

Effects of capacity limits, memory loss, and sound type in change deafness

Melissa K. Gregg¹ · Vanessa C. Irsik² · Joel S. Snyder²

Published online: 30 August 2017
© The Psychonomic Society, Inc. 2017

Abstract Change deafness, the inability to notice changes to auditory scenes, has the potential to provide insights about sound perception in busy situations typical of everyday life. We determined the extent to which change deafness to sounds is due to the capacity of processing multiple sounds and the loss of memory for sounds over time. We also determined whether these processing limitations work differently for varying types of sounds within a scene. Auditory scenes composed of naturalistic sounds, spectrally dynamic unrecognizable sounds, tones, and noise rhythms were presented in a change-detection task. On each trial, two scenes were presented that were same or different. We manipulated the number of sounds within each scene to measure memory capacity and the silent interval between scenes to measure memory loss. For all sounds, change detection was worse as scene size increased, demonstrating the importance of capacity limits. Change detection to the natural sounds did not deteriorate much as the interval between scenes increased up to 2,000 ms, but it did deteriorate substantially with longer intervals. For artificial sounds, in contrast, change-detection performance suffered even for very short intervals. The results suggest that change detection is generally limited by capacity, regardless of sound type, but that auditory memory is more enduring for sounds with naturalistic acoustic structures.

Keywords Auditory memory · Change deafness · Processing capacity

A surprising type of perceptual error that occurs during auditory tasks is the inability to detect large changes to objects in scenes. This phenomenon, known as change deafness (Eramudugolla, Irvine, McAnally, Martin, & Mattingley, 2005), is intriguing on a theoretical level because such large perceptual errors suggest fundamental limitations to our perceptual and mnemonic representations of the environment. Several lines of research on this topic have shown that change deafness (and its visual analog, change blindness) is partially dependent on attentional limitations (Rensink, O'Regan, & Clark, 1997) and on a limited capacity short-term/working memory (Luck & Vogel, 1997). However, it remains unclear whether loss of information over time, due to decay (Peterson & Peterson, 1959), interference (Keppel & Underwood, 1962), or sudden death (Zhang & Luck, 2009) is also a major contributing factor to change deafness, or for that matter, to change blindness.

Multiple studies have provided evidence that change deafness is at least partly caused by a limitation in processing capacity, including limitations in attending to, encoding, or maintaining multiple sounds in auditory memory. For example, change deafness increases as the number of sounds within a scene is increased (Eramudugolla et al., 2005; McAnally et al., 2010). In addition, experimental manipulations that alleviate general processing constraints on change detection have been shown to reduce change deafness. Attentional limitations can be reduced by focusing attention via a valid cue indicating the name of the object that will change (Eramudugolla et al., 2005) or by presenting a cue indicating a particular spatial location where the change will occur (Backer & Alain, 2012). Capacity limitations are apparent in

✉ Melissa K. Gregg
greggm@uwp.edu

✉ Joel S. Snyder
joel.snyder@unlv.edu

¹ Department of Psychology, University of Wisconsin–Parkside, Kenosha, WI 53144, USA

² Department of Psychology, University of Nevada, Las Vegas, NV 89154, USA

other auditory perception tasks as well, such as in auditory working memory tests (Li, Cowan, & Saults, 2013) and informational masking paradigms in which listeners are unable to identify a target tone in the presence of a masker that is distinct in frequency from the target (Durlach et al., 2003a, b). Collectively, these results suggest that capacity might limit perception of complex auditory scenes in general.

The extent to which other processing limitations contribute to change deafness is not well understood. For example, an additional cause of change deafness could be information loss over time. If the memory of the scene decays or is interfered with during the course of a trial, then change detection would fail. Information loss has not been systematically studied in change blindness, and there are only two research groups, to our knowledge, that have systematically investigated the issue of memory loss over time in change deafness. These research efforts have produced conflicting results. A series of experiments by Demany, Trost, Serman, and Semal (2008) provided evidence for information loss over time using simple stimuli. In one such experiment, listeners were presented with successive pairs of nonrecognizable chords composed of several simultaneous pure tones, one of which could change on a given trial. The successive chords were separated by varying delays (ranging from 0 to 2000 ms) and composed of a varying number of pure tones (e.g., 4, 7, or 12). Demany et al. (2008) found that change-detection performance declined as the number of pure tones within each chord increased and as the delay between chords increased, indicating important roles for processing capacity and information loss over time, respectively.

In contrast, a study using complex, naturalistic sounds demonstrated no apparent information loss over time (Pavani & Turatto, 2008). In this study, listeners were asked to detect changes to scenes of three or four animal sounds. The scenes were separated by 500 ms of silence, 500 ms of noise, or no delay interval, the latter to evaluate the potential role of auditory transients in change deafness rather than information loss. Change-deafness rates were higher for scene sizes of four sounds than for three sounds, but change deafness was just as prevalent with no delay interval between scenes as it was with a 500 ms silent or noise interval. This finding suggests that change deafness with natural sounds, which are more typically used in change-deafness studies (see reviews by Dickerson & Gaston, 2014; Snyder & Gregg, 2011; Snyder, Gregg, Weintraub, & Alain, 2012), does not result from information loss over time.

These previous studies illustrate that there is not yet a clear answer as to whether change deafness is partly caused by information loss over time. One potential reason for the discrepancy across studies is that different types of stimuli were used. There is evidence outside the change-deafness literature suggesting that stimulus type may modulate the encoding and the maintenance of auditory memories. Enhanced recognition

memory has been shown to occur across a long interval with speech sounds relative to naturalistic sounds (Cohen, Horowitz, & Wolfe, 2009) and vocal melodies relative to piano melodies (Weiss, Vanzella, Schellenberg, & Trebut, 2015), which suggests specialized memory for particular types of familiar sounds. Electrophysiological data have also been reported showing that different categories of complex sounds are processed by distinct neural networks as early as 70 ms after stimulus onset (Murray, Camen, Gonzalez Andino, Bovet, & Clarke, 2006). Together, these findings suggest that the auditory system has distinct mechanisms for processing certain classes of spectro-temporally complex sounds, which could lead to perceptual and memory advantages over more artificial sounds (for a similar finding in vision, cf. Brady, Störmer, & Alvarez, 2016).

In addition, the simple chord stimuli used in Demany et al. (2008) may not be appropriate for measuring general mechanisms of object memory loss over time in a change-detection task. Such stimuli may allow for a specialized frequency-change detection mechanism that is sensitive to very small frequency changes (Demany et al., 2008). Thus, it is important to evaluate this issue using a variety of sound types, including simple, artificial sounds (such as pure tones and noise rhythms) and more naturalistic, complex sounds (sounds composed of multiple time-varying components) within the same paradigm. Comparisons across studies are also difficult, given the different range of delay intervals used: Pavani and Turatto (2008) only compared delay intervals of 0 ms and 500 ms, while Demany et al. (2008) tested a wider range of delay intervals (0–2,000 ms). Thus, it is important to carefully evaluate this issue using multiple sound types and a comparable range of delay intervals.

