 36-46.
- BROWN, J., & CARR, T. (1993). Limits on perceptual abstraction in reading: Asymmetric transfer between surface forms differing in typicality. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1277-1296.
- CAVE, C. B. (1997). Very long-lasting priming in picture naming. *Psychological Science*, **8**, 322-325.
- CHURCH, B., & SCHACTER, D. L. (1994). Perceptual specificity of auditory priming: Memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 521-533.
- CRAIK, F. I. M., & KIRSNER, K. (1974). The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology*, **26**, 274-284.
- GEISELMAN, R. E., & BELLEZZA, F. S. (1977). Incidental retention of speaker's voice. *Memory & Cognition*, **5**, 658-665.
- GIBSON, J. (1966). *The senses considered as perceptual systems*. Boston: Houghton-Mifflin.
- GOLDINGER, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 1166-1183.
- GOLDINGER, S. D. (1997). Speech perception and production in an episodic lexicon. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 33-66). New York: Academic Press.
- GOLDINGER, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, **105**, 251-279.
- GOLDINGER, S. D., PISONI, D., & LOGAN, J. (1991). On the nature of talker variability effects on serial recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 152-162.
- GRAF, P., & RYAN, L. (1990). Transfer-appropriate processing for implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 978-992.
- HERDMAN, C. M. (1992). Attentional resource demands of visual word recognition in naming and lexical decisions. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 460-470.

- HERDMAN, C. M., & DOBBS, A. R. (1989). Attentional demands of visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 124-132.
- HINTZMAN, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, **93**, 411-428.
- HINTZMAN, D. L., BLOCK, R., & INSKIP, N. (1972). Memory for mode of input. *Journal of Verbal Learning & Verbal Behavior*, **11**, 741-749.
- HOLENDER, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral & Brain Sciences*, **9**, 1-23.
- JACKSON, A., & MORTON, J. (1984). Facilitation of auditory word recognition. *Memory & Cognition*, **12**, 568-574.
- JACOBY, L. L. (1983a). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 21-38.
- JACOBY, L. L. (1983b). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning & Verbal Behavior*, **22**, 485-508.
- 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, & Cognition*, **14**, 240-247.
- JACOBY, L. L., & DALLAS, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306-340.
- JACOBY, L. L., & HAYMAN, C. (1987). Specific visual transfer in word identification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 456-463.
- JACOBY, L. L., KELLEY, C. M., & DWYAN, J. (1989). Memory attributions. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391-422). Hillsdale, NJ: Erlbaum.
- JACOBY, L. L., & WHITEHOUSE, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, **118**, 126-135.
- JACOBY, L. L., WOLOSZYN, V., & KELLEY, C. M. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, **118**, 115-125.
- JOHNSON, M. K. (1983). A multiple-entry, modular memory system. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 17, pp. 81-123). New York: Academic Press.
- JOHNSON, M. K., HASHTROUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.
- JOHNSTON, W. A., DARK, V., & JACOBY, L. L. (1985). Perceptual fluency and recognition judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 3-11.
- JOHNSTON, W. A., HAWLEY, K. J., & ELLIOTT, J. (1991). Contributions of perceptual fluency to recognition judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 210-233.
- KELLEY, C. M., & JACOBY, L. L. (1990). The construction of subjective experience: Memory attributions. *Mind & Language*, **5**, 49-68.
- KELLEY, C. M., & JACOBY, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory & Language*, **35**, 157-175.
- KELLEY, C. M., JACOBY, L. L., & HOLLINGSHEAD, A. (1989). Direct and indirect tests of memory for source: Judgments of modality. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 1101-1108.
- KELLEY, C. M., & LINDSAY, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory & Language*, **32**, 1-24.
- KIRSNER, K. (1973). An analysis of the visual component in recognition memory for verbal stimuli. *Memory & Cognition*, **1**, 449-453.
- KOLERS, P. A., & OSTRY, D. J. (1974). Time course of loss of information regarding pattern analyzing operations. *Journal of Verbal Learning & Verbal Behavior*, **13**, 599-612.
- KRULIE, G., TONDO, D., & WIGHTMAN, F. (1983). Speech perception as a multilevel processing system. *Journal of Psycholinguistic Research*, **12**, 531-554.
- KUČERA, H., & FRANCIS, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- LINDSAY, D. S., & KELLEY, C. M. (1996). Creating illusions of familiarity in a cued recall remember/know paradigm. *Journal of Memory & Language*, **35**, 197-211.
- LOGAN, G. D., & ETHERTON, J. L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1022-1050.
- LOGAN, G. D., TAYLOR, S. E., & ETHERTON, J. L. (1996). Attention in the acquisition and expression of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 620-638.
- MANDLER, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, **87**, 252-271.
- MANDLER, G., NAKAMURA, Y., & VAN ZANDT, B. J. S. (1987). Nonspecific effects of exposure on stimuli that cannot be recognized. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 646-648.
- MANSO DE ZUNIGA, C., HUMPHREYS, G., & EVETT, L. (1991). Additive and interactive effects of repetition, degradation, and word frequency in the reading of handwriting. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading* (pp. 10-33). Hillsdale, NJ: Erlbaum.
- MASSON, M. E. J., & FREEDMAN, L. (1990). Fluent identification of repeated words. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 355-373.
- MCNEILL, D., & LINDIG, K. (1973). The perceptual reality of phonemes, syllables, words, and sentences. *Journal of Verbal Learning & Verbal Behavior*, **12**, 419-430.
- MEEHAN, E. F., & PILOTTI, M. (1996). Auditory priming in an implicit memory task that emphasizes surface processing. *Psychonomic Bulletin & Review*, **3**, 495-498.
- 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.
- RICHARDSON-KLAVEHN, A., & BJORK, R. A. (1988). Measures of memory. *Annual Review of Psychology*, **39**, 475-543.
- ROEDIGER, H. L., III (1996). Memory illusions. *Journal of Memory & Language*, **35**, 76-100.
- 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, WELDON, M. S., & CHALLIS, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3-14). Hillsdale, NJ: Erlbaum.
- SALASOO, A., SHIFFRIN, R. M., & FEUSTEL, T. C. (1985). Building permanent memory codes: Codification and repetition effects in word identification. *Journal of Experimental Psychology: General*, **114**, 50-77.
- SCHACTER, D., & CHURCH, B. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 915-930.
- TENPENNY, P. L. (1995). Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin & Review*, **2**, 339-363.
- VAN ORDEN, G. C., & GOLDINGER, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 1269-1291.
- WHITTLESEA, B. W. A. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 3-17.
- WHITTLESEA, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1235-1253.
- WHITTLESEA, B. W. A., & BROOKS, L. R. (1988). Critical influence of particular experiences in the perception of letters, words, and phrases. *Memory & Cognition*, **16**, 387-399.
- WHITTLESEA, B. W. A., & CANTWELL, A. L. (1987). Enduring influence of the purpose of experiences: Encoding-retrieval interactions in word and pseudoword perception. *Memory & Cognition*, **15**, 465-472.

