Influence of source location and temporal structure on spatial auditory saliency

Podwinska, Z, Fazenda, BM and Davies, WJ

http://dx.doi.org/10.1121/1.4987852

Title	Influence of source location and temporal structure on spatial auditory saliency
Authors	Podwinska, Z, Fazenda, BM and Davies, WJ
Publication title	The Journal of the Acoustical Society of America (JASA)
Publisher	Acoustical Society of America
Type	Article
USIR URL	This version is available at: http://usir.salford.ac.uk/id/eprint/43808/
Published Date	2017

USIR is a digital collection of the research output of the University of Salford. Where copyright permits, full text material held in the repository is made freely available online and can be read, downloaded and copied for non-commercial private study or research purposes. Please check the manuscript for any further copyright restrictions.

For more information, including our policy and submission procedure, please contact the Repository Team at: library-research@salford.ac.uk.

PM2pPPc27:

INFLUENCE OF SOURCE LOCATION AND TEMPORAL STRUCTURE ON SPATIAL AUDITORY SALIENCY

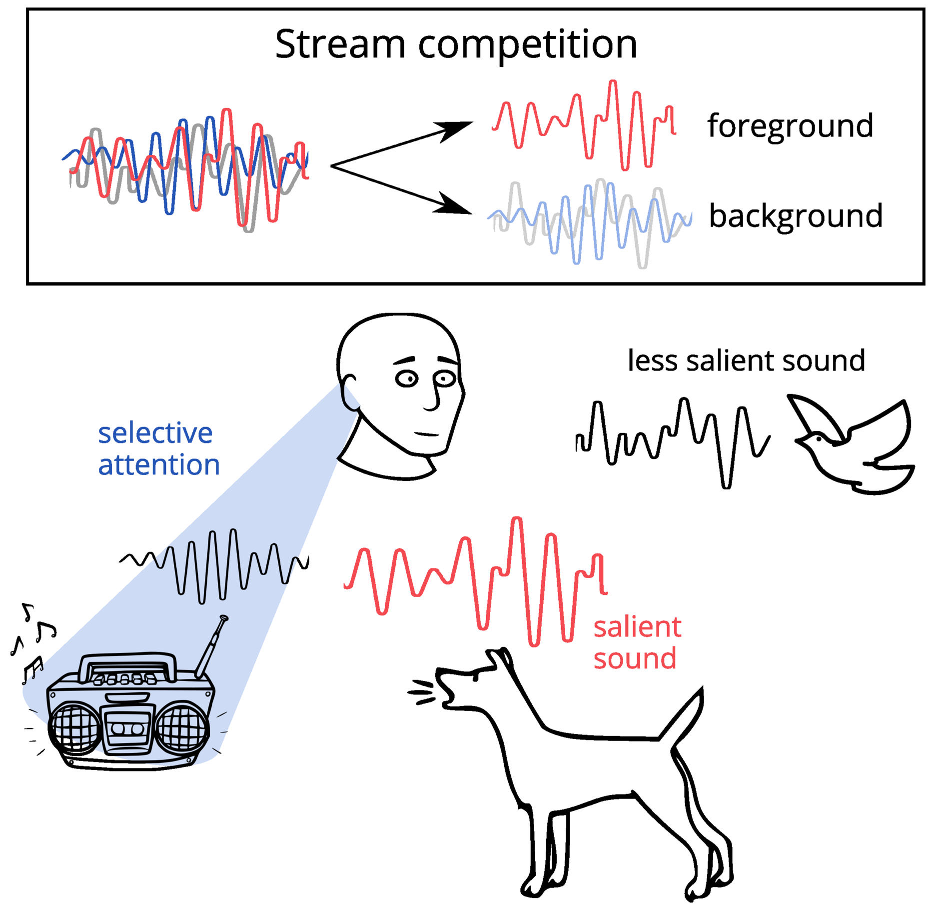
University of Salford MANCHESTER

Zuzanna M. Podwinska, Bruno M. Fazenda, William J. Davies
Acoustics Research Centre

BACKGROUND

A crucial part of auditory scene analysis is attention:

- top-down, attentional 'searchlight' controlled by the listener, and
- bottom-up, which relies on physical features of sound. The property of sound which makes it stand out in a scene is **saliency**.



