Short-Term Memory for Spectrally and Temporally Complex Sounds: Comparing Apples to Apples

G77

Kristina M. Visscher[1], Michael J. Kahana[2], Robert Sekuler[1,3]
[1]Volen Center for Complex Systems, Brandeis University, Waltham, MA
[2]Department of Psychology, University of Pennsylvania, [3]Department of Cognitive and Neural Systems, Boston University

INTRODUCTION

The auditory and visual systems share common constraints and processing goals, and may thus perform many of the same operations on their respective input streams.

Initially it may appear obvious that input from different modalities is treated differently by the brain: early sensory processors (eyes and ears) are obviously different.

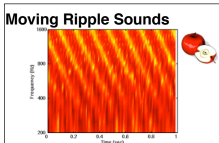
Relatively recent data have shown that early transformations performed by the nervous system on auditory and visual data appear to be quite similar (Shamma, 2001, deCharms, Blake and Merzenich, 1998).

We ask whether the processing performed during memory for hearing and vision are also analogous. In order to compare processing performed during auditory and visual memory, the same memory task should be performed with each, and auditory and visual stimuli must be as analogous in their early processing as possible. We use moving ripple sounds because they share several common properties with the sinusoidal gratings frequently used in psychophysical research.

Ripple Stimuli

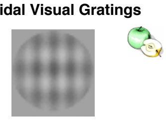
'Moving ripple sounds' are broad band stimuli that vary sinusoidally in time and frequency content (Shamma, 2001). These stimuli and the memory task in which we studied them allowed detailed comparisons of auditory memory and visual memory.

Moving Ripple Sounds



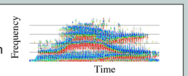
- Vary sinusoidally in frequency content and time
- Cochlea organized by frequency
- Stimuli can serve as basis set for arbitrary sound
- Primary cortex:
  - •Center-surround receptive fields
  - •Responds optimally, selectively see deCharms et al, 1998

Sinusoidal Visual Gratings



- Vary sinusoidally in space
- Retina organized by space
- Stimuli can serve as basis set for arbitrary visual image.
- Primary cortex:
  - •Center-surround receptive fields
  - •Responds optimally, selectively

This spectrogram for the spoken word 'real' illustrates that natural speech shows frequency bands and temporal variation on the scale of the ripple sounds examined.

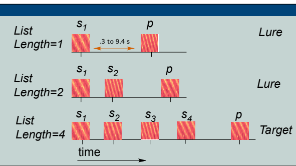


From http://cslu.cse.ogi.edu/tutordemos/SpectrogramReading/ipa/ipadefault.html

References:
deCharms, R. C., Blake, D. T., & Merzenich, M. M. (1998). Optimizing sound features for cortical neurons. Science, 280(5368) 1439-1443.
Kahana, M. J., & Sekuler, R. (2002). Recognizing spatial patterns: a noisy exemplar approach. Vision Research, 42(18), 2177-192.
Nosofsky, R. (1985). Overall similarity and the identification of separable-dimension stimuli: a choice model analysis. Percept & Psychophys, 38(5), 415-32.
Shamma, S. (2001). On the role of space and time in auditory processing. Trends in Cognitive Science, 5(8), 340-348.
Sternberg, S. (1966). High-speed scanning in human memory. Science, 153, 652-654.

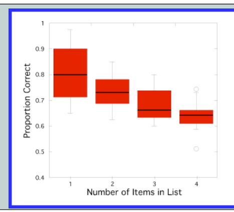
Supported by: NIH grants MH068404 and T32 NS07292.
Presented CNS, 2006
Contact: visscher@brandeis.edu

Examine memory for ripple sounds and visual gratings using same method



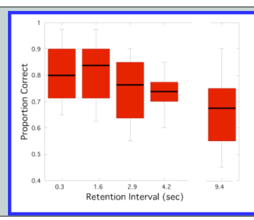
The Sternberg recognition paradigm allows quantitative comparison between memory for auditory and visual stimuli.
Used to examine memory for visual gratings (see Kahana & Sekuler, 2002).
Lists of 1 to 4 items ($s_1 \ldots s_4$) were presented, followed by a probe (p). Subjects responded whether p matched ANY stimulus item.
12 subjects performed 20 trials of each condition, randomized.

