Familiarity, Similarity and Memory for Speech Events

Sonya M. Sheffert and Richard M. Shiffrin

Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405

1 This research was supported by NIH-NIDCD Research Grant DC00111, NIH-NIMH Research Grant MH12717 and NIH-NIDCD Training Grant DC00012 to Indiana University, Bloomington.

2 Luther Dana Waterman Professor of Psychology, Indiana University, Bloomington.
Familiarity, Similarity and Memory for Speech Events

Abstract. The experiments described in the present report were designed to explore the "registration and learning" (Hintzman, Curran & Oppy, 1992) of instance-specific acoustic information. The purpose of this investigation was to determine if familiarity with a spoken word affects the acquisition of the perceptual details specifying a talker's voice and a word's plural form. It is generally assumed that the more we experience a word, the more complete our knowledge about the word will be. To test this hypothesis, the frequency of study presentation, the similarity between study and test items, and the instructions at retrieval were manipulated. Participants studied a list of words in which target items were repeated various numbers of times. In the test phase, participants heard a list of words that contained new items as well as old targets that were either the same as the studied item or were in a different voice, in a different plurality, or in a different voice and a different plurality. Subjects estimated the frequency of each word, with the additional caveat that they restrict their positive frequency judgments to items that were in the same form as the study item. The results revealed that presentation frequency improved listeners' knowledge that a word occurred, without improving their ability to discriminate it from a perceptually different word. This dissociation between knowing that a spoken word occurred and knowing perceptual details about the word suggests that instance-specific perceptual attributes of a spoken word are less likely to be encoded if the word has become very familiar over several repetitions. The collective import of this research is show that "registration without learning" occurs for auditorily presented words, for simple and complex stimulus dimensions, and across various experimental settings. The findings add to a growing literature on the factors that are important in determining what knowledge a subject encodes during spoken word recognition, and help to delimit theoretical interpretations of spoken language processing and memory.

Introduction

Among the most reliable findings in the literature on learning and memory is that familiarity facilitates recognition. Items that have become familiar over several prior presentations are recognized more accurately than new items or items presented only once. The research to be described here was motivated by an interest in discovering the changes in information processing that occur when an item is repeatedly presented to an observer. The aim was to determine if familiarity with a spoken word results in more efficient and complete acquisition of detailed perceptual information, or if instead, familiarity hinders the acquisition of perceptual information.

Memory for Repeated Events

A method researchers use to directly manipulate the amount of experience a subject has with an item is to vary the number of study repetitions the item receives. After the study list has been presented, a recognition memory or frequency judgment test can then be administered. Typically, well-known or often experienced items are recognized more accurately and given higher frequency estimates than items experienced once. An advantage the frequency judgment task has over a test of recognition memory is that the frequency judgment task provides information that is specific to the number of traces retrieved by a test item, rather than merely indicating the presence or absence on a trace. Moreover, the fact that judgments of
frequency are quite accurate under a range of encoding and retrieval circumstances (Flexser & Bower, 1975; Greene, 1984; Harris, Begg & Mitterer, 1980; Hintzman & Block, 1971; Hockley, 1984; Naveh-Benjamin & Jonides, 1986; Rose & Rowe, 1976; Rowe, 1974), has led some theorists to assume that the information that underlies frequency judgments probably accures naturally as an automatic consequence of event perception (Hasher & Zacks, 1974; 1984).

Registration Without Learning

It is generally assumed that the more we experience an item, the more complete our knowledge about the item will be. If it really is the case that “practice makes perfect,” one would expect that our memory for highly familiar items should be very accurate. This intuitive idea was tested in a series of recent experiments by Hintzman, Curran & Oppy (1992). In one study, pictures were presented 1, 3, 8 or 15 times in a list. At test, the pictures were presented in the same orientation as at study or in a different (mirror reversed) orientation. Subjects estimated the frequency with which a study item occurred, with the additional requirement that positive frequency judgments should be given only to items that were in the same orientation. Hintzman et al. found that frequency judgments increased with presentation frequency. This indicates that subjects “registered” or catalogued the study repetitions in memory. However, given that an “old” judgment was made, the frequency ratings were similar for pictures tested in the same orientation or different orientation. In particular, Hintzman et al. found that there was no improvement in subjects’ ability to distinguish same-orientation items from different-orientation items beyond the third repetition. This finding constitutes what has since come to be known as “registration without learning.”

What is surprising about this result, according to the authors, is that someone can see a picture 15 times, “each time attending to it at least well enough to record the fact of repetition, without improving measurably in the ability to discriminate it from its mirror image” (Hintzman et al., 1992, p. 679). The same results have been found with other visual materials (singular and plural printed words). These effects are independent of listwise serial position, arising even when the first presentation of an item occurs near the end of the study list.

Hintzman & Curran (1995) have also shown that this bias against learning perceptual features on later repetitions is difficult to counteract. For example, simply telling subjects to attend to a particular stimulus dimension during study does not improve subjects’ learning of the perceptual feature. Hintzman and Curran found that subjects must be made aware of the incompleteness of their knowledge before they show improvement in discrimination skill. This was accomplished by providing subjects with the opportunity to assess their own knowledge about the perceptual form of the word and, when incorrect, to focus additional attention on the relevant feature. As a consequence of this manipulation, their subjects showed a constant rate of learning over presentation frequencies.

Encoding of Repeated Items

Why would subjects fail to process a repetition as thoroughly as they do a first occurrence? The explanation proposed by Hintzman et al. (1992) is that repetition affects subjects’ encoding processes. Specifically, familiarity leads to a redistribution of perceptual processing that changes the nature and extent of lower-level analyses (cf. DiGirolamo & Hintzman, 1997; Johnston & Hawley, 1994; Kraut, Smothergill & Farkas, 1981; Nickerson & Adams, 1979; Posner & Boies, 1971). According to this view, repeating an item activates previously stored memory traces of that item. The “priming” or activation that occurs as a result of repetition exerts a beneficial top-down influence on perceptual processing by decreasing the
amount of sensory information needed to identify the item. This improves the speed and accuracy of item recognition.

However, the perceiver pays a price for this facilitation because processing of incoming sensory information is inhibited. The greater reliance on top-down conceptual information that accompanies familiarity decreases subjects’ use of bottom-up data-driven information. Consequently, the stimulus attributes of a familiar item are less likely to be noticed. In order to draw attention away from conceptual features and towards perceptual features, aspects of the item’s context must change appreciably across repetitions. In fact, Tulving and Kroll (1997) make the strong claim that a necessary condition for long-term storage of information is the perception of novelty. The redistribution of perceptual processing brought about by familiarity helps to optimize behavior because it allows perceivers to capitalize on prior experiences, while also being able to readily detect and represent novel objects and events (Treisman, 1992).