Both affect stream competition, which determines which sound will become *foreground*, and which will be *background*.

OBJECTIVE

The objective of this study is to test the relationship between absolute stimulus location and saliency.

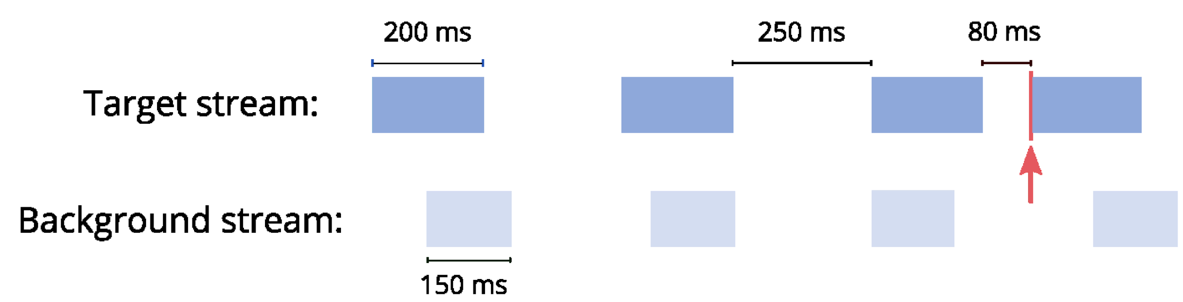
- **Hypothesis**: perceived saliency of sound changes with absolute location of stimulus around the listener.
- **Motivation**: aid development of a spatial auditory attention model.

If the hypothesis is true, this effect could be directly implemented in a binaural saliency model.

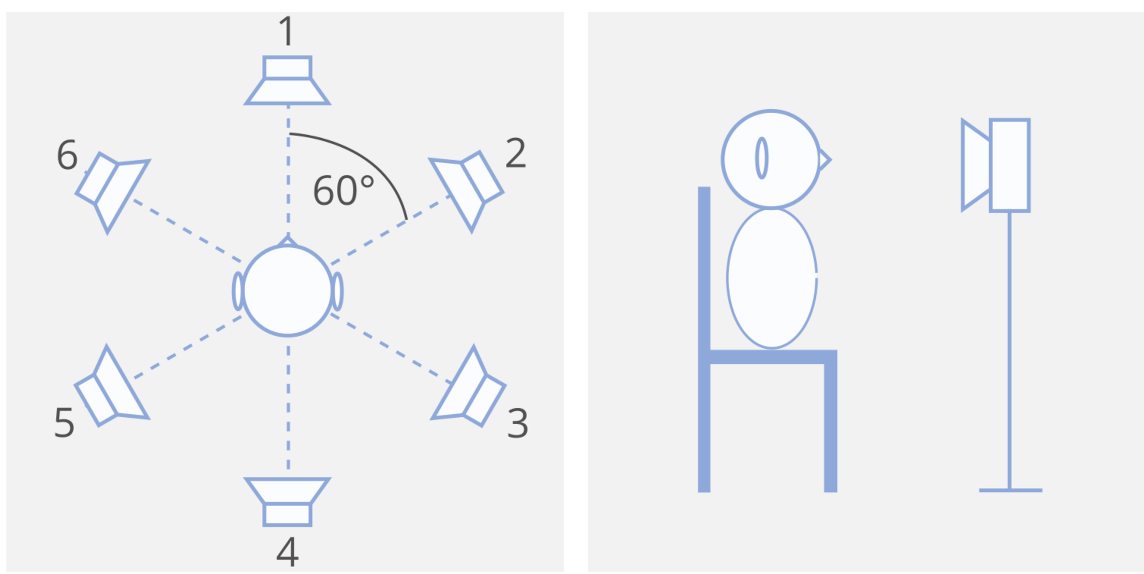
METHODS

Experiment was based on the Segregation of Asynchronous Patterns (SOAP) framework [1].

- **Assumption**: in absence of top-down effects, a more salient stream will be more likely to become foreground. Also, it will be easier to detect changes in the foreground (more salient) stream.
- **Scenario**: two simultaneous sound patterns, one of which includes a change (shortened inter-stimulus interval). Task: detect the change and determine in which pattern it originated.



