Trial-to-trial carry-over of item- and relational-information in auditory short-term memory

Kristina M. Visscher[a], **Michael J. Kahana**[b], and **Robert Sekuler**[a]

[a]Volen Center for Complex Systems, Brandeis University, MS 013, Waltham, MA, USA 02454

[b]Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA 19104

Abstract

Using a short-term recognition memory task we evaluated the carry-over across trials of two types of auditory information: the characteristics of individual study sounds (item information), and the relationships between the study sounds (relational information). On each trial, subjects heard two successive broadband study sounds and then decided whether a subsequently presented probe sound had been in the study set. On some trials, the probe item's similarity to stimuli presented on the preceding trial was manipulated. This item information interfered with recognition, increasing false alarms from 0.4% to 4.4%. Moreover, the interference was tuned so that only stimuli very similar to each other interfered. On other trials, the relationship among stimuli was manipulated in order to alter the criterion subjects used in making recognition judgments. The effect of this manipulation was confined to the very trial on which the criterion change was generated, and did not affect the subsequent trial. These results demonstrate the existence of a sharply-tuned carry-over of auditory item information, but no carry-over of relational information.

Introduction

In many different settings previously-acquired information interferes with memory for information that is acquired subsequently (Jacoby, Debner, & Hay, 2001; Postle, Brush, & Nick, 2004; Underwood, 1957). This phenomenon has recently attracted renewed attention, due in part to the realization that the inability to suppress irrelevant information, as from a previous trial, can significantly limit key cognitive functions (May, Hasher, & Kane, 1999; Lustig, May, & Hasher, 2001; Jonides & Nee, 2006). Our purpose here was to examine and compare two different types of information that might disrupt auditory short-term memory.

One type of information that may carry over from trial to trial is *item information*, which arises from the characteristics of individual study stimuli. On each trial of a typical short-term memory experiment, subjects are exposed to information arising from the values of stimuli presented on that trial. As a result, item information might come from the identities of the words heard, or the shapes of objects seen. The carry-over of such item information has been the focus of many empirical and theoretical studies and is sometimes referred to as proactive interference (reviewed in Kahana, in press).

A second type of information that may carry over from trial to trial arises from the relationships among study stimuli. We refer to this as *relational information*, a designation meant to emphasize parallels to other work that examines how relationships between stimuli can affect memory (e.g. Hunt & Einstein, 1981; Jiang, Olson, & Chun, 2000). As the term implies,

Dr. Kristina Visscher, Volen Center for Complex Systems, Brandeis University, Waltham, Massachusetts, USA 02454, e-mail: visscher@brandeis.edu, fax: (781)736-2398.

relational information reflects the relationships among stimuli, such as the number and ordering of items to be remembered, the physical similarity of one study item to another, or their semantic relationships. To be clear, we use the term in a relatively broad sense. We do not, for example, mean it strictly as a synonym for associative information (Dyne, Humphreys, Bain, & Pike, 1990). Relational information does not represent information about individual items *per se*, but nevertheless can influence subjects' recognition responses (Kahana & Sekuler, 2002; Kahana, Zhou, Geller, & Sekuler, in press; Nosofsky & Kantner, 2006; Wright, 1998).

Knowledge of the extent to which different classes of information carry over or fail to carry over from trial to trial could illuminate the character and mechanisms of proactive interference. Researchers have distinguished at least two forms of carry-over, item-specific proactive interference, which depends upon the repetition of items, and item-nonspecific proactive interference, which reflects the accumulation of information, independent of item repetition (Postle et al., 2004). Interestingly, release from these two forms of carry-over seems to depend upon the same circuit in prefrontal cortex (Postle et al., 2004).

Relational information impacts performance on a given trial (e.g. Einstein & Hunt, 1980; Nosofsky & Kantner, 2006; Kahana & Sekuler, 2002), but does not involve the repetition of an item, nor the sheer accumulation of information over trials. It is an open question, therefore, whether relational information would exhibit carry-over that characterizes proactive interference. We asked whether memory of relational information operated like item-specific proactive interference, carrying over from one trial to influence recognition performance on the next. Additionally, by using metric auditory stimuli tailored to individual subjects' discrimination thresholds, we were able to carefully explore the sensitivity of interference from item information by determining how similar an item had to be to a probe in order to interfere.

Item Information

Much research on carry-over effects has focused on item information and is often termed proactive interference (e.g. Jonides & Nee, 2006). Typically, these experiments have used verbal stimuli, such as lists of words (e.g. D'Esposito, Postle, Jonides, & Smith, 1999) or other, readily named items, such as pictures of familiar objects (e.g. Smith, Leonard, Crane, & Milner, 1995). In such experiments, the strength of carry-over often reflects the semantic connections among the verbal items presented on successive trials (Wickens, 1972), but fine-grain tests of the similarity between stimuli required for carry-over have to our knowledge not been reported. Because responses in tests of short-term memory can be affected by information retrieved from long-term memory (Huttenlocher, Hedges, & Duncan, 1991; Olsson & Poom, 2005; Visscher, Viets, & Snyder, 2003), memory tasks using easily-categorized or nameable stimulus materials allow the recruitment of information from outside of short-term memory, and, compared to non-nameable materials, are likely to recruit different areas of the brain (e.g., Ikeda & Osaka, 2007). In order to minimize semantic influences on our results, we examined carry-over effects in human short-term memory for auditory stimuli that were difficult to categorize or name in a consistent fashion.

A few studies have observed proactive interference with single or multiple auditory stimuli, such as tones, which do not demand explicit verbal mediation (e.g., Ruusuvirta, 2000; Ruusuvirta, Wikgren, & Astikainen, 2006; Wright, 1999). We extended these results by using complex auditory stimuli that make it possible to quantify and manipulate the degree of similarity between any two stimuli, and allow measurement of how similar items must be in order to interfere.

Relational Information

In addition to examining carry-over of item information, we also examined carry-over of relational information. By using the same dataset to examine both effects, the trial-to-trial influence of relational information is contrasted to that of item information.

Here we examined one particular form of relational information: the similarity of study stimuli to one another, referred to as *inter-item homogeneity*. This variable has been shown to exert a strong influence on subjects' recognition responses (Kahana & Sekuler, 2002; Kahana et al., in press; Nosofsky & Kantner, 2006; Visscher, Kaplan, Kahana, & Sekuler, 2007; Yotsumoto, Kahana, Wilson, & Sekuler, in press). Specifically, when study items are homogeneous, subjects are less likely to judge that a given probe matches a study stimulus (Kahana & Sekuler, 2002; Kahana et al., in press; Nosofsky & Kantner, 2006; Yotsumoto et al., in press). This effect has been made explicit by a computational model, called the Noisy Exemplar Model (NEMo). Nosofsky and Kantner (2006) argued that study-item homogeneity influences the subject's decision criterion, with stricter decision criteria being associated with higher levels of homogeneity between study items. Although the design of our experiment was guided by NEMo, analysis was carried out independently of the model.

The influence of response criterion has been the target of research in the sensory and memory literatures (e.g. Ball & Sekuler, 1980; Cho et al., 2002; Gorea, Caetta, & Sagi, 2005; Gorea & Sagi, 2000; Jones, Cho, Nystrom, Cohen, & Braver, 2002). In both domains, theoretical accounts of performance typically assume that a subject adopts some criterion against which stimulus item information is compared. For example, in the case of memory, such a comparison can form the basis of a recognition response, such as a judgment of a test stimulus as "old" or "new." Regardless of how a criterion value is generated, that criterion could persist for some trials, either because the relevant conditions are unchanging or because of inertia in the criterion-setting process; alternatively, the criterion might be reset anew on each trial, tracking trial-by-trial changes in task demands. The mutability of the subject's criterion is a matter of debate in the memory literature (Heit, Brockdorff, & Lamberts, 2003; Singer & Wixted, 2006; Stretch & Wixted, 1998), as it is in sensory psychophysics literature. Although some sensory studies suggest that a subject's criterion is highly mutable, adjusting to conditions from one trial to the next (Petzold & Haubensak, 2004; Treisman & Williams, 1984), other studies suggest that subjects adopt and hold a single, stable criterion across a group of trials (Gorea & Sagi, 2000). Additionally, at least one sensory study has revealed a remarkable flexibility in criterion setting, with ubjects able to develop and hold multiple, distinct criteria, drawing on any one as cued for a particular trial (Morgan, Watamaniuk, & McKee, 2000). Of course, evidence for or against the mutability of criteria in sensory tasks such as detection or discrimination does not constrain the mutability of criteria in memory tasks.