The Retrieval of Repeated Items

Registration without learning has attracted considerable interest not only because it runs counter to our intuitions, but also because it runs counter to most current memory theories. The dissociation between knowing that a word occurred and knowing details about the perceptual form of the word suggests the use of two different retrieval processes (familiarity and recall). In contrast, most contemporary memory theorists (e.g., Gillund & Shiffrin, 1984; Hintzman, 1988; Humphreys, Bain & Pike, 1989; Metcalfe, 1982; Murdock, 1982; Ratcliff, 1978; Shiffrin & Steyvers, 1997) assume that recognition decisions (and, by extension, frequency judgments) can be modeled by a single process (familiarity) based on the match of a test item to the entire contents of memory.

In these memory models, memory is conceived of as a collection of episodic memory traces that represent each occurrence of a word and include information about the word’s sensory form and context. Depending on the model, repetition can serve to strengthen the original presentation’s trace (e.g., SAM [Raaijmaker & Shiffrin, 1981; Gillund & Shiffrin, 1984]), produce a new memory trace (e.g., Minerva 2 [Hintzman, 1988]) or both (e.g., REM [Shiffrin & Steyvers, 1997]). During retrieval, information about the test item’s form and context combine into a single memory probe. Memory access is “global” in the sense that all the traces in memory are evaluated in parallel and contribute to the output of memory. The degree to which item and context features match is used as an index of the global activation or familiarity of the test item. Increases in familiarity result in an increase in the likelihood of identifying a test item as “old” and an increase in the magnitude of its frequency estimate. Likewise, similar distractor items (e.g., different orientation or different pluralization) that are falsely recognized will produce a response pattern akin to that of a less familiar target item.

Indeed, the registration without learning experiments confirmed this latter expectation (Hintzman et al., 1992). Hintzman’s experiments showed that the judgments of frequency (JOFs) for similar distractor items increase proportionally to the frequency of target items. However, the experiments also revealed a large number of JOF=0 (i.e., “correct rejections”), a finding which is not expected if judgments are based on a single process. Hintzman infers that zero judgments occurred because sometimes subjects were able to explicitly recall enough information about the target item to allow them to rule out a similar test item. This “recall-to-reject” strategy requires a recall-like search component that supports the learning and recollection of item-specific information, in addition to the direct-access familiarity component responsible for registration.
Objectives of the Present Experiments

The broad theoretical implications of Hintzman's experiments highlight the importance of replicating and extending the "registration without learning" paradigm to a different domain and to new and more complex features. In addition, the literature on the effects of a perceptual match across study and test episodes on measures of explicit memory is based almost entirely on recognition memory tests (Richardson-Klavehn & Bjork, 1988). Therefore, it would be worthwhile to examine the retention of perceptual details using a different kind of explicit memory task.

The experiments described in the present report were designed to explore how experience with a spoken word affects the acquisition of instance-specific acoustic information. The first goal was to establish the generalizability of the "registration without learning" phenomena to the auditory domain. To this end, we developed an auditory version of visual plurality manipulation reported in Hintzman et al. (1992). The only major departure from Hintzman's design is that the study-to-test items vary by two dimensions (plurality, voice, or both), rather than one dimension. The experiments assessed the effects of word repetition on subjects' ability to make fine-grained distinctions among same-form and different-form items.

We also vary two types of features: voice and plurality. The motivation for comparing the effects of plural change with the effects of voice change was a suspicion that the lack of learning for plurality in Hintzman's experiments may have been related to the fact that plurality was a minimal or single feature. In contrast, there are reasons to expect that voice learning may be more robust than other stimulus dimensions and consequently may increase over word repetitions. One reason for this expectation is that the processing of talker information and phonetic information occur in a parallel-contingent fashion (Green, Tomiak, & Kuhl, 1997; Mullennix & Pisoni, 1990), whereas other sources of acoustic variation, such as speaking rate, amplitude (Bradlow, Nygaard & Pisoni, in press) and possibly plurality may not be processed integrally.

A second reason for this expectation is that a change in voice may be perceptually more salient than a change in plurality and therefore more likely to be noticed and encoded. The acoustic correlates that result from a change in voice are very complex, involving many acoustic-phonetic features, and affect not only the perception of the individual talker, but also the perception of the phonetic segments produced by the talker (Bricker & Pruzansky, 1976; Peterson & Barney, 1952). Moreover, there is ample evidence demonstrating that listeners encode details about a talker's voice in long term memory, and can explicitly recollect voice information during a word recognition test (Bradlow et al.; in press; Craik & Kirsner, 1974; Geiselman & Bellezza, 1976; Hintzman, Block & Inskeep, 1972; Palmeri, Goldinger, & Pisoni, 1993; Sheffert, in press; Sheffert & Fowler, 1995).

In contrast to voice, our plurality manipulation is acoustically very simple and unlikely to affect the perception of the preceding phonetic segments. Specifically, in the present study, pluralization of a singular word was accomplished by a simple acoustic transformation that involved splicing a [z] sound to the end of a word. The advantage of the splicing technique was that it eliminated the possibility of coarticulatory effects present in naturally produced plural tokens, and ensured that the only acoustic difference between singular and plural items was in the presence or absence of [z].

Using these stimulus materials, we manipulated the frequency of study presentation, the similarity between study and test items, and the instructions at retrieval. In Experiment 1, participants studied a long list of words that contained 24 target words produced by a male or a female talker. These words were either
singular or plural. The targets were presented once or repeated five or ten times in the list. Participants listened to each word with the expectation that their memory would be tested after completing the study phase. In the test phase, participants heard a list of 48 successively presented words. Half were new items, and half were old targets that were either the same as studied items or were similar distractors. The degree of similarity between targets and distractors varied by one or two dimensions: Similar items were in a different voice, in a different plurality, or in different voice and in a different plurality from studied items.

In the test phase, subjects made frequency judgments based on word token information. That is, subjects estimated the frequency with which an item occurred at study, with the caveat that they restrict their positive JOF's to those items that were in the same form as the study item. For one group of subjects, "same form" meant same voice. They were told that test words were spoken in the same voice or in a different voice as at study, and that they should give "different voice" items a frequency judgment of zero. No mention was made of the plurality dimension. For another group of subjects, "same form" meant same plurality. They were told to give "different plurality" items a zero, and were not told about the plurality dimension. Subjects were only told about one of the two dimensions. This was done in order to determine if a change in the irrelevant dimension would influence frequency judgments or same-form/different-form discriminations. Together, our design allows us to replicate and extend Hintzman's experiments to the auditory modality and to determine the extent to which feature complexity, feature integrality and test instructions moderate the "registration without learning" phenomena.