In the present study, we address the contribution of capacity limitations and information loss on change deafness by using varying delay intervals and varying scene sizes with four sound types. Two types of artificial sounds were used to ensure that any pattern obtained with one type of artificial sound (e.g., noise rhythms) is generalizable to other artificial sounds (i.e., pure tone rhythms). Two types of naturalistic sounds were used: a set of recognizable environmental sounds and a set of unrecognizable sounds. The unrecognizable sounds were scrambled versions of the recognizable sounds and were used to test for the potential of listeners relying on a verbal (or semantic) memory strategy. Though it is possible to affix verbal labels to nonverbal stimuli (see Braidá, Lim, Berliner, & Durlach, 1984), our manipulation made it quite difficult to do so. If information loss over time is a major contributing factor to change deafness, then more change deafness should occur to all sound types as the delay interval between scenes is increased. If, on the other hand, auditory memory has extended storage time, then change deafness may not be affected by the delay interval between scenes. Furthermore, if memory representations of naturalistic, familiar sounds are more robust

than representations for artificial, less familiar sounds, more change deafness would be expected to occur for artificial sounds compared to more natural sounds for larger delay intervals.

Experiment 1

Method

Participants Forty-eight listeners with normal hearing participated in this experiment (24 females and 24 males; mean age = 19.4 years, range: 18–25 years). In this and in the following experiments, listeners were University of Nevada, Las Vegas (UNLV) undergraduates who received course credit for their participation. Sample size was chosen to be similar to or larger than what was used in other studies investigating scene size and scene delay (ISI) in an auditory change-detection paradigm (e.g., Demany et al., 2008). All participants provided informed consent according to a protocol approved by the UNLV Office for Research Integrity.

Stimuli The naturalistic sounds consisted of recognizable and unrecognizable environmental sounds. The recognizable sounds consisted of 15 common environmental sounds (e.g., a drill, drumming, a dog barking; see the [Appendix](#) for a complete list). The sounds were rated as highly recognizable in a previous study (Gregg, Irsik, & Snyder, 2014). The duration of each sound was 1,000 ms. All sounds were digitized to a sampling rate of 44.1 kHz, matched for RMS amplitude, filtered for noise, and a linear off-ramp to zero amplitude was imposed over the final 10 ms to avoid abrupt offsets. We chose not to impose on-ramps to better control the synchrony of sounds within a scene. All sounds were carefully examined to ensure there were no abrupt onsets. All stimuli were presented at a comfortable listening level (approximately 70 dB) in a sound attenuated chamber.

The stimuli were digitally combined to create unique scenes consisting of two, four, or six objects. Some acoustic properties of the scenes were equated by keeping the average acoustic spread among objects in the scenes similar. The average acoustic spread was created by calculating the Euclidean distance between pairs of stimuli based on two acoustic properties: fundamental frequency and harmonicity. We calculated the average acoustic spread for objects in both Scene 1 and Scene 2 by calculating the Euclidean distance between each pair of stimuli within each scene (all combinations of four objects resulted in six pairs) and then calculating the average acoustic distance between all pairs. The result was an average acoustic spread within Scenes 1 and 2. To obtain the acoustic measurements, each stimulus was submitted to Praat (Boersma & Weenink, 1992) for analysis of harmonicity (i.e., mean amount of acoustic periodicity in the signal

measured as the ratio between the power of harmonics of the fundamental frequency to the power of nonharmonic components) and fundamental frequency (Gregg & Samuel, 2008). These two properties are particularly important for sound segregation: frequency is important given the tonotopic organization of the auditory system and is well-established as a strong cue to auditory scene analysis (e.g., Bregman, 1990); harmonicity is also a strong cue to auditory scene analysis (Yost & Sheft, 1993) and sounds have been found to be automatically and preattentively assigned to categories of periodic/aperiodic (Kat & Samuel, 1984). See also Gygi, Kidd, and Watson (2007) for the importance of these two dimensions.

A set of 15 unrecognizable sounds was created from the set of recognizable sounds. We achieved this by submitting each sound to a custom program created in MATLAB. The program split each sound into fifty 20-ms chunks, randomized the order of the chunks, and then connected the randomized chunks together into a new 1-s sound. The sounds were rated as reliably unrecognizable, as reported in a previous study (Gregg et al., 2014). As with the recognizable sounds, the unrecognizable stimuli were combined to create unique scenes with two, four, or six sounds that were equated in acoustic spread using the fundamental Frequency \times Harmonicity space (fundamental frequency and harmonicity were remeasured from the scrambled sounds and those measurements were used to create a two-dimensional Euclidean space).

The artificial sounds consisted of simple pure tone rhythms and more spectrally complex noise burst rhythms. A set of 15 band-pass filtered white noise rhythms were created in Praat (Boersma & Weenink, 1992). Different rhythms were created by combining a series of noise bursts that were short (1/48 s), medium (1/24 s), or long (1/12 s) in duration. The noise bursts within a rhythm were interrupted by intervals of silence that were short (1/48 s), medium (1/24 s), or long (1/12 s) in duration. These parameters were adopted from a change-deafness study using similar stimuli (Puschmann et al., 2013). The noise bands were centered at frequencies of 200, 400, 800, 1600, 3200 Hz, or 6400 Hz, with bandwidths set to 25% of the center frequency. Each 1,000-ms noise rhythm was used to create scenes, each with a unique noise duration, silent interval duration, and frequency. The rhythms were combined to create unique scenes consisting of two, four, or six sounds. Within each scene, the noise rhythms were matched for loudness using loudness level contours (Fletcher & Munson, 1933). As with the natural sounds, the acoustic spread of the scenes was equated.

A set of 15 pure-tone rhythms were created in Praat (Boersma & Weenink, 1992) using the same durations, intervals, and frequencies as the noise rhythms, the only difference being that the noises were replaced by pure (i.e., sinusoidal) tones. The tone rhythms were combined to create unique scenes that were equated for difficulty and corrected for loudness as described above for the noise rhythms.