In order to control subjects' recognition criterion, we manipulated inter-item homogeneity, a variable linked to subjects' recognition criterion (Nosofsky & Kantner, 2006). The metric properties of our stimuli allowed us to manipulate the inter-item homogeneity on any given trial, and thus directly test whether this relational information's effect is maintained across trials. The requisite precise control over stimuli was made possible, in part, because our stimuli could be adapted for individual subjects, compensating for systematic differences in subjects' powers of discrimination. Thus we were able to control the perceptual similarity between individual stimuli on each trial, as well as the relationships between stimuli on successive trials. We exploited this stimulus control to examine how item information and relational information on trial *n* affected performance on trial *n*+1.

The Noisy Exemplar Model

The details of our experiment were guided by the Noisy Exemplar Model (NEMo), which is described fully elsewhere (Kahana & Sekuler, 2002; Visscher et al., 2007). NEMo shares with several related models the idea that summed probe-item similarity is a basis of recognition responses (e.g., Brockdorff & Lamberts, 2000; Clark & Gronlund, 1996; Humphreys, Pike, Bain, & Tehan, 1989; Lamberts, Brockdorff, & Heit, 2003; Nosofsky, 1991), but differs from other models in asserting that recognition responses are influenced also by the homogeneity of the study items. Taking account of inter-item homogeneity significantly improves the model's ability to predict subjects' performance for various classes of stimuli, including complex sounds (Visscher et al., 2007), visual gratings (Kahana & Sekuler, 2002; Kahana et al., in press), faces (Yotsumoto et al., in press), and colors (Nosofsky & Kantner, 2006).

On each trial in the experiments presented here, a pair of study items, s_1 and s_2, was presented, followed by a probe. The subject had to indicate whether the probe matched one of the stimuli in the study set. As mentioned above, NEMo assumes that recognition judgments are based on the similarity between the probe and all the list items (termed "summed similarity"), and the homogeneity of the items in the study set. Figure 1 illustrates the characteristics of a group of stimuli presented on a typical trial. The lengths of the green solid lines represent the probe-item similarity. The length of the blue dashed line in Figure 1 represents the inter-item homogeneity on that trial.

Although the model predicts performance on a current trial, NEMo is mute as to whether the inter-item homogeneity on one trial affects performance on the next trial. The fact that inter-item homogeneity can be easily manipulated on a trial-to-trial basis suggests a way to examine the question. In addition, the model's tuning with probe-item similarity suggests the idea that the effect of a stimulus from a previous trial might also be tuned.

Overview

We ask two main questions: (1) to what extent is item information maintained across trials, and (2) to what extent is relational information maintained across trials? The type of relational information we examine is inter-item homogeneity. To foreshadow, we found a distinction between the endurance of item and relational information: item information is maintained across trials, producing a modest, but reliable change in response that is sensitive to the degree of similarity between stimuli on successive trials. However, relational information seems not to be maintained across trials, but is modulated on a trial-to-trial basis. This distinction implies that the two effects stem from different mechanisms.

Materials and Methods

Moving Ripple Stimuli

In order to examine how similarity relationships among stimuli affect subjects' responses, we used moving ripple sounds, stimuli that could be continuously varied, and whose inter-item similarities could be measured. Moving ripple sounds are broadband sounds that vary sinusoidally in both time (with a period w cycles per second) and in frequency content (with a period Ω cycles per octave). In Figure 2 the horizontal axis represents time, and the vertical axis represents the frequency content of two sample stimuli. These stimuli were generated by superimposing sounds at many frequencies whose loudness at any time (t), and for any frequency (f) is defined by

$$s(t, x) = D_0 + D \cdot \cos[\,2\pi(wt + \Omega x) + \psi\,] \tag{1}$$

where x=$\log_2(f/f_0)$ and f_0 is the lowest allowed frequency. ψ is the phase of the ripple, and D is modulation depth. D_0 is the base loudness, which was set to 1.0. In order to simplify the stimulus space, only one parameter (w) was varied among the stimuli. Other parameters took fixed values: $\Omega=1$, $\psi=0$, $D=0.9$, and f_0=200 Hz. Frequencies ranged over 3 octaves above f_0, that is, from 200 to 1600 Hz. Each stimulus contained 20 logarithmically-spaced frequencies per octave. Each stimulus has a spectral profile that drifts in time, so that different frequencies are at their peaks at different times. For each stimulus, duration was set to one second. Example stimuli can be found at

http://people.brandeis.edu/~sekuler/rippleSoundFiles/movingRippleSounds.html. The advantages of this particular kind of stimulus for studying memory were described by Visscher et al. (2007), who also showed that short term memory for these stimuli exhibits strong parallels to short term memory for visual stimuli such as oriented sinusoidal gratings. An additional benefit for studying ripple sounds is that they share similarities to speech sounds (see Shamma, 2001). For example, their frequency bands modulate in time. Thus, findings using these ripple sounds are likely to generalize to speech sounds. The ripple sounds are difficult to verbalize, however, allowing examination of memory for language-like sounds independent of verbal labels.

Subjects

Subjects were between the ages of 18 and 30, and came from the student population of Brandeis University. At the outset, each potential subject underwent audiometric screening. Using a MAICO MA39 audiometer, thresholds were measured at 250, 500, 750, 1000 Subjects were between the ages of 18 and 30, and came from the student population of Brandeis University. At the outset, each potential subject underwent audiometric screening. Using a MAICO MA39 audiometer, thresholds were measured at 250, 500, 750, 2000, 3000, 4000 and 6000 Hz. Each subject had normal or above-normal hearing, that is, thresholds at or below 20 dB$_{HL}$ at each frequency.

Twelve subjects participated in eight sessions each, following an initial session in which just noticeable difference thresholds (JND) for the w parameter (cycles per second) were measured (see below), and 200 practice trials were performed. Experimental sessions, lasting about one hour each, comprised 586 trials. At the beginning of every session, each subject completed at least 30 practice trials that were excluded from data analysis. Successive sessions were separated by at least six hours, and for any subject, all sessions were completed within three weeks. Subjects participated for payment of $72, plus a performance-based bonus of up to $16.

The methods used in the study were approved by the Institutional Review Board of Brandeis University.

Apparatus and sound levels

Subjects listened to stimuli through Sennheiser Pro HD 280 headphones. Stimuli were generated by an Apple iMac computer, and Matlab, including its PsychToolbox add-on (Brainard, 1997). In order to characterize the stimulus intensity at the subject's eardrum, sound levels for this system were verified using a Knowles electronic mannequin for acoustic research. All stimuli were 79 dB$_{SPL}$, well above our subjects' hearing thresholds.

Stimulus presentation

On each trial, either one or two study items were presented, followed by a probe. The analyses presented here focus on the two-item lists. One-item trials were included in order to quantify pairwise perceived similarity, a parameter needed for the NEMo model fits presented in Visscher et al. (2007). Study lists were restricted to no more than two items in order to provide control of the variables required for the questions of experimental interest. The subject's task

was to judge whether the probe (*p*) matched any of the study items (s_1 or s_2). The response was indicated by a button press. During the presentation of study items, subjects fixated on a '+' in the center of a computer screen. Trials with one study item were intermixed among trials with two.

Each stimulus was one second in duration. When two study items were presented, they were separated by 0.25 seconds. The probe was presented 0.75 seconds after the final study item, and was accompanied by the presentation of a '?' on the computer screen. Subjects responded with a button press, indicating whether the probe matched a study item ("*Yes*") or not ("*No*"). Immediately after the subject's response, a distinctive tone provided feedback about response correctness. To increase motivation, after each trial subjects were shown their percent correct thus far in the session, and the difference between that value and their goal of at least 70% correct. Subjects were rewarded at the end of a session with a candy bar if their percent correct exceeded 70%. For every percentage point above that value, subjects received a $0.25 increment to their base payment.

