Temporal cues and modulation rate interplay with attention to detect a target sound embedded in background noise

S. Akram[1,4], J.Z. Simon[1,3], M.Elhilali [2], B.Englitz [4], S.A. Shamma[1,4]

[1] Department of Electrical and Computer Engineering, University of Maryland
[2] Department of Electrical and Computer Engineering, Johns Hopkins University
[3] Departments of Biology, University of Maryland
[4] The Institute for Systems Research, University of Maryland

Introduction

To parse a complex auditory scene, perceptual cues are extracted in an interplay of bottom-up saliency and top-down attentional modulation. Low modulation rates (2–12 Hz) are essential to integrate these different cues (Moore and Gockel, 2002). Close rates - like 4 and 7 Hz - were expected to behave quite similar, but showed significantly different characteristics in previous studies (Xiang et al. 2010, Wang et al. 2011).

We investigate these differences more closely using psychophysical and magnetoencephalographic (MEG) data in humans. The subject's attention was drawn to different features of an auditory scene, composed of a rhythmic (7 Hz) target buried in a random, irregular background, which complements an earlier study using a 4 Hz rhythm (Elhilali et al. 2009).

These differences are possibly related to the low-pass characteristic of neurons, and the functional role that the different frequencies play in global brain interactions (4 Hz is low theta, 7 Hz high theta).

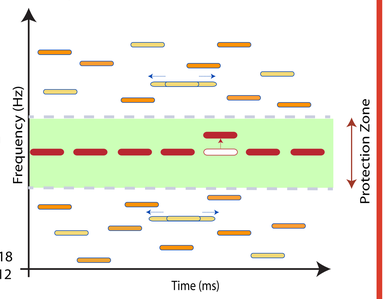
Stimulus Design

The stimulus is a rhythmic 7 Hz, regular target embedded in a random, irregular background. The target frequency is chosen randomly in the range of 250-500 Hz with a 2 semitone interval. The target is within protection zones of 4, 8 or 12 semitones.

Tasks

Target Task: Detection of a frequency-shifted deviant, randomly placed in the target sequence.

Masker task: Detection of elongated masker tones, in a single 500 ms time window, randomly chosen for each trial.



Procedure

Psychoacoustics part A: n=18
Psychoacoustics part B: n=12
MEG: n=12

Psychoacoustics were performed using Matlab in a soundproof room.
MEG recordings were conducted in a dimly lit magnetically shielded room (Yokogawa Electric Corporation) using a 160-channel whole-head system (Kanazawa Inst. of Tech.).

References

D. P. Pillips et al, The Journal of the Acoustical Society of America, 1989.
D. Poeppel, Speech Communication 2003. 1.
J. Xiang et al, The Journal of Neuroscience, 2010.
J. J. Eggermont, Hear Res 56:153–167, 1991.
M. Elhilali et al, PLoS Biology 2009.
Y. Wang et al, Journal of Neurophysiology 2011.
M.P. Kilgard and M. M. Merzenich, Hear Res 134 :16–28, 1999.

Acknowledgements

The authors thank Nai Ding and Majid Mirbagheri for interesting discussions and technical support.
This project was supported by NIH grant, 1R01AG036424.

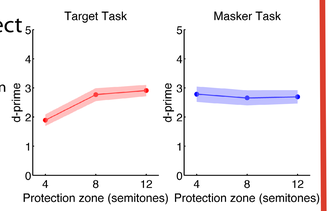
Psychoacoustics

Part A

Protection Zone effect

Wider Protection Zone :

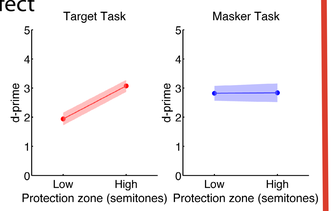
➔ Easier Target Task detection
➔ No significant effect on Masker Task



Target Frequency effect

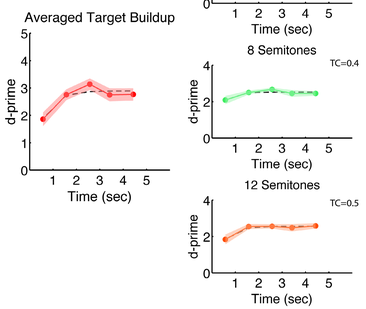
Higher Target Frequency (High :TF > 354 Hz) :

➔ Easier Target Task detection
➔ No effect on Masker Task



Build-Up effect

The most useful Protection Zone is 4 semitone interval. For 8 and 12 semitones, the easiness of the task may lead to early saturation. (Time constants estimated via exponential fits).