- **Spatial extension of SOAP**: sound patterns from 2 different location around the listener.
- **Stimuli**: patterns of noise bursts
  - 2 frequencies: high– and low-pass filtered noise, f_c = 2kHz
  - 2 pattern tempos: short and long stimuli
- **Setup**: listening room, 6 loudspeakers



- **Conditions**: 6 target locations (x 5 background locations) x 2 frequency x 2 tempo
- **Participants**: N = 19
- **Data collected**: response time, accuracy, body position data from Kinect

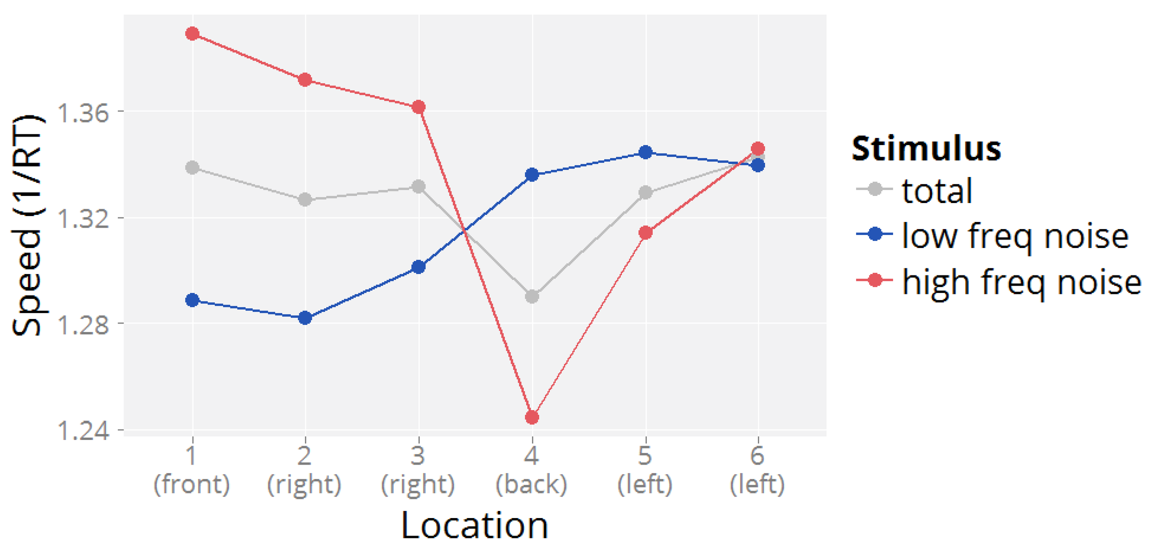
RESULTS

Response Times

Pre-processing: normalizing 1/x transformation was applied (resulting in **response speed**).

Statistical analysis: repeated-measures ANOVA to test main effects of location, frequency and tempo and their interactions on response speed:

- effect of **tempo**, F(1,18)=30.9, p<0.0001, η^2_G=0.01
- interaction effect between **location and frequency**, F(5,90)=5.1, p=0.0004, η^2_G=0.02



For high frequency sounds - effect of location on response speed, F(5,90)=3.4, p=0.008, η^2_G=0.03.

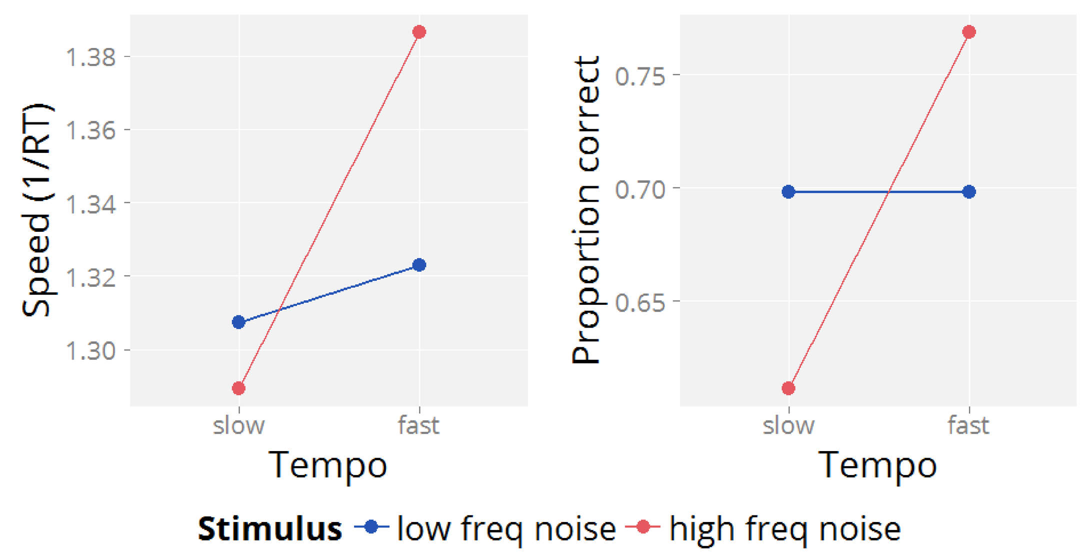
- interaction effect between **frequency and tempo**, F(1,18)=7.2, p=0.02, η^2_G=0.01