Adjustment for discrimination threshold

Stimuli were adjusted to each subject's auditory discrimination threshold, thereby eliminating one source of potential individual differences, and making the memory task comparably difficult for all subjects (Zhou, Kahana, & Sekuler, 2004). In addition, the similarity among stimuli made it difficult for subjects to use naming or categorizing strategies in a consistent, reliable fashion. In a subject's first experimental session, pairs of stimuli were presented in succession on each trial, and the subject identifed which stimulus had the faster rate of modulation. Watson and Pelli (1983)'s QUEST algorithm found the difference in modulation rate ($\frac{\Delta w}{w}$) that just permitted correct identification of the more rapidly modulated stimulus on 70% of trials. This value was taken as the just noticeable difference (JND).

This JND value was then used to generate the stimuli that would be used in subsequent sessions to test that subject's recognition memory. The lowest value of *w* was $w_0 = 7$ Hz and successive values were given by $w_0(1 + JND)^n$, where n varies from 0 to 9. This generates stimuli that increment in one JND steps. In order to reduce the possibility that subjects could memorize the set of stimuli and assign verbal labels to them, we increased the number of stimuli to which subjects would be exposed. A second set of ten stimuli was created whose values lay midway between successive stimuli in the first set; taking on values $\frac{w_0 + w_1}{2}(1+\text{JND})^n$, where *n* again varies from 0 to 9. Trials whose test items were drawn from the first series were randomly intermixed with trials whose test items came from the second series. Thus the complete collection of possible stimuli comprised twenty sounds. Items in the stimulus pool were tightly packed along the dimension *w*, with separations of just 0.5 JNDs. This tight packing was meant to make absolute identification of individual stimuli difficult. On a particular trial, stimuli were drawn from only one series or the other, meaning that a trial's stimuli [s_1, s_2, p] were always an integer number of JNDs from each other.

Trials were self-paced, each initiated by the press of a key on a computer keyboard. On equal numbers of trials, the probe matched one of the study stimuli, or did not match either of the study stimuli. We designated matching trials as *Target* trials, and non-matching trials as *Lure* trials. *Target* and *Lure* trials were randomly intermixed during memory testing.

Experimental design

In order to assess the carry-over of item information and relational information from trial to trial, we manipulated the stimulus materials that were presented on successive trials. For each trial pair, the first trial (Trial A) comprised the set up trial, intended either to establish some

particular item information or to produce some particular relational information (operationally defined by inter-item homogeneity). Following each set up trial, the response on the next, test trial (Trial B) provided an index of the influence that had been established on the preceding trial. The details of the various conditions represented in the design are given below and in Table 1.

To minimize subjects' awareness of the complex regularities in the stimulus presentation schedule, trial pairs were randomly interleaved with trials of other types (a total of 320 carefully controlled pairs of trials within the 4680 trials presented to each subject). Trials listed as "Model testing" in Table 1 were analyzed in addressing a separate issue (Visscher et al., 2007). Trials on which just one study item was followed by a probe were randomly interleaved among all trials, and were used to gauge stimulus similarity.

Table 1 summarizes the effects that were targeted by each condition in our experiment. Note that the column headed "Condition" signifies the *relationship* among presented stimuli [s_1, s_2, p] rather than specific choices of stimuli, which varied from trial to trial. Many sets of stimuli consistent with the rules defining each condition were generated; examples were generated for both *Target* trials (on which the probe replicated a study item) and *Lure* trials (on which the probe did not replicate a study item). For example, the last row in the table refers to trials on which there was just one study item (s_1) and it did not match the probe (p). As each trial's stimuli were chosen from a set of 10 stimuli, there are 90 possible pairings of s_1 and p. For the conditions represented in the bottom two rows of the table, all possible pairings were used; other conditions used only a random subset of all possible pairings.

Carry-over of item information—In order to gauge carry-over of item information from one trial to the next, pairs of successive trials were constructed so that the stimuli from the first trial in the sequence (Trial A), $\mathbf{A} = [s_1^A, s_2^A, p^A]$, were similar to the probe, p^B, on the second trial (Trial B). This condition, represented in Figure 3, is labeled *hi*Sim after the relatively high similarity of the probe from Trial B to the stimuli from Trial A. On these trials, s_1^A, s_2^A, p^A and p^B were all within 3 JNDs of each other, as seen in Figure 3). If item information were carried over from trial to trial, memory of the study items on Trial A might influence recognition, inducing subjects to judge erroneously that p^B matched a study item on Trial B (s_1^B or s_2^B). In other words, carry-over of item information from Trial A to Trial B would be characterized by the proportion of false positive recognitions.

This *hi*Sim condition was contrasted with the *lo*Sim condition, in which pairs of trials were arranged so that p^B had a low similarity to Trial A's stimuli. Trials A and B in the *lo*Sim condition were the same as in the *hi*Sim condition, except that trials were paired such that Trial A's stimuli (s_1^A, s_2^A, p^A) differed from p^B by at least 5 JNDs. Such low similarity between stimuli on subsequent trials should give rise to very little proactive interference of item information, and few false positive recognitions.

In Trial A of both *hi*Sim and *lo*Sim conditions, the probe (p^A) and both study items (s_1^A, s_2^A) were all very similar to each other (that is, within 3 JNDs of each other). On this subset of trials, the probe and both study items all took values among the three highest allowed stimulus values, or the three lowest allowed stimulus values. The following trial, Trial B, always contained study items (s_1^B, s_2^B) that were only 1 or 2 JNDs from each other. The probe (p^B) differed from the closest study item by 5 JNDs. In all conditions, only the similarity among s_1, s_2 and p was constrained; their ordering in stimulus space along the w axis was not. Thus, s_1 was equally likely to take a value greater than or less than s_2. For simplicity, Figure 3 illustrates only the case where $s_1 < s_2$. In addition, the probe's value was equally likely to be

greater than or less than the study items. On Trial A (but not Trial B), the probe could even fall at a stimulus value between two study items, or hold an identical stimulus value to one of the items.

Previously, related procedures, using a "recent negative probe" condition have been shown to provide a sensitive assay of the degree of carry-over of item information (Monsell, 1978; D'Esposito et al., 1999). If information in memory did not carry over between trials, recognition performance on instances of Trial B in *hi*Sim should be no different from instances of Trial B in the *lo*Sim condition. One difference between the "recent negative probe" design and our own is that our design controlled the similarity of the probe from Trial B (p^B) to the stimuli in **A**. Thus in the *hi*Sim condition p^B could either exactly match a stimulus from Trial A ($p^B \in$ **A**) or be highly similar, though not identical to a stimulus from Trial A ($p^B \notin$ **A**). In the *lo*Sim condition, p^B was highly dissimilar from stimuli on Trial A (*lo*Sim condition). The similarity among stimuli could be quantified, allowing evaluation of the specificity of item information maintained from previous trials. Note that Trial B consistently has a low value of homogeneity. Trial B is the same in both *hi*Sim and *lo*Sim conditions and is likely, based on previous experiments (Kahana & Sekuler, 2002; Kahana et al., in press; Nosofsky & Kantner, 2006; Visscher et al., 2007; Yotsumoto et al., in press), to give rise to a relatively low false alarm rate on those trials.

Carry-over of relational information—The relationships among study stimuli on a trial robustly affect subjects' responses on that trial (Kahana & Sekuler, 2002; Kahana et al., in press; Nosofsky & Kantner, 2006; Visscher et al., 2007). To evaluate the possibility that such information was maintained from one trial to the next (Gorea & Sagi, 2000), we used a design parallel to that described above. Again, pairs of successive trials were generated, with the set up trial (first in the pair) varying in study item homogeneity. Guided by NEMo, we generated two kinds of set up trials, which we label as *hi*Hom and *lo*Hom. On *hi*Hom trials, s_1 and s_2 were highly homogeneous, differing from one another by just one JND. These trials were expected to promote a high, stricter criterion and fewer false alarms on that trial. On *lo*Hom trials, s_1 and s_2 differed from one another by at least 4 JND units, and thus had relatively low homogeneity. These trials were expected to promote a lower, more liberal criterion and more false alarms.

Presentation of either a *hi*Hom trial or a *lo*Hom trial was followed by the presentation of a neutral, test trial. These neutral, test trials were drawn from a pool of four different *Lure* stimulus sets (four sets of values for s_1^B, s_2^B, and p^B). Study stimuli and probes were chosen randomly for each set in the pool of neutral, test trials, and each of these random trials followed *hi*Hom and *lo*Hom trials with equal frequency, 20 times each. Any systematic difference in performance on neutral, test trials after *lo*Hom trials vs. *hi*Hom trials would indicate that relational information had been maintained and carried over to the neutral, test trial. For each subject, 40 *hi*Hom and 40 *lo*Hom trial pairs were randomly intermixed among all trials. Note that these condition labels refer to the characteristics of the first trial in a pair.