WHITTLESEA, B. W. A., JACOBY, L. L., & GIRARD, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory & Language*, *29*, 716-732.

WITHERSPOON, D., & ALLAN, L. G. (1985). The effect of a prior presentation on temporal judgments in a perceptual identification task. *Memory & Cognition*, *13*, 101-111.

NOTES

1. By using S/N ratios, the present study deviated in method from Jacoby et al. (1988), who used a white-noise generator set at fixed values. We preferred S/N ratios for two reasons. First, by using each word's amplitude as its own referent, our noise values (relative to speech) were truly equivalent (across tokens) at each level. Second, our testing method required listeners to judge three noise levels on a 5-point scale. Using S/N ratios allowed natural variations in *absolute* noise levels across tokens, making our stimulus-response mappings more reasonable.

When other researchers have used similar noise manipulations, they have added procedures to verify their efficacy. For example, Whittlesea

et al. (1990) collected word-naming times and recognition memory judgments on each trial. To keep our procedures simple, we conducted a separate, auditory naming test on our stimuli. Thirty volunteers listened to all stimulus words (10 listeners per noise level), repeating each as quickly as possible. Naming times were 428 msec (soft noise), 499 msec (medium noise), and 566 msec (loud noise). These times reliably increased across noise levels [$F(2,27) = 59.02$, $MS_e = 929.30$], but error rates (5.5%, 6.8%, and 6.8%, respectively) did not. This test verified that our noise values affected perceptual fluency, without dramatically affecting accuracy.

2. In a separate analysis, mean recognition accuracy was unaffected by word status [$F(2,94) = 1.50$, $MS_e = 0.04$] but was affected by noise level [$F(2,94) = 3.95$, $MS_e = 0.01$]. Responses were slightly (2.5%) less accurate at the medium noise level, relative to the soft and loud levels. The word status \times noise level interaction was nonsignificant.

3. We also conducted this experiment with new words included. When listeners false-alarmed to new words, they had to specify either old or new voice. As expected, *old-voice* decisions were more likely in soft-noise trials (10.5%) than in loud-noise trials [2.0%; $F(1,19) = 18.09$, $MS_e = 1.77$].

APPENDIX A

Stimulus Words (and Frequencies) Used in All Experiments

Bisyllables	Frequency	Monosyllables	Frequency	Bisyllables	Frequency	Monosyllables	Frequency
High-Frequency Words (150–250, Kučera–Francis)							
river	165	stage	174	title	77	prove	53
market	155	class	207	forget	54	grass	53
police	155	sound	204	coffee	78	dust	70
figure	209	black	203	novel	59	fresh	82
beyond	175	floor	158	fashion	69	watch	81
nature	191	book	193	favor	78	knife	76
father	183	cold	171	garden	60	tone	78
spirit	182	town	212	listen	51	throat	51
music	216	ground	186	master	72	speed	83
recent	179	north	206	vision	56	lake	54
table	198	girl	220	Low-Frequency Words (<5)			
party	216	late	179	bicep	1	germ	3
report	174	wall	160	rustic	3	vest	4
picture	161	fire	187	nectar	3	dire	1
basis	184	bring	158	parcel	1	malt	1
person	175	rest	163	mingle	2	wilt	3
value	200	lost	171	staple	1	grub	2
common	223	care	162	gusto	2	soot	1
final	156	plan	205	forage	3	blur	3
single	172	hard	202	deport	1	crow	2
Medium-Frequency Words (50–100)				pigeon	3	vine	4
symbol	54	rule	73	venom	2	mule	4
dozen	52	moon	60	nugget	1	chunk	2
handle	53	safe	58	garter	2	weed	1
cousin	51	bank	83	portal	3	hoop	3
active	88	band	53	beacon	5	kelp	2
permit	77	crowd	53	patron	4	knack	4
career	67	phone	54	jelly	3	leash	3
careful	62	chair	66	cavern	1	fade	2
captain	85	tree	59	hazel	2	stale	4
balance	90	bright	87	wedlock	2	raft	4