**Experiment 1**

**Method**

**Participants**

The participants were 48 Indiana University students who volunteered for the experiment in exchange for course credit. Half received the voice instructions and half received the plural instructions. All were native speakers of English and reported no history of any speech or hearing disorders at the time of testing. The participants were tested in groups of four or fewer. The data from four voice and two plural subjects were eliminated owing to a failure to follow instructions.

**Apparatus**

Presentation of the stimulus materials and collection of the response data were carried out for each participant on an IBM-compatible personal computer. The words were presented binaurally to subjects at 75 dB SPL over matched and calibrated stereophonic headphones (BeyerDynamic DT-100).

**Stimulus Materials**

The target stimuli were 24 monosyllabic singular words produced by a male speaker and a female speaker. The words were selected from the Indiana University Multi-talker Speech Database (see Torretta, 1996, for a detailed description). The audio recordings were obtained by asking the speakers to read each

---

3 An obvious third condition is one in which subjects restrict their positive frequency judgments to those items that were in the same voice and the same plurality as the study item. This experiment was conducted. We found that under these conditions, subjects favored a strategy whereby they judged most target items to be in a different form, even those that were in exactly the same voice and plurality. This response bias produced an inordinate number of zero judgments, and effectively rendered the data uninterpretable. Consequently, the complete results from this experiment will not be reported.
item in citation form. The recordings were made in a sound-proof booth, and then digitized at a 20 kHz sampling rate and equated for peak amplitude.

Plural forms of the singular target words were created by splicing a sample of word-final frication onto a singular word. For example, /s/ was spliced onto the end of the word dog to produce dogs. To ensure that the plurality of target items was perceptually salient, a pretest was conducted which assessed the intelligibility of the singular and plural forms from each talker. A separate group of five listeners (two speech scientists and three undergraduates) identified each word. Overall intelligibility, as defined by the percentage of correct word identification, was quite high (95%), and, importantly, plural items were unanimously heard as such.

The study list consisted of 24 distinct target items, with eight assigned to each of the presentation frequencies: 1, 5, and 10. Half of the items were presented in a female voice and half in a male voice. Similarly, half of the items were singular and half were plural. The form of an item was consistent across all presentations. Four different study lists were constructed in which each of the 24 target items was rotated through a different, randomly selected condition. An additional 77 items were randomly selected as filler items, and were used to decrease the possibility of spacing effects on explicit memory. The filler items were produced in the same two voices as the target items and included both singular and plural items. Each filler item occurred only once in the study list. Subjects were randomly assigned to study lists, and the order of presentation was random for each subject. Each study list was 205 items long.

The test phase consisted of 24 target items and 24 new items. Target items were of four types: Six target items were repeated in the same voice (designated as “same” items), and were the exact tokens used in the study phase. Six target items were repeated in a different voice (“different-voice”); six were repeated in a different plurality (“different-plural”); and six were repeated in a different voice and in a different plurality (“both-different”). An additional 24 items were new fillers. Subjects received one of four different test lists, depending on the list they studied. The order of words was individually randomized for each subject.

Procedure

There were two phases in the experiment: study and test. The study list consisted of 205 successive word presentations. Subjects were told that some of the words would be presented more than once in the list. They were also told to try to remember each occurrence of the word because they would receive a memory test at the end of the study session which would require them to estimate the number a times each word was presented. After listening to each word (which typically took approximately 1.5-2 sec), participants pressed a key to signal the computer to begin playing the next word. The interval between this key press and the onset of a new trial was 1 s. The study phase was followed by a short (2-2.5 minutes) retention interval, followed immediately by the test phase.

The test list consisted of 48 successive word presentations. The order of presentation was random for each subject. Participants in the voice instruction group were told that they would hear a list of words, some of which were present on the study list, and some of which were new. They were told that their task was to estimate the number a times a word was presented during the study phase. Following Hintzman et al. (1992), subjects were explicitly instructed to pay close attention to whether the word was in the same voice as at study, and were told to only type the number of times they heard the word exactly as heard on the test list. Subjects were provided with the following example:
Your task is to estimate the number a times a word was presented during the study phase. Only type the number of times you heard the word exactly as heard on the test list. Pay close attention to whether the word was spoken by a male or female voice. For example, in the study list you could have heard the word "cat" spoken by the female twice, but never "cat" spoken by a male. The correct answer depends on whether "cat" (the female version) is presented to you in the test. If "cat" (the original female version) is in the test, you would be correct if you answered "two". If "cat" (the male version) is on the test, you should answer "zero" because that exact word was not presented. Finally, the form of a word was consistent across word repetitions.

The complement to these instructions were given to subjects in the plural instruction group. The instructions focused attention on the plurality of the item, rather than the voice. Subjects in both groups were also told that no word was ever presented more than 15 times, so all their frequency judgments should be between 0 and 15. Each time a word was presented, the subject had to type their response using the numeric keys. When they completed their frequency judgment, the subject would press the return key and the next trial would begin 1 second later. The entire experiment lasted approximately 30 minutes, after which the subjects were debriefed.

**Results and Discussion**

**Frequency Distributions**

The purpose of this measure is to determine if the shape of the response distribution differed for items that were identical to the studied words as compared to items that were similar to the studied words. If a single process (familiarity) underlies frequency judgments, the response distributions for similar items should resemble scaled down versions of the response distributions of same items.

Figures 1a and 1b displays the frequency judgment distributions for each item type from frequency = 10. These patterns are representative of all three frequency conditions. Figure 1a represents the voice instruction group. In that figure, the top two panels (same and different-plural) show a unimodal response distribution, with the majority of responses between 6 and 12. The bottom two panels (different-voice and both-different) show a bimodal response distribution, with the first mode at 0, and the second mode between 6 and 12. The large number of 0 responses reflect subjects’ ability to reject similar distractors, and indicate that some detailed information about the physical form of words was encoded during the study phase. The remaining nonzero JOFs are “false recognitions” and reflect subjects’ failure to discriminate same items from similar items.

---

Insert Figure 1a about here

---

A complementary pattern was found in the plural instruction group (shown in Figure 1b). Same and different-voice items produced unimodal distributions, and different-plural and both-different items produced a bimodal distribution.

---

Insert Figure 1b about here

---
Figure 1a. Histograms of the frequency judgment distributions for subjects in the voice instructions condition. The figure displays the number of responses for each item type at presentation frequency = 10.
Figure 1b. Histograms of the frequency judgment distributions for subjects in the plural instructions condition. The figure displays the number of responses for each item type at presentation frequency = 10.
The response distributions also reveal an effect of instructions. Figure 1a shows that subjects were largely unaffected by study-to-test changes in the plurality under conditions that required them to discriminate items based on a voice match. Compatibly, Figure 1b shows that subjects were unaffected by changes in voice under conditions that required them to discriminate words based on a plurality match. Thus, although subjects were not explicitly told to ignore the irrelevant stimulus dimension when making their same form-different form discriminations, they nevertheless did so. This pattern was found in several of the following measures.