Results

Individual thresholds and perceptual similarity

Individual subject thresholds for discriminating between ripple sounds differed. Table 2 shows the just noticeable difference for each subject. The mean JND was 0.17, with a standard deviation of 0.053. Two stimuli differing by this proportion would be discriminated correctly 70% of the time. As mentioned earlier, the stimuli used to assess memory were created according to individual subjects' JND thresholds.

Carry-over of item information across trials

Stimulus information from Trial A carried over to Trial B. The proportion of false alarms was greater when a previous trial's stimuli could be confused with the current trial's probe (proportion "*Yes*" for *hi*Sim > *lo*Sim, paired *t*-test, $T_{11} = 2.55$, $p < 0.03$). In fact, only two subjects made any "*Yes*" responses in the *lo*Sim condition, while the other ten subjects made none. Because of the low variance in the *lo*Sim condition, we double-checked the statistics using nonparametric analyses: a two-sided Wilcoxon rank sum test also showed the effect to be strong and significant ($p < 0.03$). This indicates maintenance of stimulus memory from one trial to the next (see Methods and Figure 3). False alarm rates were low, as shown in Figure 3 (mean of 4.4 % in the *hi*Sim vs. 0.4% in the *lo*Sim condition). This modest rate was expected, as the probe is relatively dissimilar from either study item (the probe is 5 JNDs from the closest study item).

Further, we examined the selectivity of this carry-over effect. On 72% of *hi*Sim trials, p^B exactly matched s_1, s_2, or p from Trial A ($p^B \in \mathbf{A}$), whereas, due to stimulus constraints, on the remaining trials, p^B was 1 JND away from the closest stimulus on Trial A ($p^B \notin \mathbf{A}$). Subjects were significantly more likely to make a false recognition judgment when $p^B \in \mathbf{A}$ than on trials in which $p^B \notin \mathbf{A}$ (mean 5.4% vs 1.6%, paired *t*-test, $T_{11} = 2.57$, $p<0.03$). In fact, the false recognition rate where $p^B \notin \mathbf{A}$ was not significantly different from zero (*t*-test, $T_{11}=1.48$, $p = 0.16$). This indicates that carry-over from item information on previous trials is very selective, affecting exact matches much more than similar, but not identical item information.

Because all stimuli from Trial A as well as p^B were constrained to fall within 3 JNDs of each other, trials of condition $p^B \notin \mathbf{A}$ could only follow a Trial A that was a *Target* trial. This constraint meant that difference between the $p^B \in \mathbf{A}$ and $p^B \notin \mathbf{A}$ conditions could have arisen from a difference in *Target* vs. *Lure* trials. However, there was no overall difference between the false alarm rates for *Target* and *Lure* trials in the *hi*Sim condition (means 3.9% and 5.3%, paired *t*- test, $T_{11} = 0.64$, $p=0.53$). This indicates that the increase in false alarm rate in cases where $p^B \in \mathbf{A}$ likely reflects increased similarity, rather than an artifact of following *Lure* trials more often. On average, 20% of the Trial A cases were *Lure* trials (simply due to the constraint that both stimuli and probe from that trial were within 3 JNDs). False alarms following *Lure* trials were not significantly more frequent than false alarms following *Target* trials, in either *hi*Sim or *lo*Sim conditions (paired *t*-test, $T_{11} = 0.64$ and 0.74, respectively and $p=0.53$ and 0.47 respectively).

Overall, the data demonstrate carry-over of item information across trials. This interference lessens when p^B is less similar to the interfering stimuli.

Carry-over of relational information across trials

Figure 4 shows the mean proportion of "*Yes*" responses on trial $n+1$, when trial n was *hi*Hom (left box and whiskers) or *lo*Hom (right box and whiskers). These two proportions were nearly identical to one another (paired *t*-test, $T_{11} = 1.0$, $p = 0.33$, suggesting that whatever effect might have been generated on trial n did not carry over to trial $n+1$. Confirming the absence of an effect here, mean reaction times also did not differ between the groups of trials whose data are shown in Figure 4 (paired *t*-test, $T_{11} = 0.56$, $p = 0.59$).

Figure 4 suggests that manipulation of inter-item homogeneity on trial n has no detectable effect on trial $n+1$. Before concluding that this represents an absence of carry-over from one trial to the next, it is essential to verify that the manipulation of inter-item homogeneity did indeed affect performance on the current trial. To this end, we examined trials that simultaneously met three criteria. For inclusion in the analysis, a trial had to (1) qualify as a *hi*Sim or a *lo*Sim trial, (2) be a Lure trial, and (3) have a summed probe-item similarity of 5

JNDs (that is, $|p\text{-}s_1| + |p\text{-}s_2| = 5$ JNDs). These constraints make it possible to examine the effect of inter-item homogeneity without confounds from other parameters, such as summed probe-item similarity, that are known to affect performance. In order to compare a subject's proportion correct for any two conditions, the probe-item similarity must be the same in both conditions. A summed probe-item similarity of 5 JNDs was chosen in constraint (3) because this value gave the largest proportion of trials. Of course, only a fraction of all trials survived the imposition of the three constraints: of the 40 total trials per condition, only an average of 12.5 *hi*Hom and 20 *lo*Hom trials per subject survived. The effect of inter-item homogeneity was robust and statistically significant despite the relatively small number of trials. The select *hi*Hom trials produced a mean false alarm rate of 15.9%; the select *lo*Hom trials on which $|s_1\text{-}s_2|=5$ JNDs produced almost twice as many false alarms, 30.6%, a statistically significant difference (paired *t*-test, $T_{11} = 2.5$, $p < 0.05$).

We should note that this robust difference between *hi*Hom and *lo*Hom trials was not some artifact of the strict selection process used to select the subset of trials we analyzed. In fact, the 14% difference in false alarms between conditions in the selected subset of data was quite in line with the analogous difference seen in the entire data-set, with no imposition of selection criteria. In particular, the overall difference between *hi*Hom and *lo*Hom trials was 12% (paired *t*-test, $T_{11} = -2.7$, $p < 0.05$). Thus, the absence of a difference between conditions in Figure 4 did not mean that the manipulation of inter-item homogeneity was ineffectual; rather it showed that the robust effect generated by the manipulation failed to survive from one trial to the next.

Discussion

Item information interferes across trials, but relational information does not

Our results show that interference from item information can operate over successive trials, allowing what was heard on trial *n* to influence recognition on trial *n* + 1. In particular, as Figure 3 shows, carry-over from item information produces more false alarms in the *hi*Sim than the *lo*Sim condition, an effect that depends strongly on the degree of similarity between stimuli on successive trials. In contrast to this strong inter-trial influence, the effect generated by the homogeneity of one trial's study stimuli fails to influence performance on the next trial: No difference was found between trials following *lo*Hom and *hi*Hom conditions (Figure 4).

Our decision to examine auditory memory was stimulated in part by the recent report of several detailed similarities between short-term memory for visual stimuli and short-term memory for comparable auditory stimuli (Visscher et al., 2007). The cross-modality similarity in so many aspects of memory invited comparison of carry-over from item information and relational information. The data presented here suggest that the mechanisms giving rise to short-term memory for item and relational information are distinct.

Carry-over of item information

Our results confirm that stimuli encountered on trial *n* can affect responses on trial *n* +1. This finding had been demonstrated earlier for verbal and visual stimuli, but not, to our knowledge, until now for abstract auditory stimuli in humans (Bennett, 1975; D'Esposito et al., 1999; Feredoes, Tononi, & Postle, 2006; Monsell, 1978; Postle et al., 2004; Wickens, 1972; Wright, 1999). Previous studies of item information carry-over showed small but significant reaction time effects and small effects on accuracy (so long as accuracy was below its upper limiting value). This pattern of results is consistent with the modest, but significant differences in performance found here in Figure 3. Finally, it is worth noting that even under conditions that generate the greatest carry-over, relatively few false recognitions result (only 4.4%). This is not attributable to some dissipation of memory over the interval between trials; in fact, we have shown that in the absence of interference, memory for the stimuli used here is very well

preserved for many seconds (Visscher et al., 2007). So, despite the highly significant carryover of item information, subjects are usually able to gate out item memory from the previous trial.