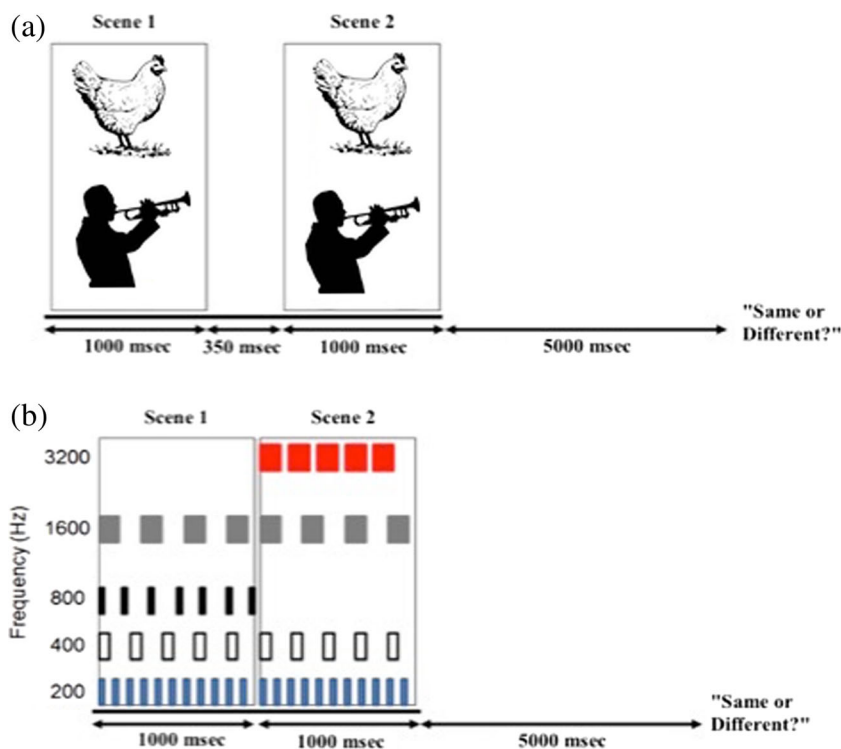


Fig. 1 Trial schematic of the change-detection task. **a** An example of a *same* trial, composed of recognizable sounds with a scene size of 2 and an ISI of 350 ms. **b** An example of a *different* trial composed of pure tones, with a scene size of 4 and ISI of 0 ms. (Color figure online)

Procedure Twelve different listeners completed the change-detection task for each sound type: recognizable sounds, unrecognizable sounds, artificial noise rhythms, and artificial pure tone rhythms (the number of participants in each condition was selected to be comparable to previous investigations of memory loss in change deafness (see Demany et al., 2008; Pavani & Turatto, 2008). Figure 1 depicts example trials of the task. Listeners were presented on each trial with Scene 1, a multiple-object scene in which the sounds were presented simultaneously for 1,000 ms binaurally through Sennheiser HD 280 Pro headphones. Scene 1 was followed by Scene 2, a 1,000-ms scene consisting of either the same sounds as Scene 1 (*same* trial) or all but one of the same sounds as Scene 1 and one new sound (*different* trial). Three different levels of Scene Size (two, four, or six objects) were orthogonally combined with four different levels of interscene interval (ISI: 0; 350; 750; and 2,000 ms). Each of the 12 combinations was presented in a separate block, resulting in 12 total blocks and 384 trials total. There were 32 trials within each block, 16 same trials and 16 different trials. Each Scene 1 was unique in terms of the combination of sounds from the set of 15 that were used. Listeners were instructed to indicate by button press whether Scene 1 and Scene 2 were the “same” or “different.” There was a 5,000-ms intertrial interval (ITI, the time from the offset of Scene 2 to the onset of Scene 1 of the next trial). Listeners performed 12 practice trials (one of each combination of scene size and ISI) before beginning the experiment.

Data analysis In all experiments, responses were used to calculate the proportion of hits (responding “Different” on *different* trials) and false alarms (responding “Different” on *same* trials). The differencing strategy was used to obtain d' scores (see Appendix A5.4 in Macmillan & Creelman, 2005). Prior to obtaining d' , any conditions having proportions of either a 0 or 1 for false alarms or hits were corrected by replacing 0 and 1 values with $1/(2N)$ and $1 - 1/(2N)$, respectively, where N equals the total number of trials on which a proportion was based (Macmillan & Kaplan, 1985). Hit and false-alarm rates are reported in Table 1. The d' scores were submitted to a three-way ANOVA, with sound type (recognizable, unrecognizable, noise, or tones) as a between-subjects factor and scene size (two, four, or six sounds) and ISI (0; 350; 750; or 2,000 ms) as within-subject factors. Pair-wise comparisons on main effects used the least-significant difference adjustment. Planned linear contrasts were used to follow up significant interactions.

Results and discussion

Change-detection performance differed by sound type, $F(3, 44) = 3.77$, $p = .017$, $\eta_p^2 = .20$. Change detection was significantly worse for the naturalistic sounds: planned comparisons indicated significantly more change detection errors (i.e., change deafness) in response to unrecognizable and recognizable sounds than to the noises or tones ($ps < .05$).

Table 1 Hit and false-alarm rates from Experiment 1

	Recognizable	Unrecognizable	Noise	Tones
0 ms ISI				
2 sounds				
Hits	0.90	0.89	0.97	0.95
False alarms	0.08	0.17	0.07	0.03
4 sounds				
Hits	0.80	0.71	0.95	0.88
False alarms	0.05	0.12	0.12	0.07
6 sounds				
Hits	0.55	0.53	0.82	0.81
False alarms	0.06	0.17	0.09	0.08
350 ms ISI				
2 sounds				
Hits	0.89	0.86	0.96	0.93
False alarms	0.05	0.05	0.05	0.04
4 sounds				
Hits	0.77	0.66	0.86	0.74
False alarms	0.04	0.05	0.06	0.04
6 sounds				
Hits	0.62	0.45	0.65	0.71
False alarms	0.08	0.04	0.07	0.04
750 ms ISI				
2 sounds				
Hits	0.91	0.87	0.95	0.91
False alarms	0.07	0.06	0.07	0.07
4 sounds				
Hits	0.76	0.64	0.85	0.74
False alarms	0.09	0.07	0.11	0.07
6 sounds				
Hits	0.55	0.47	0.64	0.64
False alarms	0.07	0.05	0.13	0.08
2,000 ms ISI				
2 sounds				
Hits	0.89	0.89	0.96	0.92
False alarms	0.09	0.05	0.06	0.06
4 sounds				
Hits	0.74	0.63	0.84	0.70
False alarms	0.05	0.09	0.1	0.06
6 sounds				
Hits	0.56	0.54	0.6	0.63
False alarms	0.07	0.06	0.1	0.08

Change detection across all sound types was affected by the number of sound objects, which indicates that capacity limitations have a general effect on change detection. As can be seen in Fig. 2, change-detection performance decreased as the number of sound objects increased, $F(2, 88) = 226.26, p < .001, \eta_p^2 = .84$. Planned comparisons indicated that change detection was worse when the scene size was six sounds

(mean $d' = 3.43, SD = 0.95$) than when the scene size was four sounds (mean $d' = 4.24, SD = 1.07, p < .05$), and two sounds (mean $d' = 5.57, SD = 1.14, p < .05$).