Mean Frequency Judgments

This measure provides information about subjects' sensitivity to variation in presentation frequency and whether JOFs overestimate or underestimate the actual presentation frequency. Because the response distributions of the different-form items were not normally distributed, only the mean JOF's from the same items were analyzed. Mean JOFs for same items are displayed as a function of frequency and instruction condition in Figure 2. The positive slope of the lines indicates that subjects in both groups were sensitive to presentation frequency. Mean JOFs across the three frequency levels (collapsed over instructional group) were 1.25, 4.2 and 8.1. An ANOVA on the frequency estimates from each test group revealed a highly significant effect of frequency for the voice group \( F(2, 38) = 39.45, p < .0001 \) and for the plural group \( F(2, 22) = 82.56, p < .0001 \). A separate analysis comparing frequency estimates across the two groups showed that there was no difference across instructional conditions (this holds true for all subsequent analyses).

Nonzero JOFs

This measure assessed whether items that were falsely recognized (given a JOF of greater than 1, rather than 0) were perceived as less familiar, and consequently given a lower JOF than items that were correctly recognized. The analysis excluded all JOF = 0, and analyzed only the nonzero frequency judgments.

Figure 3 displays the mean nonzero JOFs as a function of test condition, presentation frequency and item type. The figure shows that JOFs increased as actual frequency increased, and that overall performance levels were similar across the two test conditions (the mean nonzero frequency judgment was 4.6 for the voice group and 4.7 for the plural group). The figure also shows differences among the four item types at frequency = 10, but only in the voice instruction group.

An ANOVA was conducted on the nonzero frequency judgments at each frequency level, separately for each test condition. Each analysis included the within-subject factor of item type. Subjects who failed to contribute at least one nonzero judgment for each item type at each frequency were excluded from the analysis. This was necessary to prevent subject selection artifacts from contaminating the inferential statistics (see Hintzman and Curran, 1994).
Figure 2. Mean judgments of frequency (JOFs) are displayed as a function of actual frequency and test instruction group, collapsed over test item type.
Figure 3. Mean judgments of frequency (JOFs) greater than zero are displayed as a function of presentation frequency and test item. The upper panel displays frequency judgments for subjects in the voice instructions condition. The lower panel displays frequency judgments for subjects in the plural instructions condition.
For subjects in the voice instruction condition, item type was significant at frequency = 10 [F(3, 42) = 3.06, p < .04], reflecting the fact that the frequency judgments for same items were higher than frequency judgments for similar items. Specifically, nonzero frequency judgments were higher for same items than for different-voice items [t(17) = 2.41, p < .03], and were higher for different-voice items than for different-plural items [t(17) = 2.88, p < .01]. Thus, study-to-test changes in voice led to lower JOF's, presumably because the words were less familiar than same-voice items. This perceptual match effect is similar to the pattern obtained in Hintzman et al. (1992; 1995), except that his data reveal differences between same and similar items at lower frequencies.

For the plural instruction group, the similarity between study and test items was not reliable at any frequency, which indicates that same and similar items were perceived as equally familiar. One reason for this difference across instructional conditions may be that plural information is less likely to be encoded in memory than voice information. This hypothesis is addressed in a subsequent word recognition measure and directly tested in Experiment 2.

Alternatively, subject variability may be the reason for the null effect of item similarity in the plural group, as well as the weak effect at lower frequencies in the voice group. Examination of the individual subject data reveals extremely variable performance across subjects. Consequently, it may be necessary to run as many subjects as Hintzman (typically 70 subjects per condition) in order to replicate his findings.

Word Recognition

In this analysis, frequency judgments were collapsed into a binary recognition measure. The recognition measure was obtained by classifying all words given a judgment greater than 0 as "old" and those given a judgment of 0 as "new". Figure 4 displays the proportion of JOFs that were greater than 1 (JOF > 0) as a function of presentation frequency and item type. The top panel shows the data from the voice instruction condition and the lower panel shows the data from the plural instruction condition. The corresponding $d'$ values for both conditions are provided in Table 1.

For subjects in the voice instruction condition, if a test word was in the same voice as at study (which would be the case for same and different-plural words), subjects typically followed the instructions and gave it a frequency judgment that was greater than zero. As the figure shows, there was a large proportion of JOF > 0 for same voice items.

The figure also shows a smaller proportion of JOF > 0 for items in which the voice changed across study and test episodes. The difference between same and different-voice items shows that voice information is in fact encoded in memory and can affect recognition performance. However, the response curves for the different-voice items remained flat, rather than decreasing in slope. This indicates that as presentation frequency increased, subjects' ability to reject similar test items as new did not improve. The same pattern is found in the plural instruction condition (see Figure 4, lower panel).
Figure 4. Proportions of judgments of frequency (JOFs) greater than zero are displayed as a function of presentation frequency and test item. The upper panel displays proportions for subjects in the voice instructions condition. The lower panel displays proportions for subjects in the plural instructions condition.
Table 1.

d' for Item Types and Test Condition in Experiment 1.

<table>
<thead>
<tr>
<th>Item Type and Test Condition</th>
<th>Frequency at Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>One</td>
</tr>
<tr>
<td>Same VOICE</td>
<td>1.20</td>
</tr>
<tr>
<td>Same PLURAL</td>
<td>.48</td>
</tr>
<tr>
<td>Different plural VOICE</td>
<td>1.30</td>
</tr>
<tr>
<td>Different plural PLURAL</td>
<td>.41</td>
</tr>
<tr>
<td>Different voice VOICE</td>
<td>.40</td>
</tr>
<tr>
<td>Different voice PLURAL</td>
<td>1.08</td>
</tr>
<tr>
<td>Both different VOICE</td>
<td>.97</td>
</tr>
<tr>
<td>Both different PLURAL</td>
<td>.41</td>
</tr>
</tbody>
</table>

Note: d' was derived from individual subject performance. Hits of 1.00 were truncated to .95; false alarms (FA's) of 0.00 were truncated to .05. FA's are based on "new" filler items; FA rate for the VOICE instruction condition = .28. FA rate for PLURAL instruction condition = .38.