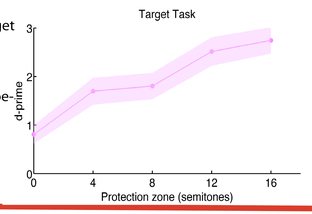
Part B

Protection Zone Investigation

To test the effect of different Protection Zone widths, the range was expanded from 0 to 16 semitones

with a constant rate of 7 Hz . (target task only)

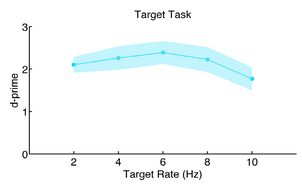
Wider Protection Zone ➔ better behavioral performance



Target Rate Investigation

To investigate the effect of slow and fast rates on the subject's performance, we fixed the Protection Zone to 8 semitones and varied the rates from 2 to 10 Hz in steps of 2 Hz. (target task only)

➔ No significant performance change as a function of target rate.



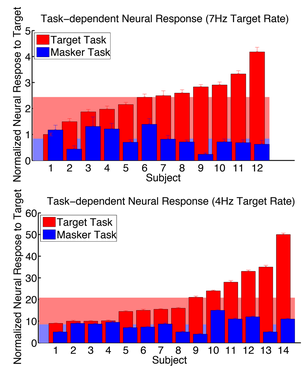
Neural Analysis

Task-Dependent Neural Response

Neural Responses to the target normalized by average of neighboring frequencies.

Note scale change for different vertical axes.

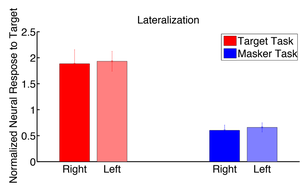
Error bar = 1 standard error
Background bars = average response



Lateralization

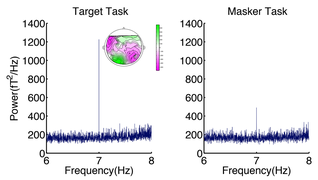
No significant lateralization found (in contrast to strong lateralization to the right for 4hz exp., Elhilali et al. 2009)

probably caused by susceptibility of right/left activity to modulation rate of the stimuli (Wang et al. 2011)



Neural Response

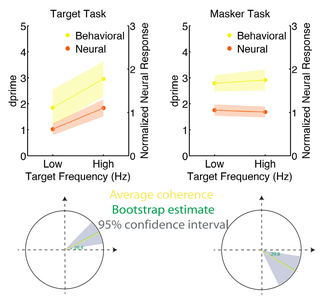
Power spectral density averaged over subjects (Inset: corresponding magnetic field distribution for a representative subject).



Bottom-Up Saliency of Target

Neural response and behavioral performance to the target rhythm as a function of target frequency, followed by their correlation estimate.

➔ Positive and negative correlations for Target and Masker task, respectively.



Power Enhancement

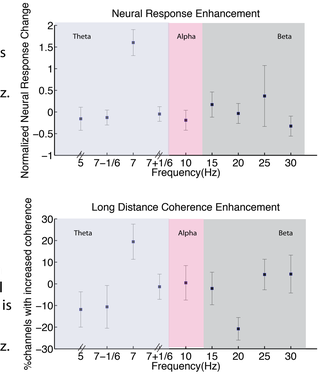
Difference between neural responses of Target relative to Masker task.

Significant enhancement only at 7 Hz.

Phase Enhancement

Difference between the number of long-range channel pairs with increased and decreased coherence in target task, normalized over the total number of long-range channel pairs is shown in the figure.

Significant enhancement only at 7 Hz.

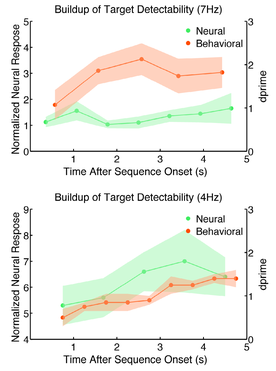


Neural Build-Up Investigation

Normalized neural responses and the behavioral performance to the target rhythm for Target task.

➔ Constant neural response and no correlation with behavioral buildup (despite significant buildup for 4 Hz case, correlated with behavioral response, Elhilali et al. 2009).

➔ Possibly caused by faster buildup of higher rates leading to a flat buildup after an early sharp increase, in response.



Discussion

- Wider Protection Zone ➔ Easier detection for the Target Task
  ➔ No significant effect on Masker Task
  ➔ Increased neural response for Target Task
- Higher target frequency ➔ Easier detection for the Target Task
  ➔ No significant effect on Masker Task
  ➔ Increased neural response for Target Task
  (Pos./neg. correlations of behavioral and neural responses for Target/Masker Task)
- Behavioral buildup for Target Task mostly because of 4 semitone Protection Zone (No significant buildup for 8 and 12 semitones, probably due to easiness of the task)
- No behavioral performance change with respect to different target rates
- Decrease in neural response compared to 4Hz target rate
- Increased neural response in target task over masker task, i.e. attended versus unattended task (Power and phase enhancement)
- No lateralization
- No temporal neural buildup