Memory trace for interfering item information is sharply tuned—The carry-over of item information depends on how closely the previous trial's stimuli match the current trial's probe. Previous experiments showed that subjects use summed similarity to judge whether the probe matches a stimulus in the current study list (Brockdorff & Lamberts, 2000; Clark & Gronlund, 1996; Humphreys et al., 1989; Kahana & Sekuler, 2002; Kahana et al., in press; Lamberts et al., 2003; Nosofsky, 1986, 1991; Nosofsky & Kantner, 2006; Yotsumoto et al., in press). Our data extend this point, suggesting that a full computational account of short-term recognition memory must take account of a probe's similarity to stimuli not only on the current trial, as NEMo does, but also on a previous trial or trials.

The data indicate that trial-to-trial interference from item information is highly selective, as moving p^B only one JND away from the closest stimulus on the previous trial caused a relatively large and reliable decrease in false recognitions, from 5.4% to 1.6%. This shows that the memory trace for the interfering item information is sharply tuned for perceptual similarity.

Further studies using metric stimuli even less than 1 JND apart could more finely examine the tuning curve. Related further studes could compare the tuning of interfering memory to the tuning for memories of stimuli on the current trial. Additionally, such experiments could determine how closely the "recent negative effects" observed in other paradigms depend on exact identity between a probe and a recent stimulus, or whether rough similarity to the probe is sufficient.

The stimulus-tuning of a possibly related effect, "release" from proactive interference (Wickens, 1972; Wickens, Born, & Allen, 1963; Bunting, 2006), could also be profitably examined using appropriate metric stimuli. In the "release" paradigm, subjects are presented successive lists of words containing semantically related items. For example, successive lists might contain the names of animals. As expected, recall worsens over successive trials, as proactive interference builds up. However, recall is abruptly and dramatically improved by the presentation of one set of stimuli, for example, a list of vegetable names, that are semantically unrelated to the preceding stimuli. It is not known whether this release from proactive interference requires stimuli that are categorizable (as vegetable or animal names); nor is it known precisely how large the difference must be between stimuli that generate the proactive interference and stimuli that "release" it.

Our results demonstrate that proactive interference is strongest when the probe stimulus precisely matches a study item from the preceding trial (Figure 3). This suggests that stimulus values are maintained in memory with considerable fidelity even across trials. Some researchers have suggested that carry-over of item-specific information reflects a residual item-memory trace generated on a previous trial (Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998). Other researchers have suggested that such proactive interference reflects a reliance on general familiarity information that is available when explicit recollection has failed (Jacoby et al., 2001;Tulving, 1985). The current data add to this debate by showing that stimulus information is precisely maintained by the memory mechanism responsible for proactive interference.

Extensions of proactive interference effects

Our experimental design's reliance on *pairs* of trials limits assessment of interference effects to just one previous trial. Of course, it is certainly possible that some interference effects persist beyond one trial. For example, Cho et al. (2002) make the case that interference can persist from individual stimuli several trials removed from the current trial. Further research might

explore the number of trials over which carry-over effects for complex auditory stimuli can act (Petzold & Haubensak, 2001). Another possible source of proactive interference is the accumulation of interfering information across many items from many previous trials, often called item-non-specific interference (Postle et al., 2004; Postman & Keppel, 1977). Both item-specific and item-non-specific forms of interference are thought to be mediated by the same regions of the brain (Postle et al., 2004). Studies in monkeys demonstrated item-non-specific interference for auditory stimuli (non-confusable environmental sounds) (Wright, 1999). Such non-specific effects may be relevant for the present experiments as well, as stimuli were chosen to be somewhat confusable, and to resist perfect, consistent categorization. Thus, the item-specific proactive interference effects we observed may be operating against a background of item non-specific effects.

Proactive interference from stimulus information is but one of many ways in which stimulus information on one trial might affect a subject's response on a subsequent trial. For example, repeated presentation of a stimulus affects a subject's judgment of subsequent stimuli, so that later stimuli are recalled differently depending on their relationship to the well-studied stimulus (Huttenlocher et al., 1991; Visscher et al., 2003). In addition, the much-studied phenomenon of priming (Henson, 2003), in which perception of a stimulus is enhanced based on input from a previous trial, also depends on the maintenance of stimulus information from one trial to the next. Most salient for the issues we address are demonstrations that various forms of stimulus information can be maintained across trials for priming, including non-verbal information such as spatial frequency and color (Huang, Holcombe, & Pashler, 2004; Maljkovic & Nakayama, 1994).

All of these effects (proactive interference, effects of a well-studied stimulus, and priming) reflect the preservation of stimulus representations over trials. In each case, remembered stimulus information affects later recognition responses. The relationship among study stimuli constitutes a different type of information, relational information, which does not directly come from the individual stimulus values. It is notable that relational information is treated differently from many other types of trial-related information and is not maintained across trials.

Item information and relational information are encoded separately

The data show that stimulus information generated on one trial carries over to interfere with recognition judgments on the succeeding trial, but that the effect of inter-item homogeneity does not. This indicates that item-specific information, not information about relationships between study items, carries over between trials.

The recognition task used here imposes artificial temporal, episodic boundaries defining which remembered stimuli are relevant for the current trial. Specifically, only the two stimuli seen most recently are relevant. The implicit reward structure of the experiment punishes subjects for allowing information acquired on trial n to affect responses on trial $n+1$. However, in normal, everyday application of short-term memory, temporal boundaries are less distinct, and maintenance of stimulus information can be advantageous. Therefore it makes sense that despite the expected reward structure, subjects will not show a perfect ability to exclude recent, but no longer relevant information from memory. Without maintaining item information across episodes, we would not be able to convert episode information into more general knowledge. For example, without the ability to generalize across episodes, one may not be able to infer from a previous episode that the roar of a river indicates the presence of water nearby. This might degrade the ability to find water. On the other hand, rapid adjustments based on relational information would be useful in real-world situations, for example, orienting toward the river on a still day requires a relatively lax criterion for identifying water sounds among few distractors, while moments later, if the wind picks up, discriminating the water sounds from the rustling of leaves requires a much more stringent criterion. Thus, rapid, situation- and task-

dependent changes in responses to relational information would be advantageous in a way that complete inhibition of item information would not.

Summary

We found that remembered information about stimulus properties (item information about auditory stimuli) carries over from one trial to the next, as indexed by the recognition judgment on the next trial. This form of proactive interference appears to be relatively narrowly tuned. In contrast, we have also demonstrated that at least one form of relational information is temporally restricted, and does not carry over into the successive episode. Further experiments are needed to examine whether other forms of relational information (such as serial position information, for example) are similarly restricted in scope.

Acknowledgments

References

Ball K, Sekuler R. Models of stimulus uncertainty in motion perception. Psychological Review 1980;87 (5):435–69. [PubMed: 7413887]

Bennett RW. Proactive interference in short-term memory: Fundamental forgetting processes. Journal of Verbal Learning & Verbal Behavior 1975;14:573–584.

Brainard D. The psychophysics toolbox. Spatial Vision 1997;10(4):433–6. [PubMed: 9176952]

Brockdorff N, Lamberts K. A feature-sampling account of the time course of old-new recognition judgments. Journal Experimental Psychology: Learning, Memory, & Cognition 2000;26:77–102.

Bunting M. Proactive interference and item similarity in working memory. Journal of Experimental Psychology: Learning, Memory, & Cognition 2006;32(2):183–196.

Cho RY, Nystrom LE, Brown ET, Jones AD, Braver TS, Holmes PJ, et al. Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. Cognitive and Affective Behavioral Neuroscience 2002;2(4):283–99.

Clark SE, Gronlund SD. Global matching models of recognition memory: How the models match the data. Psychonomic Bulletin & Review 1996;3:37–60.

D'Esposito M, Postle B, Jonides J, Smith E. The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related functional MRI. Proceedings of the National Academy of Sciences, USA 1999;96(13):7514–9.