Change detection was also affected by the delay between scenes: there was a significant effect of ISI, $F(3, 132) = 16.89, p < .001, \eta_p^2 = .28$. This finding suggests that information loss over time contributes to change detection errors; however, an interaction between ISI and sound type, $F(9, 132) = 4.89, p < .001, \eta_p^2 = .25$, revealed that information loss only affected change detection to certain sound types. Planned linear contrasts indicated that the ISI significantly influenced change detection only to the artificial sounds (i.e., tones and noise rhythms). Change-

detection performance to the scenes composed of tones and noise rhythms was worse at the 750-ms and 2,000-ms ISIs than at the 0-ms and 350-ms ISIs ($ps < .05$). For the recognizable and unrecognizable sounds, there were no significant differences in change-detection performance across the four ISIs. The three-way interaction among ISI, sound type, and scene size was not significant, $F(18, 264) = 1.03, p = .431, \eta_p^2 = .06$. Thus, although capacity limitations seem to have a general effect on change detection, the role of information loss in change detection seems to be specific to artificial sounds, and more complex environmental sounds appear to be relatively resistant to information loss over time. This latter finding is admittedly restricted to a time course spanning 2,000 ms. Experiment 2 was conducted to examine potential information loss over a longer time period.

Experiment 2

Experiment 1 demonstrated that change deafness to complex, naturalistic sounds is not due to information loss over time, with delays ranging from 0 ms to 2,000 ms. Experiment 2 was conducted to examine whether change-detection performance for such sounds remains unaffected by longer delays between scenes. In this experiment, we determine whether information is lost during a change-detection task when the delay is as long as 6,000 ms.

Method

Participants Fifty-two listeners with normal hearing participated in this experiment (28 females and 24 males; mean age = 20.32 years, range: 18–35 years). Data from four participants were not included in analyses due to prior health history ($n = 1$, brain tumor), technical malfunction ($n = 2$), or because they pushed buttons randomly ($n = 1$). A total of 48 participants (12 in each condition) were included in statistical analyses.

Stimuli Stimuli were identical to Experiment 1.

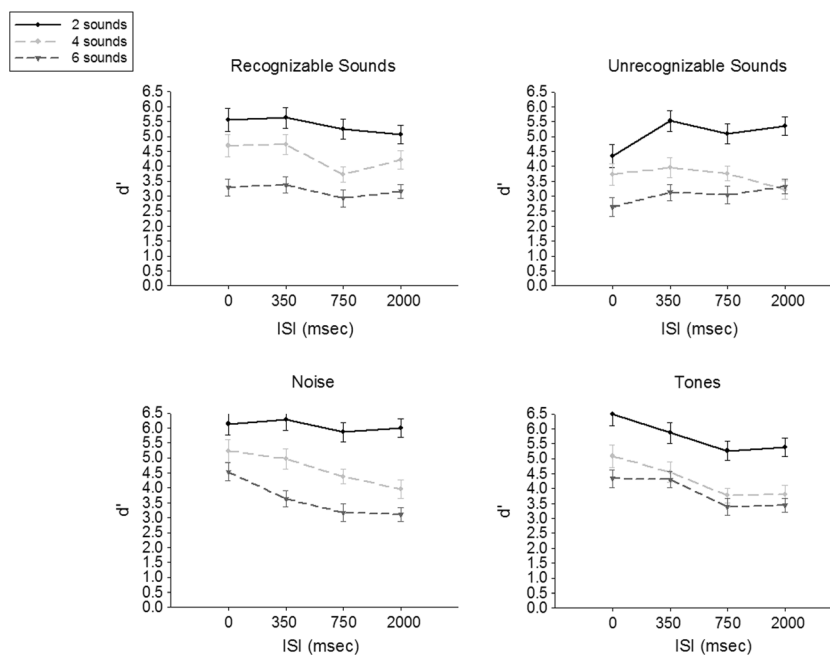


Fig. 2 Change-detection performance to all sound types in Experiment 1 (measured by d'). Error bars represent the standard error of the mean

Procedure The procedure was identical to Experiment 1, except different ISIs were used (0; 100; 750; 6,000 instead of 0; 350; 750; 2,000 ms).

Results and discussion

Hit and false-alarm rates from Experiment 2 are reported in Table 2. As in Experiment 1, change-detection performance was worse for the naturalistic sounds. This was indicated by an effect of sound type, $F(3, 44) = 14.36, p < .001, \eta_p^2 = .50$, as well as planned comparisons indicating more change-detection errors in response to unrecognizable and recognizable sounds than to the noises and tones ($ps < .05$).

Also consistent with Experiment 1, change detection across all sound types was affected by the number of sound objects, indicating again that capacity limitations have a general effect on change detection (see Fig. 3). This finding was indicated by an effect of scene size, $F(2, 88) = 414.46, p < .001, \eta_p^2 = .90$. Planned comparisons indicated that change-detection performance was lower when the scene size was six sounds (mean $d' = 3.09, SD = 0.71$) than when scene size was four sounds (mean $d' = 3.93, SD = 0.78, p < .05$), and two sounds (mean $d' = 4.91, SD = 0.62, ps < .05$).

Change detection was also affected by the delay between scenes, as indicated by a significant effect of ISI. Mauchly’s test indicated that the assumption of sphericity had been violated, therefore p values were corrected using Greenhouse–Geisser estimates of sphericity: $F(3, 132) = 144.36, p < .001, \eta_p^2 = .77$. There was also an interaction between ISI and sound type, $F(9, 132) = 3.94, p < .001, \eta_p^2 = .21$, and a three-way interaction between ISI, sound type, and scene size,

$F(18, 264) = 2.12, p = .006, \eta_p^2 = .13$. Planned linear contrasts on the three-way interaction indicated that the interaction was driven by the marked drop in change-detection performance at the 6,000-ms ISI, especially at larger scene sizes (see Fig. 3). For all sound types and all scene sizes, performance was significantly worse during the 6,000-ms delay than during the shorter delays (0, 100, and 750 ms). This finding suggests that the effect of ISI is more general across sound types when longer delay intervals are examined. However, one potential problem in Experiment 2, that was not observed in Experiment 1, was a high false-alarm rate for trials with a 6,000-ms ISI (mean = 23.9%, compared to a false alarm rate of 11% for the longest ISI in Experiment 1). This could have occurred because the intertrial interval for long ISI trials was actually shorter than the delay interval between scenes (ITI was held constant at 5,000 ms). This may have caused Scene 2 to become paired in memory with Scene 1 of a new subsequent trial as a result of closer temporal proximity (cf. Cowan, Saults, & Nugent, 1997), which would in turn cause interference during discrimination judgments. Therefore, Experiment 3 was conducted to address this concern.