The statistical analysis confirmed these patterns. Separate ANOVAs on the voice group and plural group data were conducted, with the within subject factors of frequency and item type. For subjects in the voice group, there was a main effect of presentation frequency [F(2, 38) = 4.61, p < .01]. However, post-hoc tests comparing performance across frequency conditions for each item type revealed that the effect of frequency arose entirely from differences between frequency = 1 and frequency = 5 [t(79) = 2.88, p < .0003]). This indicates that there was no additional learning of voice or plural information beyond frequency = 5. There was also a main effect of item type [F(3, 57) = 17.34, p < .0001]. In general, items that preserved voice across study and test were better recognized than items that changed voice. Same items were recognized better than different-voice [t(59) = 5.24, p < .0001] and both-different items [t(59) = 5.29, p < .0001]. Different-plural items were recognized better than different-voice [t(59) = 5.01, p < .0001] and both-different items [t(59) = 4.09, p < .0001]. There was no interaction between frequency and item type in either group.

For subjects in the plural group, there was an effect of frequency [F(2, 80) = 9.42, p < .0004]. Post-hoc tests confirmed that there was no additional improvement in subjects learning of the same perceptual form-different perceptual form distinction beyond frequency = 5. The effect of item type was significant [F(3, 63) = 11.02, p < .0001]. Recognition was more accurate on same items than on different-plural items [t(65) = 3.40, p < .001] or both-different items [t(65) = 2.75, p < .0001]. Different-voice items were recognized more accurately than same items [t(65) = 2.34, p < .02], different-plural items [t(65) = 5.22, p < .0001], or both-different items [t(65) = 5.49, p < .0001]. There were no other effects.
In sum, the word recognition analysis revealed that when frequency judgments were recoded as hits, the benefits of a perceptual match across study and test were evident in both groups, and directly related to the test instructions: Subjects who were told to focus on voice information showed an effect of voice match; subjects who were told to focus on plurality information showed an effect of plurality match. In both groups, subjects largely ignored the irrelevant stimulus dimension when making their same form-different form discriminations. The fact that the encoding instructions were identical across both conditions suggests that the use of perceptual information during explicit retrieval is not automatic, but is moderated by the task demands and retrieval intentions of the subject.

As a whole, the results obtained in this experiment replicate and extend the earlier findings of Hintzman et al. (1992). The hallmarks of the “registration without learning” effect are an increased in the frequency judgments for targets and similar distractors, coupled with attenuated learning of the feature necessary to discriminate targets from similar distractors. In the present experiment, both patterns were present. Subjects knowledge that a word occurred continued to increase as presentation frequency increased, whereas their ability to reject similar distractors did not increase proportionally.

We also found no difference between voice learning and plurality. This finding was not expected. We predicted voice to be preserved to a greater extent than plurality because voice information is acoustically and perceptually richer and more complex, and because voices and words are processed in integral, parallel-contingent fashion. This latter fact suggested the possibility that the word information would “carry” voice information, and lead to proportional learning of both dimensions. The data showed instead that voice learning, like plural learning, stabilized after the first few repetitions. In Experiment 2, we further evaluate the retention of voice and plural information using several new procedures.

Although there is ample evidence that many acoustic-phonetic details are represented in memory along with more abstract phonological and semantic information (see Pisoni, 1993 for a review), this evidence is derived from experiments that examine the effects of stimulus variability over a single word repetition. It is important to point out that the presence of perceptual specificity effects in memory does not necessarily mean that perceptual information is represented completely and without error. To the contrary, the research findings described in Experiment 1 suggest that it is likely that such knowledge is incomplete and fragmentary. The results also show that the extent to which episodic information about the perceptual form of a word is encoded varies across instances of an item. Indexical information about a talker’s voice, or acoustic-phonetic information about plurality, is most likely to be encoded and stored in memory when a word is novel, rather than highly familiar, and tends to largely remain constant across repetitions.

**Experiment 2**

Our first experiment revealed a dissociation between knowing that a word occurred and knowing details about the perceptual form of the word. A question to ask, then, is whether the dissociation occur because registration is truly without learning, or if instead, the results are merely a consequence of the retrieval task. In particular, the frequency judgment test required subjects to make two kinds of decisions simultaneously: Judge the familiarity of the item, and recollect perceptual information. In this way, the task incorporates both automatic and intentional retrieval processes. These processes can operate synergistically, or they can act in opposition, such as when a subject rejects a familiar distractor because they recall information about the target item (Jacoby, 1991). It is possible that greater learning of perceptual information would be apparent when voice or plurality discrimination is assessed separately from memory for frequency.
The first objective of Experiment 2 was to separate the effects of knowing a word occurred (regardless of its form) from the effects of knowing format-specific details about the word. Subjects were allowed to ignore study-to-test changes in the perceptual form of a test item when making their frequency judgments. We assessed memory for voice and plural information independently by asking subjects if the test word was presented in the same or different voice or plurality as the study item. The second objective of Experiment 2 was to compare explicit memory for voice and plurality information. Recall that a null effect of item type on frequency judgments was found among subjects in the plural instruction group (e.g., Figure 3). That measure also showed no effect of plural change on frequency judgments among subjects in the voice instruction group. Experiment 2 assessed whether these effects were the result of differences in memory for voice and plurality information.

To ensure that differences across experiments were due solely to retrieval processes and not encoding processes, the study instructions used in Experiment 2 were identical to those used in Experiment 1. Participants simply listened to the words on the study list with the expectation that they would be tested on the presentation frequencies of the words. After the study task, the participants answered four questions. The first two questions were confidence ratings and frequency estimates. Each provides a converging measure of subjects’ ability to retrieve word type information. Subjects were told to give a positive confidence rating or frequency judgment even if the voice or plurality of the item had changed across study and test. The remaining two questions assessed subjects’ ability to retrieve token-specific information. Subjects were asked to decide if the test word was in the same voice and the same plurality as the studied word. This allowed us to compare voice and plurality information by determining the extent to which each source is encoded and represented in memory. We hoped that the combination of these four questions would allow us to separate the effects that frequency and similarity have on the retrieval of type and token information.

Method

Participants

The participants were 20 Indiana University students who volunteered for the experiment in exchange for course credit. All were native speakers of English and reported no history of speech or hearing disorders. The participants were tested in groups of four or fewer. Four subjects were excluded from the analysis for failing to follow directions.

Apparatus

The stimulus materials were presented using an IBM-compatible personal computer. The words were presented binaurally to subjects at 75 dB SPL over matched and calibrated stereophonic headphones (BeyerDynamic DT-100).

Stimulus Materials

The stimulus materials were identical to those used in Experiment 1. In the study phase, participants heard a list of 205 successively presented words. The list contained 24 singular and plural target items produced by either a male or female talker. Target items were presented once or repeated five or ten times in the list. The study list also contained 77 nonrepeating fillers. In the test phase, participants heard a list of 48 successively presented words. Half were new items, and half were targets. There were four types of target items: “same,” “different-voice,” “different-plural,” and “both-different.”
Procedure

The study instructions were identical to those used in Experiment 1. Subjects were told to listen carefully to each item because they would be tested on the presentation frequencies of the words. Following the study phase, a test form was distributed. Participants gave four responses for each test word, taking as much time as necessary for each.

