# Perceptual similarity predicts item recognition errors but not serial order errors in auditory working memory

Siimilarity among elements of a scene profoundly affects perceptual organization, as in Gestalt processing (Baker 2012), gist estimation and ensemble coding (Oliva & Torralba 2006; Xu 2002), or auditory scene analysis (Bregman 1990).

Perceptual organization of stimulus elements into objects affects memory representations, as in chunking (Miller 1956) or structured memory representations (Brady & Tenenbaum 2013).

Similarity among elements in a memory set also directly affects memory representations, as in false memory research (Deese 1959, Roediger & McDermott 1995) or exemplar models (Kahana & Sekuler, 2002). Memory structures afford different kinds of retrieval, such as single-item, whole-display, or serial order

(Noyce, Varghese et al. 2024).

But most of this work has only explored **visual stimuli** and **visual memory**!

How do inter-item similarity, perceptual organization, & task demands interact in nonspeech auditory working memory?



Thank you to Sahil Luthra for advice on measuring similarity, and to Cherie Hua for early work on a precursor to this project. Funded in part by the Office of Naval Research (MURI N00014-19-12332) and by an NSERC postgraduate fellowship to ESB



Abigail L. Noyce, Will (Yuhang) Li, Eli S. Bulger, Barbara G. Shinn-Cunningham

### Carnegie Mellon University; University of Wisconsin Madison

Two online experiments (via Gorilla + Prolific) probed the interaction between task (item recognition, serial order) and stimulus set.

**Expt 2: Cat meows** & complex tones, N = 41. **Expt. 1: Everyday sounds** & **complex tones**, N = 40. In each experiment, every participant completed both tasks with both stimulus sets. Per stimulus set, each participant completed 40 trials of item recognition and 36 trials of serial order.



### **Item recognition task**

Sternberg-type old-new judgment on 4-item memory sequences. 50% "old" probability.



Baker, ed. (2012). *The Oxford Handbook of the History of Psychology: Global* Noyce, Varghese, et al. (2024). Perceptual organization and task demands Perspectives.

Bregman (1990). Auditory Scene Analysis: The Perceptual Organization of Deese (1959). On the prediction of occurrence of particular verbal

intrusions in immediate recall. JEP.

Kahana & Sekuler (2002). Recognizing spatial patterns: A noisy exemplar approach. Vision Research.

Noyce et al. (2017). Sensory-biased and multiple-demand processing in human lateral frontal cortex. Journal of Neuroscience. Noyce et al. (2022). Extended frontal networks for visual and auditory working memory. Cerebral Cortex.

## **Auditory Working Memory**

### Serial order task

Serial order same-different judgment on 4-item memory sequences. 50% "same" probability.



jointly shape auditory working memory capacity. JASA Express Letters. Oliva & Torralba (2006). Building the gist of a scene: The role of global image features in recognition. Progress In Brain Research. Piczak (2015). Environmental sound classification with convolutional neural networks. IEEE International Workshop on Machine Learning for Signal Processing. IEEE: MLSP.

Roediger & McDermott (1995). Creating false memories: Remembering words not presented in lists. JEP:LMC. Xu (2002). Limitations in object-based feature encoding in visual short-

term memory. *JEP:HPP*.

in item recognition



**Perceptual similarity ratings** between the probe and the most-similar sequence item (Target-absent trials) or next-mostsimilar sequence item (Target-present trials) **significantly predict false alarm rates for everyday sounds** and **cat** meows, and predict both hit and false alarm rates for complex tones.

# in serial order



Mean adjacent-stimuli similarity

**Perceptual similarity ratings** between adjacent stimuli (Same-order trials) or the position-swapped stimuli (Differentorder trials) **do not predict hit rates** for any stimulus set, and **only predict false alarm rates for everyday sounds**.

**Item recognition** depends on "ungrouped" memory representations, with elements stored as single items. Recognition responses are predicted by similarity between the probe item and the memoranda.

**Serial order** depends on structured memory representations, which are facilitated by perceptual grouping. For complex tones, a sequence (melody) is held as a single object.

understand auditory working memory.

# **Probe-item similarity predicts "old" responses**

Target-present trials

Target-absent trials

### Inter-item similarity does not predict "same" responses

## Different-order trials



Swapped-stimuli similarity

## Conclusions

# Sounds have inescapable temporal structure that must be considered to