Experiment 3

Experiment 3 was conducted to further explore the general effect of the delay between scenes found in Experiment 2. Specifically, this experiment allowed us to determine whether the decrease in change-detection performance at longer delays for all sound types was a result of information loss rather than an artifact of across-trial interference. In Experiment 3, we

Table 2 Hit and false-alarm rates from Experiment 2

	Recognizable	Unrecognizable	Noise	Tones
0 ms ISI				
2 sounds				
Hits	0.96	0.86	0.96	0.93
False alarms	0.05	0.07	0.04	0.07
4 sounds				
Hits	0.87	0.69	0.93	0.90
False alarms	0.04	0.07	0.05	0.07
6 sounds				
Hits	0.73	0.50	0.82	0.77
False alarms	0.05	0.08	0.05	0.08
350 ms ISI				
2 sounds				
Hits	0.96	0.89	0.95	0.93
False alarms	0.04	0.06	0.04	0.04
4 sounds				
Hits	0.85	0.64	0.84	0.85
False alarms	0.04	0.05	0.04	0.06
6 sounds				
Hits	0.69	0.41	0.74	0.73
False alarms	0.06	0.08	0.04	0.04
750 ms ISI				
2 sounds				
Hits	0.94	0.85	0.96	0.94
False alarms	0.05	0.08	0.04	0.05
4 sounds				
Hits	0.86	0.63	0.79	0.69
False alarms	0.04	0.06	0.04	0.03
6 sounds				
Hits	0.67	0.41	0.65	0.64
False alarms	0.06	0.07	0.04	0.03
2,000 ms ISI				
2 sounds				
Hits	0.95	0.82	0.94	0.86
False alarms	0.10	0.17	0.17	0.17
4 sounds				
Hits	0.80	0.61	0.76	0.68
False alarms	0.10	0.21	0.26	0.17
6 sounds				
Hits	0.60	0.56	0.56	0.62
False alarms	0.21	0.21	0.27	0.20

modified the paradigm used in Experiment 2 to better control the ISI:ITI ratios.

Method

Participants Fifty-two listeners with normal hearing participated in this experiment (39 females and 13 males; mean age

= 20.76 years, range: 18–35 years). Data from four participants were not included in analyses due to prior health history ($n = 3$, multiple sclerosis, head injury, and seizures), or because they pushed buttons randomly ($n = 1$). A total of 48 participants (12 per sound type) were included in statistical analyses.

Stimuli Stimuli were identical to Experiment 1.

Procedure The procedure was nearly identical to Experiment 1. In this experiment, different ISIs were used: 0 ms; 100 ms; 1,500 ms; 6,000 ms. Intertrial intervals (ITI) were also changed in order to retain temporal distinctiveness of Scene 2 to Scene 1 of a following trial, and to have constant ISI:ITI ratios for at least two ISI conditions. The ITI was set to 2,000 ms for ISI conditions with shorter delays (0; 100; 1,500 ms), and 8,000 ms for the longest ISI condition (6,000 ms), allowing the 1,500-ms and 6,000-ms conditions to have an equivalent ISI:ITI ratio (3:4). It was not possible to have the same ratios for the two shortest ISIs because the resulting ITIs would not be long enough for participants to respond.

Results and discussion

Hit and false-alarm rates are reported in Table 3. Consistent with Experiments 1 and 2, change-detection performance was worse for the naturalistic sounds: There was a significant effect of sound type, $F(3, 44) = 3.69$, $p = .019$, $\eta_p^2 = .20$. Planned comparisons indicated lower change-detection performance in response to unrecognizable and recognizable sounds than to the noises and tones ($ps < .05$).

Change detection across all sound types was affected by the number of sound objects, indicating again that capacity limitations have a general effect on change detection. As can be seen in Fig. 4, change-detection performance decreased as the number of sound objects increased, $F(2, 88) = 248.45$, $p < .001$, $\eta_p^2 = .85$. Planned comparisons indicated that change-detection performance was lower when the scene size was six sounds (mean $d' = 2.93$, $SD = 0.8$) than when the scene size was four sounds (mean $d' = 3.75$, $SD = 0.79$, $p < .05$), and two sounds (mean $d' = 4.70$, $SD = 0.73$, $p < .001$).

As in Experiment 1, change detection was affected by the delay between scenes composed of artificial sounds and less so to natural, complex sounds: this was indicated by a significant effect of ISI, $F(3, 132) = 80.71$, $p < .001$, $\eta_p^2 = .65$, an interaction between ISI and sound type, $F(9, 132) = 4.94$, $p < .001$, $\eta_p^2 = .25$, and a three-way interaction between ISI, sound type, and scene size, $F(18, 264) = 1.73$, $p = .034$, $\eta_p^2 = .11$. Planned comparisons to explain the three-way interaction indicated that change-detection performance during the recognizable and unrecognizable scenes was relatively similar across the three shorter ISIs (0; 100; and 1,500 ms), but dropped significantly at the 6,000-ms ISI ($ps < .05$).

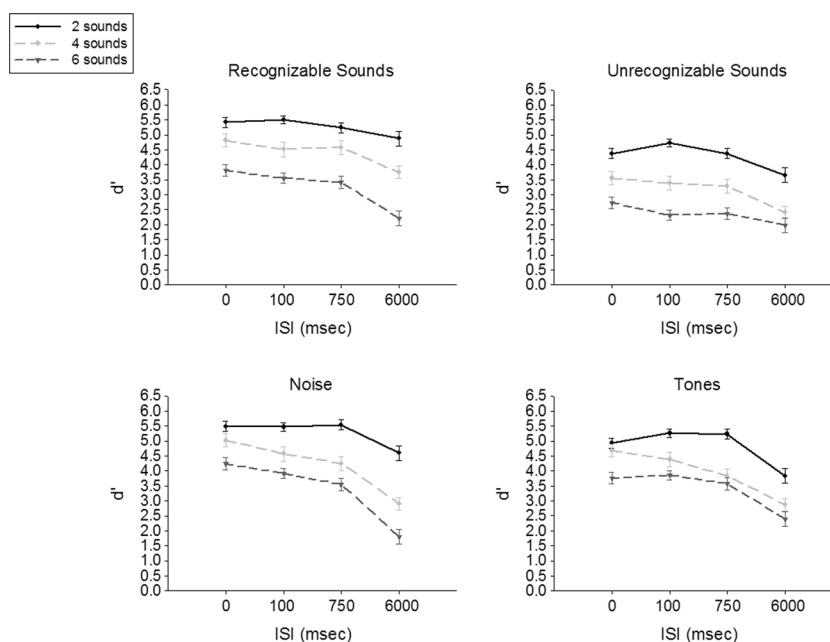


Fig. 3 Change-detection performance to all sound types in Experiment 2 (measured by d'). Error bars represent the standard error of the mean

Change-detection performance during the scenes composed of noise and tone rhythms dropped more linearly as ISI increased, (d' at 0 and 100 ms was higher than at 1,500 and 6,000 ms, $p < .05$, and d' at 1,500 ms was higher than at 6,000 ms, $p < .05$).