Questions 1 and 2 consisted of a confidence rating and a frequency judgment. Both questions assessed subjects’ ability to retrieve word type information. Question 1 was a confidence rating on a -3 to +3 scale of the certainty that some form or version of the test word had been heard during the study phase. The following example was provided:

_There are some words on the test list that were also on the study list in a different form. That is, the word may now be in a different plural form, or in a different voice, but was nevertheless present on the study list in some form. For example, you may hear “CAT” spoken in the female voice on the test, and remember that “CAT” was present in the study list, only it was spoken by a male voice. In this case, you should circle a number on the positive side of the response scale (+1, +2 or +3), depending on how certain you are that you heard some form of CAT on the study list. If you don’t remember the word, you should circle a number on the negative side of the response scale (-1, -2 or -3), depending on how certain you are that you did not hear some form of the word on the study list._

Question 2 was a frequency judgment in which subjects were asked to estimate the frequency of occurrence for each test word type. Subjects were instructed to circle a number from 1 to 15 if any version of the word was studied, and to circle zero if the word type was not studied.

Questions 3 and 4 were source judgments which assessed subjects’ ability to retrieve token-specific information. Specifically, subjects were asked to decide if the test word was in the same voice and the same plurality as the studied word. Subjects made their voice judgments by checking either “male voice” or “female voice” on a response sheet. Similarly, they made their plurality judgments by checking either “singular” or “plural” on a response sheet.

Subjects were instructed to answer each question, regardless of their responses to the previous questions, and to guess if they were unsure. The order of questions 3 and 4 were counterbalanced, as was the order of the response alternatives within each question. Subjects were carefully instructed as to how to answer each of the four test questions, and were given several examples to ensure that they understood the test tasks. The entire experimental session lasted less than 30 minutes, after which the subjects were debriefed.

Results

Frequency Judgments

Mean Frequency Judgments. Because the histograms revealed a unimodal frequency distribution for each item type, the analysis of the mean frequency estimates included all four item conditions. Subjects showed sensitivity to differences in the study presentation frequencies \( F(2, 38) = 39.45, p < .0001 \). Mean JOFs across the three frequency levels were 1.5, 5.7, and 9.1, which reveals a close relationship between
judged frequency and actual frequency. However, there was no effect of item type, and it did not interact with frequency.

Nonzero JOFs. As in Experiment 1, an analysis was conducted on the nonzero JOFs in order to determine if frequency judgments made to similar items were lower than judgments made to same items. This analysis excluded four subjects who failed to contribute at least one nonzero judgment for each item type at each frequency. The analysis of the nonzero JOFs (see Figure 5) revealed a highly significant effect of frequency \[ F(2, 22) = 70.53, p < .0001 \]. There were no other effects.

Insert Figure 5 about here

word recognition. Frequency judgments were collapsed into a binary recognition measure. Recognition of a word type as “old” was defined as a positive frequency judgment. Figure 6 shows the proportion of correct recognitions as a function of presentation frequency and item type (see Table 2 for the corresponding \( d' \) values). The figure shows that word type recognition improved with presentation frequency \[ F(2, 30) = 40.20, p < .0001 \]. In fact, performance is close to ceiling after the fifth presentation. Such high performance levels were unanticipated given that the methodology of this experiment is very similar to Experiment 1, where ceiling effects were not found.

Another difference between the present experiment and Experiment 1 is that there was no effect of item type in the present experiment. This indicates that recognition memory for word types was not affected by the physical similarity between the study tokens and test tokens in this task, and confirms that subjects were following directions.

Insert Figure 6 about here

Confidence Judgments

Mean Frequency Judgments. Mean confidence ratings are displayed in Figure 7 as a function of presentation frequency and item type. An analysis was conducted on the confidence judgments in order to determine if confidence increased with study repetitions and if different form items were given lower confidence ratings than same form items. There was, of course, a highly significant effect of frequency \[ F(2, 22) = 70.53, p < .0001 \], indicating that subjects’ confidence that a word type was “old” increased dramatically with presentation frequency. The effect of item type was not significant, which indicates that recognition of same and similar items was accomplished with equal confidence. Frequency and item type did not interact.

Insert Figure 7 about here
Figure 5. Mean judgments of frequency (JOFs) greater than zero are displayed as a function of presentation frequency and test item.
Figure 6. Proportions of judgments of frequency (JOFs) greater than zero are displayed as a function of presentation frequency and test item type.
Figure 7. Mean confidence ratings are displayed as a function of presentation frequency and test item type.
Table 2.

d’ for Item Types and Test Condition in Experiment 2.

<table>
<thead>
<tr>
<th>Item Type and Test Condition</th>
<th>One</th>
<th>Five</th>
<th>Ten</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same JOF</td>
<td>1.11</td>
<td>1.52</td>
<td>1.93</td>
</tr>
<tr>
<td>Same CONF</td>
<td>1.26</td>
<td>1.67</td>
<td>2.08</td>
</tr>
<tr>
<td>Different plural JOF</td>
<td>.90</td>
<td>1.21</td>
<td>1.93</td>
</tr>
<tr>
<td>Different plural CONF</td>
<td>1.05</td>
<td>1.34</td>
<td>2.08</td>
</tr>
<tr>
<td>Different voice JOF</td>
<td>1.10</td>
<td>1.10</td>
<td>1.93</td>
</tr>
<tr>
<td>Different voice CONF</td>
<td>.85</td>
<td>1.98</td>
<td>2.08</td>
</tr>
<tr>
<td>Both different JOF</td>
<td>.90</td>
<td>1.82</td>
<td>1.72</td>
</tr>
<tr>
<td>Both different CONF</td>
<td>1.05</td>
<td>1.98</td>
<td>1.88</td>
</tr>
</tbody>
</table>

*Note: “CONF” = recognition data from the confidence judgments. “JOE” = recognition data from the judgments of frequency. d’ was derived from individual subject performance. Hits of 1.00 were truncated to .95; false alarms (FA’s) of 0.00 were truncated to .05. FA’s are based on “new” filler items. FA rate for JOE = .40. FA rate for CONF = .34.*

**Word Recognition.** Confidence judgments were also reduced to an old/new recognition measure. Positive confidence ratings (1, 2, or 3) were recoded as “old” responses, whereas negative confidence ratings (-1, -2, or -3) were recoded as “new” responses. The proportion of positive confidence responses increased systematically with presentation frequency [F(2, 30) = 45.33, p < .0001] but did not differ among the test items. There were no other effects.