Dyne A, Humphreys M, Bain J, Pike R. Associative interference effects in recognition and recall. Journal of Experimental Psychology: Learning, Memory, & Cognition. 1990

Einstein G, Hunt R. Levels of processing and organization: Additive effects of individual-item and relational processing. Journal of Experimental Psychology: Human Learning and Memory 1980;6 (5):588–598.

Feredoes E, Tononi G, Postle B. Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory. Proceedings of the National Academy of Sciences, USA 2006;103(51):19530–4.

Gorea A, Caetta F, Sagi D. Criteria interactions across visual attributes. Vision Research 2005;45(19): 2523–32. [PubMed: 15950255]

Gorea A, Sagi D. Failure to handle more than one internal representation in visual detection tasks. Proceedings of the National Academy of Sciences, USA 2000;97(22):12380–4.

Heit E, Brockdorff N, Lamberts K. Adaptive changes of response criterion in recognition memory. Psychonomic Bulletin & Review 2003;10(3):718–23. [PubMed: 14620369]

Henson RNA. Neuroimaging studies of priming. Progress in Neurobiology 2003;70(1):53–81. [PubMed: 12927334]

Huang L, Holcombe A, Pashler H. Repetition priming in visual search: episodic retrieval, not feature priming. Memory & Cognition 2004;32(1):12–20.

Humphreys MS, Pike R, Bain JD, Tehan G. Global matching: A comparison of the SAM, Minerva II, Matrix, and TODAM models. Journal of Mathematical Psychology 1989;33:36–67.

Hunt RR, Einstein GO. Relational and item-specific information in memory. Journal of Verbal Learning & Verbal Behavior 1981;20:497–514.

Huttenlocher J, Hedges L, Duncan S. Categories and particulars: prototype effects in estimating spatial location. Psychological Review 1991;98(3):352–76. [PubMed: 1891523]

Ikeda T, Osaka N. How are colors memorized in working memory? A functional magnetic resonance imaging study. Neuroreport 2007;18(2):111–114. [PubMed: 17301673]

Jacoby L, Debner J, Hay J. Proactive interference, accessibility bias, and process dissociations: valid subjective reports of memory. Journal of Experimental Psychology: Learning, Memory & Cognition 2001;27(3):686–700.

Jiang Y, Olson IR, Chun MM. Organization of visual short-term memory. Journal of Experimental Psychology: Learning, Memory & Cognition 2000;26(3):683–702.

Jones AD, Cho RY, Nystrom LE, Cohen JD, Braver TS. A computational model of anterior cingulate function in speeded response tasks: effects of frequency, sequence, and conflict. Cognitive and Affective Behavioral Neuroscience 2002;2(4):300–17.

Jonides J, Nee D. Brain mechanisms of proactive interference in working memory. Neuroscience 2006;139(1):181–93. [PubMed: 16337090]

Jonides J, Smith E, Marshuetz C, Koeppe R, Reuter-Lorenz P. Inhibition in verbal working memory revealed by brain activation. Proceedings of the National Academy of Sciences, USA 1998;95(14): 8410–3.

Kahana, MJ. Foundations of human memory. Oxford University Press; in press

Kahana MJ, Sekuler R. Recognizing spatial patterns: a noisy exemplar approach. Vision Research 2002;42(18):2177–192. [PubMed: 12207978]

Kahana MJ, Zhou F, Geller A, Sekuler R. Lure-similarity affects visual episodic recognition: Detailed tests of a noisy exemplar model. Memory & Cognition. in press

Lamberts K, Brockdorff N, Heit E. Feature-sampling and random-walk models of individual-stimulus recognition. Journal of Experimental Psychology: General 2003;132(3):351–78. [PubMed: 13678373]

Lustig C, May CP, Hasher L. Working memory span and the role of proactive interference. Journal of Experimental Psychology: General 2001;130(2):199–207. [PubMed: 11409099]

Maljkovic V, Nakayama K. Priming of pop-out: I. role of features. Memory & Cognition 1994;22(6): 657–72.

May CP, Hasher L, Kane MJ. The role of interference in memory span. Memory & Cognition 1999;27 (5):759–767.

Monsell S. Recency, immediate recognition memory, and reaction time. Cognitive Psychology 1978;10:465–501.

Morgan MJ, Watamaniuk SN, McKee SP. The use of an implicit standard for measuring discrimination thresholds. Vision Research 2000;40(17):2341–9. [PubMed: 10927119]

Nosofsky R. Attention, similarity, and the identification-categorization relationship. Journal of Experimental Psychology: General 1986;115(1):39–61. [PubMed: 2937873]

Nosofsky R. Tests of an exemplar model for relating perceptual classification and recognition memory. Journal of Experimental Psychology: Human Perception & Performance 1991;17(1):3–27. [PubMed: 1826320]

Nosofsky R, Kantner J. Exemplar similarity, study list homogeneity, and short-term perceptual recognition. Memory & Cognition 2006;34(1):112–24.

Olsson H, Poom L. Visual memory needs categories. Proceedings of the National Academy of Sciences, USA 2005;102(24):8776–80.

Petzold P, Haubensak G. Higher order sequential effects in psychophysical judgments. Perception & Psychophysics 2001;63(6):969–78. [PubMed: 11578058]

Petzold P, Haubensak G. The influence of category membership of stimuli on sequential effects in magnitude judgment. Perception & Psychophysics 2004;66(4):665–78. [PubMed: 15311665]

Postle B, Brush L, Nick A. Prefrontal cortex and the mediation of proactive interference in working memory. Cognitive and Affective Behavioral Neuroscience 2004;4(4):600–8.

Postman L, Keppel G. Conditions of cumulative proactive-inhibition. Journal of Experimental Psychology: General 1977;106(4):376–403.

Ruusuvirta T. Proactive interference of a sequence of tones in a two-tone pitch comparison task. Psychonomic Bulletin & Review 2000;7(2):327–331. [PubMed: 10909141]

Ruusuvirta T, Wikgren J, Astikainen P. Proactive interference in a two-tone pitch-comparison task without additional interfering tones. Psychological Research. 2006

Shamma S. On the role of space and time in auditory processing. Trends in Cognitive Science 2001;5(8): 340–348.

Singer M, Wixted JT. Effect of delay on recognition decisions: evidence for a criterion shift. Memory & Cognition 2006;34(1):125–137.

Smith ML, Leonard G, Crane J, Milner B. The effects of frontal- or temporal-lobe lesions on susceptibility to interference in spatial memory. Neuropsychologia 1995;33(3):275–85. [PubMed: 7791996]

Stretch V, Wixted JT. Decision rules for recognition memory confidence judgments. Journal of Experimental Psychology: Learning, Memory & Cognition 1998;24(6):1397–1410.

Treisman M, Williams TC. A theory of criterion setting with an application to sequential dependencies. Psychological Review 1984;91(1):68–111.

Tulving E. How many memory systems are there. American Psychologist 1985;40(4):385–398.

Underwood BJ. Interference and forgetting. Psychological Review 1957;64:49–60. [PubMed: 13408394]

Visscher K, Kaplan E, Kahana MJ, Sekuler R. Auditory short-term memory behaves like visual short-term memory. Public Library of Science, Biology 2007;5(3):e56.

Visscher K, Viets E, Snyder L. Effects of training on memory-guided saccade performance. Vision Research 2003;43(19):2061–71. [PubMed: 12842159]

Watson AB, Pelli DG. QUEST: a Bayesian adaptive psychometric method. Perception & Psychophysics 1983;33(2):113–120. [PubMed: 6844102]

Wickens, DD. Characteristics of word encoding. In: Melton, A.; Martin, E., editors. Coding processes in human memory. Winston; 1972.

Wickens DD, Born DG, Allen CK. Proactive inhibition and item similarity in short-term memory. Journal of Verbal Learning & Verbal Behavior 1963;2(56):440–445.

Wright AA. Auditory and visual serial position functions obey different laws. Psychonomic Bulletin & Review 1998;4:564–584.