General discussion

In this study, we examined the contribution of capacity limitations and information loss over time to change deafness during scenes composed of four different sound types: recognizable environmental sounds, unrecognizable environmental sounds, tone rhythms, and noise rhythms. The results of the present study consistently revealed that a capacity limitation contributes to change deafness for all sound types. This finding is consistent with a large body of perceptual research demonstrating that attention, perception, and memory are all limited by the number of objects that can be simultaneously processed. For example, our ability to consciously detect auditory targets in complex backgrounds is limited by the number of competing sounds, a phenomenon referred to as informational masking (Dickerson & Gaston, 2014; Durlach et al., 2003a, b; Lutfi, Chang, Stamas, & Gilbertson, 2012). Though the existence of a fixed limit in working memory has been challenged (see Brady et al., 2016; van den Berg, Awh, & Ma, 2014), the results of the present study indicate a novel difference between auditory and visual working memory that should be noted. Recent evidence in the visual domain suggests that visual working memory has a larger capacity for naturalistic objects compared to artificial objects (Brady

et al., 2016). Our results suggest that the capacity of auditory memory is similar for naturalistic and artificial sounds (though the duration that these stimuli can be maintained in auditory memory does differ, as we discuss below). One important endeavor for future research is to further explore this potential difference in the way that auditory memory and visual memory store information, the role of stimulus complexity in storage, and the stages of processing in which the capacity limitation in change deafness (and change blindness) arises.

The results of the present study suggest that change deafness occurs not only because of limitations in auditory memory processing capacity but also because of loss of information in memory, especially for artificial sounds. Change detection of the recognizable and unrecognizable environmental sounds did not vary much as a function of the delay interval between scenes if the interval was 2,000 ms or less. Only when the interval between scenes was extended to 6,000 ms did change-detection performance to the environmental sounds decline substantially (in Experiments 2 and 3). Change detection of the artificial sounds (i.e., tones and noise bursts) was more affected by the increasing interval between scenes, and this was despite the fact that overall change-detection performance was better for artificial sounds. The differences in detection of changes to artificial and environmental sounds suggest that change detection for environmental, spectrally complex sounds has access to memory mechanisms that are more persistent than those for artificial sounds. This finding is somewhat at odds with other discrimination or segregation studies that show a more gradual decline in auditory short-term or implicit memory over long intervals for artificial sounds; however, performance during these studies may have been aided by the use of even simpler sounds

Table 3 Hit and false-alarm rates from Experiment 3

	Recognizable	Unrecognizable	Noise	Tones
0 ms ISI				
2 sounds				
Hits	0.87	0.90	0.95	0.95
False alarms	0.05	0.06	0.11	0.05
4 sounds				
Hits	0.73	0.71	0.92	0.86
False alarms	0.08	0.07	0.08	0.06
6 sounds				
Hits	0.52	0.43	0.77	0.78
False alarms	0.07	0.06	0.10	0.04
350 ms ISI				
2 sounds				
Hits	0.89	0.92	0.94	0.93
False alarms	0.06	0.03	0.05	0.04
4 sounds				
Hits	0.76	0.69	0.86	0.83
False alarms	0.04	0.04	0.08	0.03
6 sounds				
Hits	0.60	0.43	0.61	0.77
False alarms	0.06	0.04	0.04	0.06
750 ms ISI				
2 sounds				
Hits	0.90	0.87	0.92	0.89
False alarms	0.07	0.03	0.07	0.04
4 sounds				
Hits	0.71	0.61	0.80	0.66
False alarms	0.06	0.04	0.06	0.05
6 sounds				
Hits	0.54	0.37	0.50	0.60
False alarms	0.07	0.04	0.07	0.04
2000 ms ISI				
2 sounds				
Hits	0.90	0.86	0.94	0.88
False alarms	0.10	0.20	0.19	0.24
4 sounds				
Hits	0.68	0.67	0.78	0.74
False alarms	0.13	0.16	0.32	0.19
6 sounds				
Hits	0.63	0.52	0.64	0.74
False alarms	0.17	0.18	0.36	0.29

perceived as one (McKeown & Mercer, 2012; Mercer & McKeown, 2014) or two objects (Snyder & Weintraub, 2013). Given that the memory load would be minimal for comparing so few objects, it may have been easier to maintain stimulus details in memory for longer periods of time. This is consistent with our findings, as listeners had significant difficulty comparing larger scene sizes when the delay interval

was long. One important endeavor for future research will be to determine why there are differences in the way auditory memory retains environmental and artificial sounds. For example, environmental sounds may recruit larger neural populations than artificial sounds, which could contribute to the differences in representational strength. A related possibility is how the varying degrees of spectral overlap in environmental and artificial sounds contribute to change deafness.

The more surprising finding of this study was that change deafness to environmental sounds was not affected much by the delay interval between scenes when the interval was as long as 2,000 ms. We suggest that although auditory short-term memory capacity is limited, memory for naturalistic sounds is remarkably enduring, and information loss over time (e.g., due to decay, interference, or sudden death) does not cause change deafness to natural sounds up to 2,000 ms. Given the present results, efforts to improve auditory change detection in natural settings should boost processing capacity more so than the ability to retain information over short amounts of time. It is worth noting that the duration of each scene in this study was held constant at 1,000 ms. There are studies suggesting different types of memory-encoding strategies in response to short stimuli at short ISIs than to long stimuli at long ISIs (McDermott, Schemitsch, & Simoncelli, 2013) as well as evidence that scene length can affect change deafness (Eramudugolla, McAnally, Martin, Irvine, & Mattingley, 2008; McAnally et al., 2010). It will be important for future research to determine how information loss is affected in a change-detection task when both scene duration and ISI are manipulated. Future studies should also determine if information loss over time affects change blindness; to our knowledge, the delay between scenes has not been systematically manipulated in a change-blindness paradigm.

Information loss to all sound types was apparent when the delay interval between scenes was extended to 6,000 ms; this loss was particularly large at larger scene sizes. The constant ITI ratio in Experiment 3 suggests that the drop in successful change-detection performance from 2,000 to 6,000 ms was a result of true information loss, through decay or sudden death, rather than interference. Recent research on auditory memory suggests that one potential reason for the information loss is an inverse relationship between the number of objects in auditory working memory and the fidelity of each object representation (Joseph, Kumar, Husain, & Griffiths, 2015b). If this were the case in the present study, then increasing the delay between scenes would be more detrimental to performance as scene size increases because the quality of the object representations at scene sizes of six sounds would be poorer than representations of two or four sound objects. Other recent work on auditory memory has revealed that memory is better when an integrated auditory object must be held in auditory memory rather than auditory features (Joseph, Iverson, et al., 2015). This finding could explain why increasing the delay interval