So far, the data from several measures convey the same story: The more often a given word was encountered at study, the higher the frequency estimate was for that item at test. Similarly, frequently encountered words were also more likely to be recognized, and the recognition judgments were accompanied by greater certainty that the item was old. Accordingly, these findings reflect the expected effect of study repetition on explicit memory performance.

The recognition data derived from the confidence judgments were nearly identical to the recognition data derived from the frequency judgments. Of course, differences in the two measures are partly obscured by the very high performance levels at frequency = 10. Nevertheless, subjects appear to have used the two response scales similarly, despite the fact that the frequency scale has twice the number of response categories (Proctor, 1977). Although the high performance levels in the present experiment do not allow us to definitively conclude that the same processing mechanism underlies judgments of presentation frequency and recognition memory, our results are consistent with many other experiments that do subscribe to this view (e.g., Harris et al., 1980; Hintzman, 1988).
In addition, word type recognition accuracy and confidence were the same across the four item types. Presumably, token-specific information is irrelevant to a task that requires the retrieval of word type information. That is, performance was unaffected by study-to-test changes in voice and plurality because subjects were asked to ignore such changes and responded accordingly. This led to a large number of positive IOF's accompanied by high levels of confidence that the item types were present on the study list.

Source Judgments

To analyze the explicit voice and plural source judgments, the proportion of correct same and different responses were calculated for trials in which the word was correctly recognized. Figure 8 displays the proportion of correct voice and plural judgments as a function of frequency. The figure shows that explicit memory for voice and plural information was overall above chance and slightly increased with study frequency.

Two statistical analyses were performed on the source judgments. One analysis conditioned the judgments on correct recognition from the confidence ratings, whereas a second analysis conditioned the judgments on correct recognition from the frequency judgment task. Since both analyses led to the same conclusions, only the statistical analysis based on the confidence ratings is presented here.

--------------------
Insert Figure 8 about here
--------------------

An ANOVA with the factors source judgment (voice or plural) and presentation frequency (1, 5, 10) showed a significant effect only for frequency [F(2, 60) = 12.67, p < .0001]. Accuracy was the same for the plural and voice source judgments (overall proportion correct was .63 for both judgments). Planned t-tests comparing voice recognition at each frequency level showed that overall voice recognition performance was significantly different from chance in all three frequency conditions ([t(15) = 2.68, p < .01] for frequency = 1; [t(15) = 3.87, p < .001] for frequency = 5; [t(15) = 6.54, p < .0001] for frequency = 10). In contrast, performance in discriminating same and different-plural items across study and test exceeded chance only at frequency =10, [t(15) = 9.46, p < .0001].

The source judgments directly assessed the encoding and retrieval of token-specific perceptual information. The memorability of voice and plurality information was equal, and source judgment accuracy showed only a small improvement even after 10 repetitions of a word-voice/plurality pair. The fact that there was an increase (albeit a small one) in source judgment accuracy as a function of frequency shows that registration is not entirely without learning. However, the amount of learning is modest, and certainly not proportional to word learning. Taken together, the explicit source judgments from this experiment and the discrimination data from Experiment 1 lead to the same conclusion: The enhanced ability to recognize a word that has been presented many times does not guarantee that the perceptual information is represented with any more detail than a word presented only once.

General Discussion

The present research sought to extend the "registration without learning effect" (Hintzman et al., 1992) to a new domain, for a new and more complex feature and with several new procedures. To accomplish this, we had subjects listen to a list of words spoken by a male or female talker, and presented
Figure 8. Proportion of correct voice and plural explicit source judgments are displayed as a function of presentation frequency.
in singular or plural form. Target items were presented various times in the study list. In Experiment 1, subjects estimated the frequency with which a word occurred at study, excluding items that were in a different voice from the study item (voice instructions group) or in a different pluralization from the study item (plural instructions group). In Experiment 2, the test task required subjects to estimate the frequency with which a word type occurred at study, ignoring differences in voice or plurality. Subjects also provided confidence ratings and made explicit voice and plurality judgments about the test words. In both experiments, the test list contained items that differed in their physical similarity to the study items by one or two dimensions (voice, plurality or both). Both experiments measured the effects of familiarity and similarity on subjects' memory for frequency and their ability to make fine-grained distinctions among similar items.

The major findings were as follows: (1) Frequency judgments, recognition accuracy, and confidence ratings for word tokens and word types were higher for words repeated many times during study as compared to words presented just once. (2) In contrast, repeating a word many times during study did not improve subjects' memory for the perceptual details associated with the word. (3) The asymptotic learning function for the feature that discriminates same form items from similar distractors was not specific to the frequency judgment task used in Experiment 1. Explicit source judgments from Experiment 2 also revealed attenuated learning. (4) The complexity of the stimulus feature did not affect learning. (5) The effect of the similarity between study and test items was dependent on instructions at retrieval. Together, these findings demonstrate that the acquisition of instance-specific perceptual knowledge varies across word repetitions.

Repetition and Perceptual Processing

The present results are consistent with the idea that the gains in word recognizability that occur when an item is repeated are attributable to the listener's failure to engage in bottom-up processing of surface information. The idea here is that perceptual processing is suppressed for inputs that match expectations and accentuated for inputs that are unexpected or novel (Hintzman et al., 1992; Johnston & Hawley, 1994). For example, in the present experiments, when listeners encountered a word that was presented earlier in the study list, they presumably realized (either consciously or unconsciously) that they had heard the word earlier in the list, and consequently, didn't allocate processing resources reaffirming details about a word they already knew. The perceived familiarity of the item served to disguised the fact that their knowledge about the structure of the word was incomplete. Consequently, the subject later had difficulty distinguishing between the studied word and a similar distractor because attention to perceptual features had largely been truncated after the first few repetitions (DiGirolamo & Hintzman, 1997; Tulving & Kroll, 1997).

In the present study, variation in voice or plurality occurred across study and test phases of the experiment. It is worth noting that DiGirolamo and Hintzman (1997) have recently obtained similar effects when the attributes of a repeated object are varied within the study phase. In their study, DiGirolamo and Hintzman varied the perceptual form of an object either early or late in the study list. For some objects, the orientation of the first presentation differed from presentations 2-5, whereas for other objects, the fifth presentation differed from presentations 1-4. Subjects were then given a forced choice discrimination test which assessed their ability to remember whether an item occurred in one orientation or both orientations. DiGirolamo and Hintzman found that subjects reported seeing both orientations more often if orientation changed early in the list rather than later in the list. They argue that their findings are not the result of a simple primacy effect, but stem from the fact that information processing on early repetitions of an item was qualitatively different from information processing on later repetitions of the same item.
An alternative perspective from which to view the effects of repetition is to suppose that repetition serves to build a "unitized" response code (Feustel, Shiffrin & Salasoo, 1983; Shiffrin & Lightfoot, 1997). Unitization refers to a process that integrates co-occurring parts of an event into a single functional unit. Unitization can be conceived of as an abstraction process whereby features that are not relevant to the task are less likely to become integrated. Unitization allows a complex stimulus to be identified based on the most relevant or diagnostic features. However, once an item is unitized, it becomes more difficult to detect incidental or features on subsequent repetitions. Thus, unitization may play a role in accounting for the registration without learning phenomena.