Recognition decreases as a function of list length



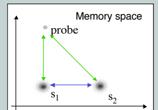
By repeated measures ANOVA, these are different ($p<1*10^{-7}$)
Similar to published results using gratings (Kahana & Sekuler, 2002).

Recognition changes little with retention interval



By repeated measures ANOVA, p=.02.
Similar to published results using gratings (Kahana & Sekuler, 2002).

Noisy Exemplar Model (NEMo)



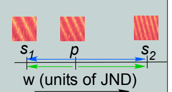
Subjects' response depends on Summed probe-item similarity and inter-item similarity.

Assumes that stimuli are represented in memory with some noise, ε. NEMo has been used to accurately describe visual memory (Kahana and Sekuler, 2002). It is an extension Nosofky et al's GCM model(1985).
Subjects respond that a probe matches an item if the following value exceeds some criterion.

$$\underbrace{\sum_{i=1}^{L} \alpha_i \eta(\mathbf{p}, \mathbf{s_i} + \varepsilon_i)}_{\text{Summed probe-item similarity}} + \underbrace{\frac{2}{L(L-1)} \beta \sum_{i=1}^{L-1} \sum_{j=i+1}^{L} \eta(\mathbf{s}_i + \varepsilon_i, \mathbf{s}_j + \varepsilon_j)}_{\text{Mean interitem similarity}}$$

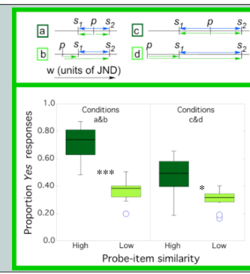
α describes the degree of forgetting of previous stimuli. L is the list length, and β determines the direction and amplitude of the effect of inter-item similarity. η defines the similarity between stimuli, in memory space. $\eta = Ae^{-\tau|s_i - s_j|}$

Testing NEMo



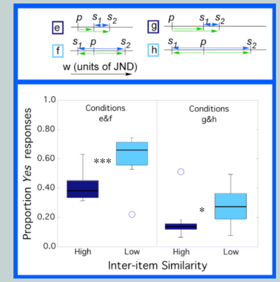
By carefully selecting stimuli, we created conditions in which either summed probe-item similarity or inter-item similarity was varied, while the other parameters (including similarity of the closest probe) were held constant.
Stimulus differences were individually scaled for each subject's just noticeable difference threshold (70% correct).
12 subjects performed 4688 trials over 8 sessions.

When summed probe-item similarity is high, subjects more likely to make false alarms



Implies summed similarity (not just similarity of closest item) influences subjects' decisions, as found for visual gratings (K&S, 2002)
Lure trials only. *=p<.05, ***=p<.001 by paired T-test.

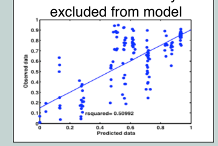
When list items are similar, subjects less likely to make false alarms



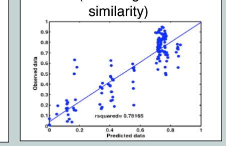
Implies β parameter < 0 as found for visual gratings (K&S, 2002).
Lure trials only. *=p<.05, ***=p<.001 by paired T-test.

Inter-item similarity necessary to predict subjects' response

Inter-Item Similarity excluded from model



NEMo (including inter-item similarity)



Data from this experiment were fit to nested variants of NEMo. Fits were far better when Inter-item similarity was included, by r^2 and Akaike Information criterion measures (controls for generalizability of the results).

CONCLUSIONS

Ripple sounds support detailed comparisons between auditory and visual memory
Short term memory for ripple sounds is similar in key respects to memory for visual gratings. For both auditory ripples and visual gratings:
•Recognition decreases as a function of the number of items in memory
•Retention interval has only a modest effect on recognition memory
•Summed similarity between probe and *all* stimuli influences subjects' response
•Inter-item similarity influences subjects' response.