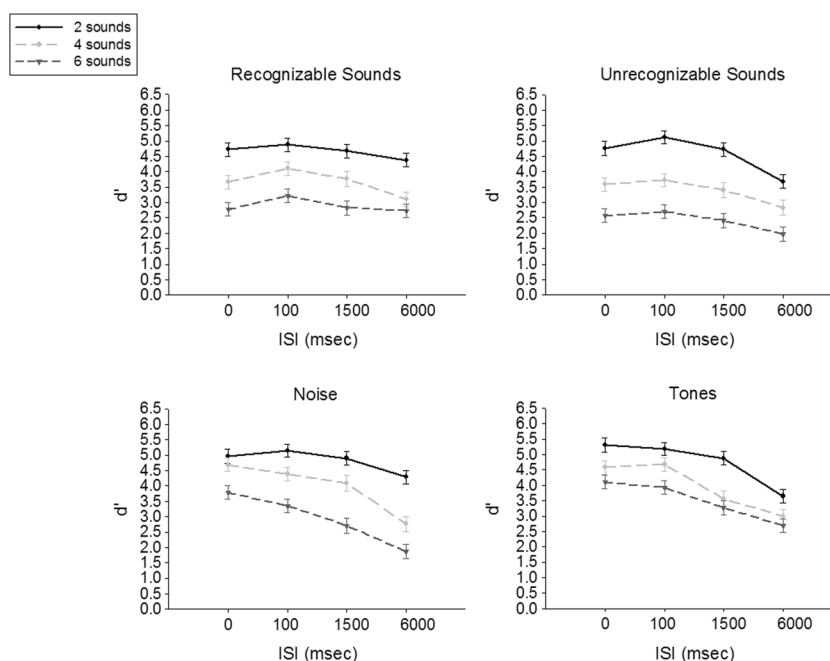


Fig. 4 Change-detection performance to all sound types in Experiment 3 (measured by d'). Error bars represent the standard error of the mean

was more detrimental to performance when the sounds were simple noise and tone rhythms, as it is possible that listeners are more likely to consider natural sounds as auditory “objects,” even if they are unrecognizable. Also, the complexity and multiple segregation cues in the auditory objects could create more durable representations. It should be noted that while listeners were better able to hold environmental sounds in memory relative to artificial sounds, they had more difficulty remembering unrecognizable relative to recognizable environmental sounds over long intervals. It is possible that the scrambled temporal structure of unrecognizable sounds made it difficult to appropriately group acoustic details and form objects on more difficult trials, motivating listeners at times to focus on auditory features.

The results of the present investigation indicate a potential difference in the auditory and visual memory processes that support change detection. A well-demonstrated finding in vision is accurate change-detection performance at very short delays between scenes (less than 100 ms) that is not affected by the number of objects in the display. However, once the delay between scenes exceeds 100 ms, change-detection performance begins to decline as the number of objects within the scenes increase (Phillips, 1974). Better change-detection performance at short delays presumably reflects an unlimited-capacity, short-duration sensory memory; meanwhile, performance at longer delays that declines with increasing scene sizes reflects a limited-capacity, long-duration visual working memory system (Luck & Vogel, 1997).

In addition, however, a similar interaction does not occur. There was no interaction between scene size and the delay interval between scenes in Experiments 1 and 3 of this study:

Demany et al. (2008) also failed to find an interaction between scene size and the delay interval between scenes in several auditory change-detection experiments. The different pattern of results in vision and audition do not necessarily mean that there are fundamental differences in the visual and auditory memory processes that support change detection. For example, it is possible that visual and auditory change-detection mechanisms are essentially the same, or at least analogous, but different patterns emerge because auditory sensory memory is longer lasting than visual sensory memory (Demany et al., 2008). The difference in auditory and visual sensory memory is well-suited for the nature of auditory and visual stimuli. Sounds are quite transient and need to be held in memory in order to be temporally integrated with subsequent sounds. Without this ability, it would be nearly impossible to understand a spoken sentence, or to organize and integrate continuous sounds as coming from a single source. Temporal integration of auditory information requires a long-duration sensory storage, but visual objects can typically be viewed for extended periods of time, making a long sensory storage unnecessary.

An additional possibility is that there is not as sharp a distinction between auditory sensory and short-term/working memory (e.g., Jones, Hughes, & Macken, 2006; Nicholls & Jones, 2002) as there is in vision. Another issue to consider is that visual stimuli in change-detection paradigms are usually static, unlike sounds, which are dynamic. It will be important for future visual change-detection studies to compare performance to static (e.g., pictures) and dynamic (e.g., videos) stimuli to further explore the issue of whether auditory and visual change-detection processes are similar.

In summary, change deafness to environmental sounds is largely due to a capacity limitation and not to loss of memory, except when using intervals of a few seconds or more. Change deafness to simple, artificial sounds, however, is caused by capacity limitations and loss of memory over time, even for short intervals between scenes. Previous investigations of change deafness have limited manipulations to only artificial sounds (e.g., Cervantes Constantino, Pinggera, Paranamana, Kashino, & Chait, 2012) or to only naturalistic sounds (e.g., Vitevitch, 2003). The present investigation is the first study, to our knowledge, to directly compare change-detection performance to spectrally complex environmental sounds and spectrally simple artificial sounds: a comparison that allowed us to address the extent to which memory loss occurs in change-deafness tasks with different stimuli. The differences in change-detection performance to the different sound types found in this experiment demonstrate the importance of using multiple sound types to fully understand the mechanisms underlying change deafness.

Appendix

List of recognizable environmental sounds used in all experiments

- Dog barking
- Chant
- Man coughing
- Baby crying
- Door creaking
- Drum beat
- Spoon hitting a frying pan
- Lighting a match
- Motorcycle engine
- Owl hoot
- Footsteps on stairs
- Rocking chair
- Knife sharpening
- Train
- Tuba melody

Author note This work was supported by the Army Research Office under Award No. W9IINF-I2-I-0256. We would like to thank Sebastian Puschmann for his advice on the creation of our noise and tone stimuli.