The unitization view also suggests that perceptual information, such as information about a talker's voice, may continue to be encoded across repetitions if it is perceived as being diagnostic of a word's identity. For example, imagine if each of the 24 target words used in the present experiment were produced by one of 24 different voices. If a specific voice reliably co-occurred with just one word within a list, listeners may detect the relevance of the voice as a retrieval cue. Consequently, voice information may become integrated with the word in such a way that promotes learning of voice information across repetitions. We are currently exploring this possibility.

Of course, these accounts of registration without learning rests entirely on data derived from explicit memory tasks. Explicit tasks are largely driven by the match between conceptual or contextual information at study and test and tend to be unaffected by perceptual information. Implicit tasks tap different kinds of cognitive operations than do explicit tasks and thus, recover different information (for a review, see Roediger & McDermott, 1993; Schacter, 1987). In particular, implicit test performance tends to be facilitated by perceptual consistency across study and test. It is possible that perceptual information associated with a familiar item is represented veridically, but that the experimental task used in the present experiments provided only a limited window on the underlying information. Further research is needed in order to determine if the effects reported here and those reported by Hintzman and colleagues are less about "registration without learning" and more about "learning without retrieval".

Retrieval Models

The results of the present experiments underscore the importance of a retrieval model based on two processes (familiarity and recall). The need for two processes is necessary because discriminative responding of the kind required by the frequency task used in Experiment 1 could not be mediated by a unidimensional familiarity signal. A second mechanism is needed that supports the recollection of instance-specific details from a prior episode, and allows one to know that a highly familiar item was not actually on the study list.

Further empirical support for the distinction between familiarity and recall comes from a recent experiment that used a response-signal paradigm (Hintzman & Caulton, 1997; cf. Hintzman & Curran, 1994; Mulligan & Hirshman, 1995). This method is an elegant way to explore how the retrieval of a word unfolds over time. Hintzman and Caulton found that the retrieval dynamics of familiarity and recall can be differentiated by their speed-accuracy retrieval patterns and by their sensitivity to the effects of repetition and similarity. Although a few qualitative models (Johnston, Jacoby & Dark, 1975; Mandler, 1980) and formal quantitative models of memory include familiarity and recall processes (e.g., Minerva 2, Hintzman, 1988; REM, Shiffrin & Steyvers, 1997; SAM, Gillund & Shiffrin, 1984), the extent to which recall contributes to recognition memory and frequency judgments is currently not well specified. This timely and interesting topic is worthy of further research and simulation.
Encoding of Talker Information

The findings from Experiments 1 and 2 have implications for the manner in which we characterize the processing or "normalization" of voices. Recent research demonstrates that familiarizing subjects with the voices in which speech stimuli are presented improves word recall (Lightfoot, 1989) and speech intelligibility (Nygaard, Sommers & Pisoni, 1994). For example, listeners who learn to identify a set of talkers from sentence length utterances are able to transcribe new sentences presented in white noise more accurately than subjects who are unfamiliar with the talkers (Nygaard & Pisoni, 1995). These experiments clearly show that familiarity with a voice has a relationship to the ongoing analysis of speech. Nygaard & Pisoni (1995) argue that the increase in sensitivity to talker-specific information brought about by training automatically increases sensitivity to the linguistic information in the signal.

Their results are grounded in the proceduralist model described by Kolers (1979; Kolers & Roediger, 1984), in which the pattern-analyzing operations or "procedures" involved in encoding an item serve as the basis for the memory representation of the event. These procedures are invoked to supplement the analysis of a repeated event. The fact that pattern-analyzing operations do not have to be constructed anew means that familiar items, or new items presented in a familiar script or voice, can be identified more efficiently. An important, albeit implicit, assumption that underlies this view is that the effect of practice is to change how information is processed, rather than which information is processed. Listeners attend to, encode, and process the same amount of information regardless of their familiarity with the stimulus item. Accordingly, the gain in the efficiency of processing does not imply that the processing of perceptual information is in any way curtailed or attenuated.

However, the experiments in this report, as well as those by Hintzman and his colleagues raise the possibility that the improvement in spoken language processing resulted not from better use of voice information, but from less use of voice information. That is, the effect of training and practice may have been to induce a systematic reduction in the amount of perceptual information processed, rather than to change the efficiency with which perceptual analysis was performed.

The notion that perceptual details are less memorable if they occur on later repetitions may also shed light on the interaction between talker variability and list position (Martin, Mulvennix, Pisoni & Summers, 1989), on differences in the encoding and retrieval of single-talker word lists as compared to multiple-talker lists (Mulvennix, Pisoni & Martin, 1989), and on the mechanisms assumed to underlie speaker normalization (see Pisoni, 1996). We claim that what is commonly referred to as "speaker normalization" may simply be a general epiphenomenon of redistributed perceptual processing or learned inattention that commonly occurs in response to many objects and events, rather than a specific linguistic process dedicated to the recovery of canonical phonetic units from speech.

Repeated exposure may also affect the influence of speaker information on other tasks. For instance, Walker, Bruce & O'Malley, (1995) found that subjects who are familiar with the face producing an audiovisual syllable are less susceptible to influences of facial speech information on auditory speech perception. That is, familiar faces produce weaker McGurk effects. The relative ease with which faces become familiar as compared to voices may also play a role in the weak effects of face variation relative to voice variation on spoken word recognition (Sheffert & Fowler, 1995). Further research is needed to understand whether familiarity directly affects the use of talker-specific information during bimodal speech processing.
Conclusion

The experiments reported here examined the perception and encoding of instance-specific perceptual information as a function of the amount of prior exposure a listener has with a particular item. The results replicate and extend the prior findings of Hintzman, et al. (1992) by revealing a dissociation between knowing that a spoken word occurred and knowing perceptual details about the word. The collective import of the studies is to show that "registration without learning" occurs for auditorily presented words, for simple and complex stimulus dimensions, and across various experimental settings. The findings add to a growing literature on the factors that are important in determining what knowledge a subject encodes during spoken word recognition, and help to delimit theoretical interpretations of memory.

References