Wright AA. Auditory list memory and interference processes in monkeys. Journal of Experimental Psychology: Animal Behavior processes 1999;25(3):284–96. [PubMed: 10423854]

Yotsumoto Y, Kahana MJ, Wilson H, Sekuler R. Recognition memory for realistic synthetic faces. Memory & Cognition. in press

Zhou F, Kahana MJ, Sekuler R. Short-term episodic memory for visual textures: a roving probe gathers some memory. Psychological Science 2004;15(2):112–8. [PubMed: 14738518]
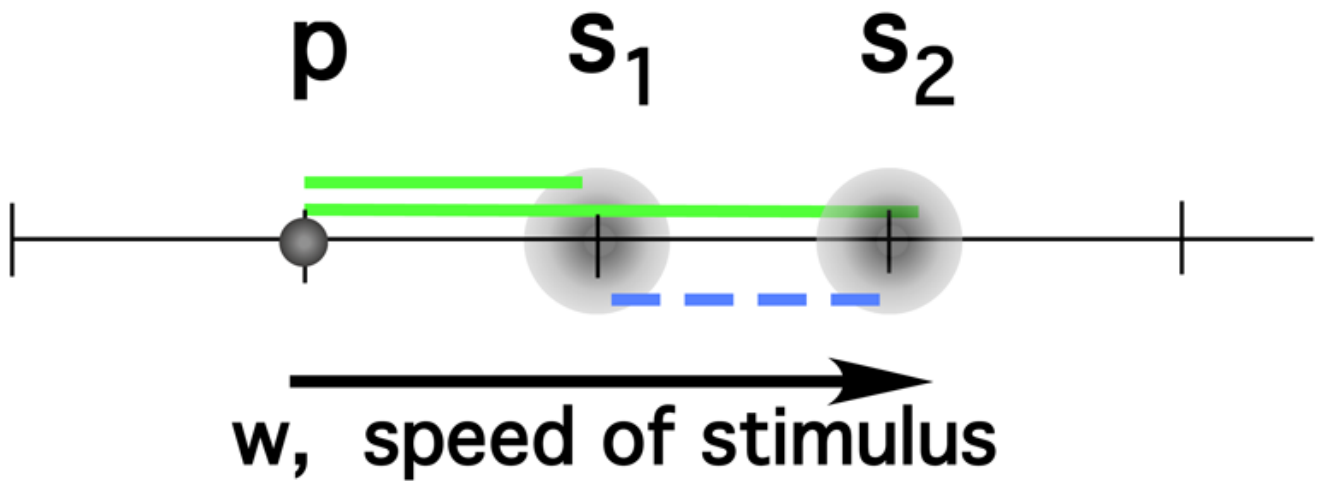
Figure 1.
Schematic illustration of elements entering into a summed similarity computation. Perceptual representations for two study items, s_1 and s_2, are defined along a single stimulus dimension. On any trial, the memory of some study item is a random sample from a distribution (the probability density function of possible memories for that item). The diameters of the schematic "clouds" signify the noise or variability associated with the memory of each stimulus item. The solid lines represent the similarity of each remembered stimulus to the probe (p). The dashed line represents relational information: the homogeneity of the two study items.
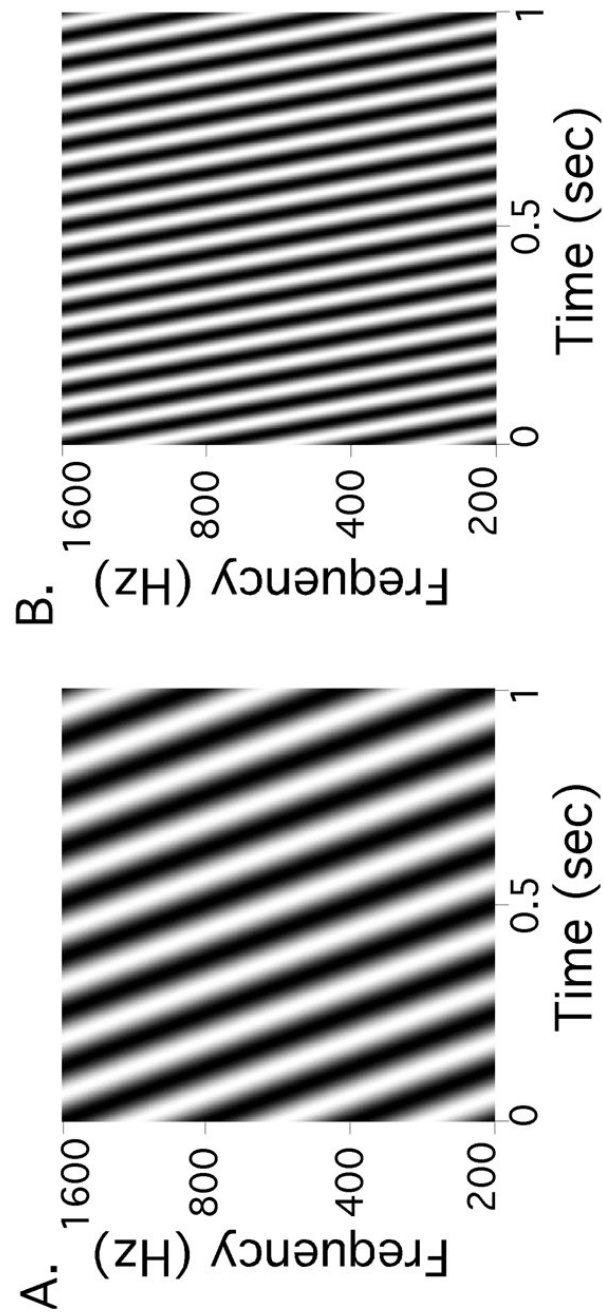
Figure 2.
Spectro-temporal plots of ripple sounds. The horizontal axis shows time in seconds, while the vertical axis shows frequency content in Hertz. Darker colors represent sounds of greater amplitude. Modulations over time are referred to as the ripple's velocity, with units of sinusoidal frequency w, while modulations over frequency are referred to as spectral density, with units of sinusoidal frequency Ω. Part A represents a stimulus with $w = 8$ Hz; Part B represents a stimulus with $w = 16$ Hz. In our experiment, other stimulus parameters are held constant, for example, $\Omega = 1$ cycle per octave.
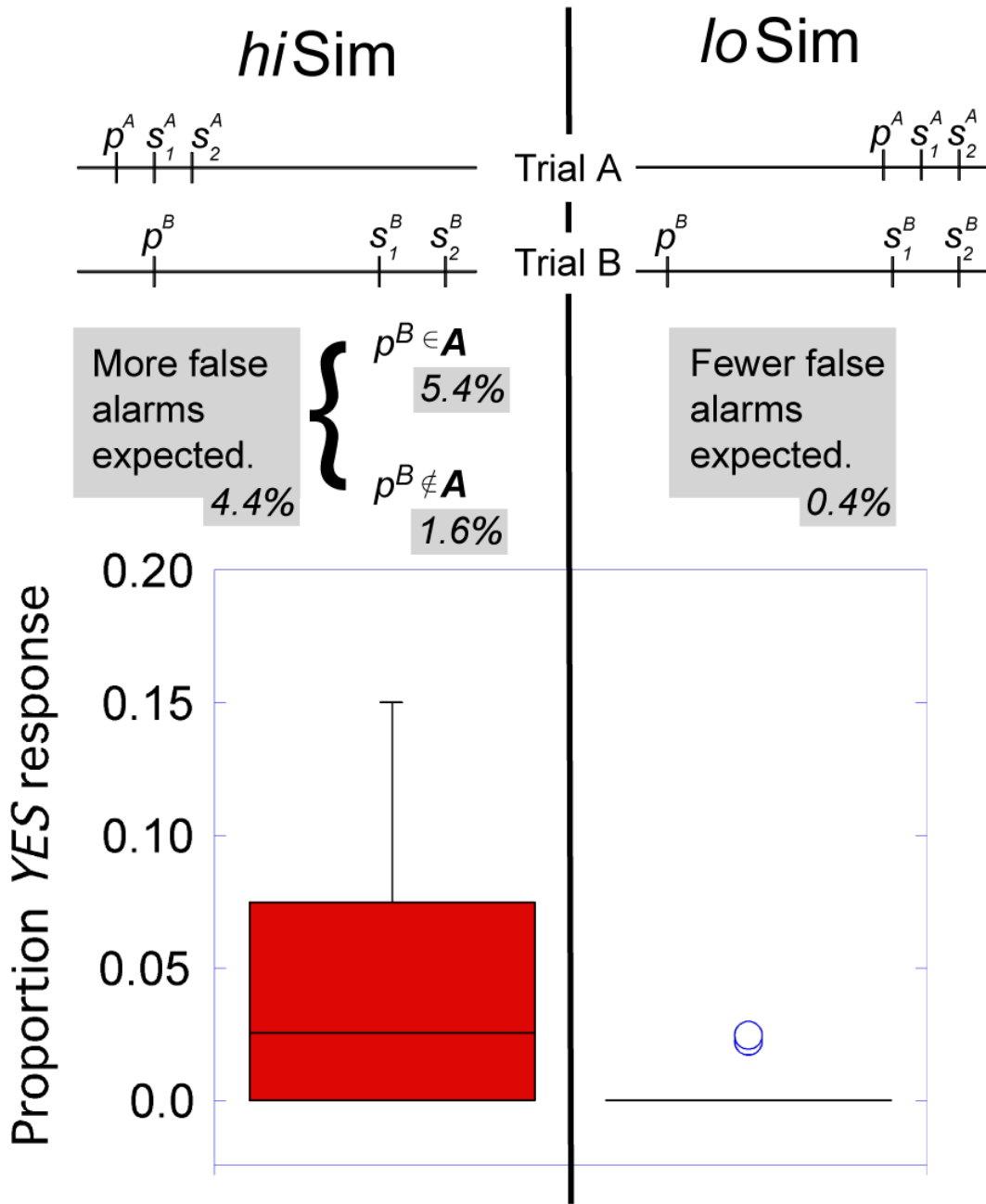
Figure 3.
Upper panel: Schematic diagram of design examining maintenance of item information across trials. Trial A immediately precedes Trial B. In the *hi*Sim condition, the probe for Trial B (with value p^B) is very similar in perceptual space to the stimuli from Trial A (study stimuli and, with probe, s_1^A and s_2^A, with probe, p^A). The horizontal axis represents the stimulus space, items closer to each other are more similar. In the *lo*Sim condition, p^B is different from s_1^A, s_2^A, and p^A. If item information were maintained from trial to trial, more false alarms (on Trial B) would be expected for the *hi*Sim condition, as interference between the probe and stimuli from a previous trial should be greater in that condition than the *lo*Sim condition. Note that Trial B is