References

- Backer, K. C., & Alain, C. (2012). Orienting attention to sound object representations attenuates change deafness. *Journal of Experimental Psychology: Human Perception and Performance*, 38(6), 1554–1566. doi:<https://doi.org/10.1037/a0027858>
- Boersma, P., & Weenink, D. (1992). Praat: Doing phonetics by computer (Version 4.3.2) [Computer software and manual]. Retrieved from www.praat.org.
- Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences* 113, 7459–7464.
- Braida, L. D., Lim, J. S., Berliner, J. E., & Durlach, N. I. (1984). Intensity perception: XIII. Perceptual anchor model of context?coding. *Journal of the Acoustical Society of America* 76, 722–732.
- Bregman, A. S. (1990). Auditory scene analysis. Cambridge, MA: MIT Press.
- Cervantes Constantino, F., Pinggera, L., Paranamana, S., Kashino, M., & Chait, M. (2012). Detection of appearing and disappearing objects in complex acoustic scenes. *PLOS ONE*, 7(9). doi:<https://doi.org/10.1371/journal.pone.0046167>
- Cohen, M. A., Horowitz, T. S., & Wolfe, J.M. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences* 106, 6008–6010.
- Cowan, N., Saults, S., & Nugent, L. D. (1997). The role of absolute and relative amounts of time in forgetting within immediate memory: The case of tone-pitch comparisons. *Psychonomic Bulletin & Review*, 4, 393–397.
- Demany, L., Trost, W., Serman, M., & Semal, C. (2008). Auditory change detection: Simple sounds are not memorized better than complex sounds. *Psychological Science*, 19(1), 85–91. doi:<https://doi.org/10.2307/40064804>
- Dickerson, K., & Gaston, J. R. (2014). Did you hear that? The role of stimulus similarity and uncertainty in auditory change deafness. *Frontiers in Psychology*, 5, 1125. doi:<https://doi.org/10.3389/fpsyg.2014.01125>
- Durlach, N. I., Mason, C. R., Kidd, G., Jr., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. G. (2003a). Note on informational masking. *Journal of the Acoustic Society of America*, 113(6), 2984–2987.
- Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., & Kidd, G. (2003b). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *The Journal of the Acoustical Society of America*, 114(1), 368–379. doi:<https://doi.org/10.1121/1.1577562>
- Eramudugolla, R., Irvine, D. R. F., McAnally, K. I., Martin, R. L., & Mattingley, J. B. (2005). Directed attention eliminates ‘change deafness’ in complex auditory scenes. *Current Biology*, 15(12), 1108–1113.
- Eramudugolla R., McAnally, K. I., Martin, R. L., Irvine, D. R., Mattingley, J. B. (2008). The role of spatial location in auditory search. *Hearing Research*, 238, 139–146.
- Fletcher, H., & Munson, W. A. (1933). Loudness, its definition, measurement and calculation. *The Journal of the Acoustical Society of America* 5(2), 82–108. doi:<https://doi.org/10.1121/1.1915637>
- Gregg, M. K., Irsik, V. C., & Snyder, J. S. (2014). Change deafness and object encoding with recognizable and unrecognizable sounds. *Neuropsychologia*, 61(0), 19–30. doi:<https://doi.org/10.1016/j.neuropsychologia.2014.06.007>
- Gregg, M. K., & Samuel, A. G. (2008). Change deafness and the organizational properties of sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 34(4), 974–991. doi:<https://doi.org/10.1037/0096-1523.34.4.974>
- Gygi, B., Kidd, G., & Watson, C. (2007). Similarity and categorization of environmental sounds. *Perception & Psychophysics*, 69(6), 839–855. doi:<https://doi.org/10.3758/BF03193921>
- Jones, D., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a

- perceptual-gestural view of short-term memory. *Journal of Memory and Language*, 54, 265–281.
- Joseph, S., Iverson, P., Manohar, S., Fox, Z., Scott, S., & Husain, M. (2015a). Precision of working memory for speech sounds. *Quarterly Journal of Experimental Psychology*, 11, 1–19. doi:<https://doi.org/10.1080/17470218.2014.1002799>
- Joseph, S., Kumar, S., Husain, M., & Griffiths, T. D. (2015b). Auditory working memory for objects vs. features. *Frontiers in Neuroscience*, 9, 13. doi:<https://doi.org/10.3389/fnins.2015.00013>
- Kat, D., & Samuel, A. G. (1984). More adaptation of speech by non-speech. *Journal of Experimental Psychology: Human Perception & Performance*, 10, 512–525.
- Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, 1(3), 153.
- Li, D., Cowan, N., & Sauls, J. S. (2013). Estimating working memory capacity for lists of nonverbal sounds. *Attention, Perception, & Psychophysics*, 75(1), 145–160. doi:<https://doi.org/10.3758/s13414-012-0383-z>
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–281.
- Lutfi, R. A., Chang, A. C., Stamas, J., & Gilbertson, L. (2012). A detection-theoretic framework for modeling informational masking. *Journal of the Acoustic Society of America*, 132(2), EL109–E113. doi:<https://doi.org/10.1121/1.4734575>
- Macmillan, N. A., & Creelman, C. D. (2005). Detection theory: A user's guide (2nd ed.). Mahwah, NJ: Erlbaum.
- Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false alarm rates. *Psychological Bulletin*, 98, 185–199.
- McAnally, K. I., Martin, R. L., Eramudugolla, R., Stuart, G. W., Irvine, D. R., & Mattingley, J. B. (2010). A dual-process account of auditory change detection. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 994–1004.
- McDermott, J. H., Schemitsch, M., & Simoncelli, E. P. (2013). Summary statistics in auditory perception. *Nature Neuroscience*, 16, 493–498.
- Mercer, T., & Mckeown, D. (2014). Decay uncovered in nonverbal short-term memory. *Psychonomic Bulletin & Review*, 21, 128–135.
- Mckeown, D., & Mercer, T. (2012). Short-term forgetting without interference. *Journal of Experimental Psychology*, 38(4), 1057–1068.
- Murray, M. M., Camen, C., Gonzalez Andino, S. L., Bovet, P., & Clarke, S. (2006). Rapid brain discrimination of sounds of objects. *Journal of Neuroscience*, 26(4), 1293–1302.
- Nicholls A. P., & Jones D. M. (2002). Capturing the suffix: Cognitive streaming in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28,(1), 12–28.
- Pavani, F., & Turatto, M. (2008). Change perception in complex auditory scenes. *Perception & Psychophysics*, 70(4), 619–629. doi:<https://doi.org/10.3758/PP.70.4.619>
- Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58(3), 193–198. doi:<https://doi.org/10.1037/h0049234>
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, 16(2), 283–290. doi:<https://doi.org/10.3758/BF03203943>
- Puschmann, S., Sandmann, P., Ahrens, J., Thorne, J., Weerd, R., Klump, G., . . . Thiel, C. M. (2013). Electrophysiological correlates of auditory change detection and change deafness in complex auditory scenes. *NeuroImage*, 75(0), 155–164. doi:<https://doi.org/10.1016/j.neuroimage.2013.02.037>
- Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8(5), 368–373. doi:<https://doi.org/10.2307/40063214>
- Snyder, J. S., & Gregg, M. K. (2011). Memory for sound, with an ear toward hearing in complex auditory scenes. *Attention, Perception, & Psychophysics*, 73, 1993–2007. doi:<https://doi.org/10.3758/s13414-011-0189-4>
- Snyder, J. S., Gregg, M. K., Weintraub, D. M., & Alain, C. (2012). Attention, awareness, and the perception of auditory scenes. *Frontiers in Psychology*, 3, 15. doi:<https://doi.org/10.3389/fpsyg.2012.00015>
- Snyder, J. S., & Weintraub, D. M. (2013). Loss and persistence of implicit memory for sound: Evidence from auditory stream segregation context effects. *Attention, Perception, & Psychophysics*, 75, 1059–1074. doi:<https://doi.org/10.3758/s13414-013-0460-y>
- van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review*, 121, 124–149. doi:<https://doi.org/10.1037/a0035234>
- Vitevitch, M. S. (2003). Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 333–342.
- Weiss, M. W., Vanzella, P., Schellenberg, E. G., & Trehub, S. E. (2015). Pianists exhibit enhanced memory for vocal melodies but not piano melodies. *The Quarterly Journal of Experimental Psychology*, 68(5), 866–877.
- Yost, W. A., & Sheft, S. (1993). Auditory perception. In W. A. Yost, R. R. Fay, & A. N. Popper (Eds.), *Human psychoacoustics* (pp. 193–236). New York, NY: Springer-Verlag.
- Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science*, 20(4), 423–428.