the same in each case. p indicates the probe frequency, s_1 and s_2 indicate the frequencies of the first and second study stimuli. Within the *hi*Sim condition, on some trials p^B matched a stimulus from Trial A ($p^B \in \mathbf{A}$), while on other trials, p^B did not match any stimulus from Trial A ($p^B \notin \mathbf{A}$). The percentage of false alarm trials observed for each condition are shown in filled boxes. Lower panel: Box plot shows median (thick bar); middle 50% of data are encompassed by boxes. The whiskers include all data points that are not outliers. Circles represent outliers, defined as points $>1.5\times$ the inter- quartile range from the median. Note that more false alarms were made in the *hi*Sim condition.
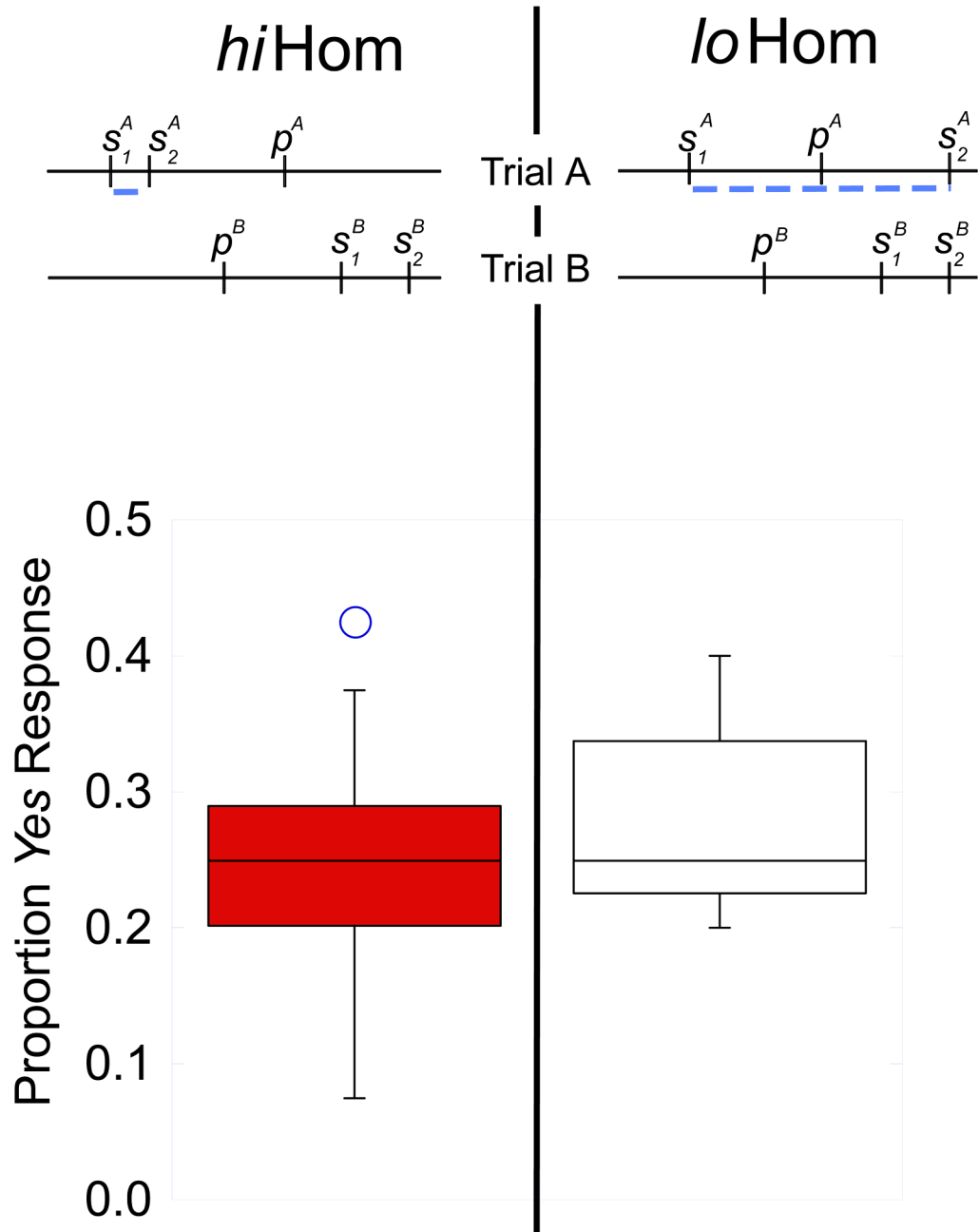
Figure 4.
Upper panel: Schematic diagram of design examining maintenance of relational information across trials. Trial A immediately preceded Trial B. In the *hi*Hom condition, the stimuli in Trial A were homogeneous (blue dashed line is short), while in Trial B, the two stimuli were different from one another (blue dashed line is longer). The horizontal axis represents the stimulus space, items closer to each other are more similar. If relational information were maintained from trial to trial, more false alarms (on Trial B) would be expected for the *lo*Hom condition, as the false alarm rate on Trial A is greater in that condition. Note that Trial B is the same in each case. p indicates the probe frequency, s_1 and s_2 indicate the frequencies of the first and second study stimuli. Lower panel: Identical trials following trials of high inter-item homogeneity (*hi*Hom)

or low inter-item homogeneity (*lo*Hom) showed no difference in proportion correct. Symbol conventions as in Figure 3. This form of relational information is not carried over from trial to trial.

Table 1

Trial types in experimental design. Columns refer to the effect examined for each group of trials, the condition, the number of types of lure trials (combinations of two list items and a probe), number of types of target trials, and the number of repetitions of each type of trial

Effect examined	Condition	Lure	Target	Reps
Item Information	*hi*Sim	4	4	10
Item Information	*lo*Sim	4	4	10
Relational Information	*hi*Hom	4	4	10
Relational Information	*lo*Hom	4	4	10
Model testing	Other 2-item lists	64	64	20
Define perceptual similarity	One study item (Targets)	-	10	90
Define perceptual similarity	One study item (Lures)	90	-	10

Table 2

Thresholds for 70% correct recognition performance

Subject	JND	Subject	JND
1	0.22	7	0.08
2	0.11	8	0.17
3	0.26	9	0.22
4	0.23	10	0.13
5	0.17	11	0.14
6	0.16	12	0.15