Accuracy

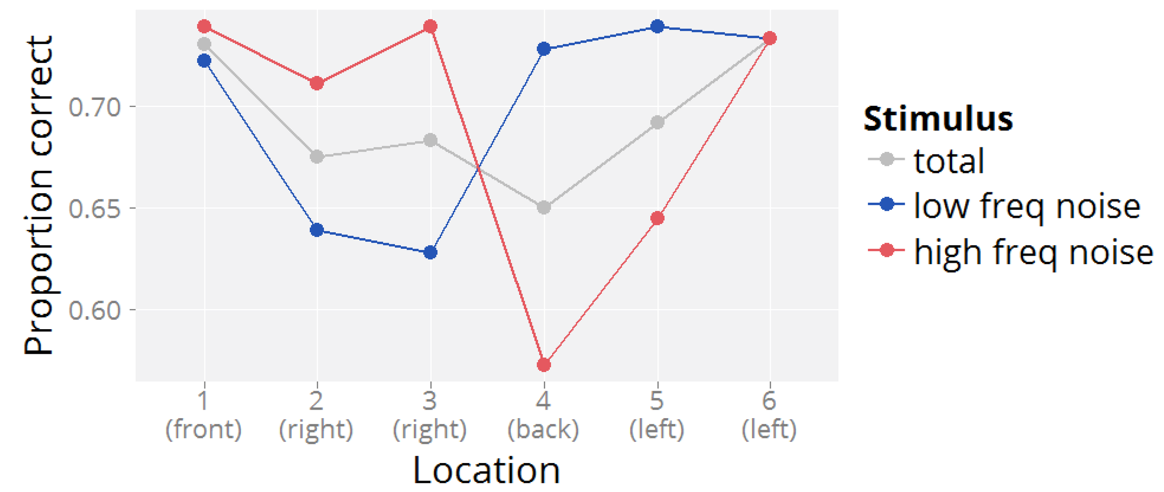
Pre-processing: personalized acceptance window, removed on average 27% of each participant's correct responses (N=18).

Statistical analysis: Generalized Estimating Equations, significant predictors:

- interaction between **frequency and tempo** (β=0.76, SE=0.18, p<0.0001)



- interaction between **location 4 (back) and frequency** (β=-0.82, SE=0.41, p=0.045).



DISCUSSION

Summary:

- Participants generally responded quicker to the **fast pattern** than to the slow pattern (median difference, md = 40ms), but not more accurately.
- The effect of **tempo** was smaller for low frequency (md = 20ms) than for high frequency sounds (md = 57ms). The responses were also more likely to be correct for high frequency fast, than low frequency slow patterns.
- For high frequency stimuli, responses were slower (md = 84ms) if target was **behind the listener**, than if it was in front of them, and they were also more likely to be incorrect.
- The results suggest that spatial saliency is related to the spectral content of the sound. For high (but not low) frequency noise, sounds in the back are less salient than those in the front.

Future work:

- Analyse head movement data from Kinect
- Similar experiment but with natural sounds and/or including visual attention effects

REFERENCE

[1] F. Tordini, A. S. Bregman, J. R. Cooperstock, A. Ankolekar, and T. Sandholm, "Toward an improved model of auditory saliency.", in *Proc. of the 19th International Conference on Auditory Display*, Lodz, Poland, 2013

z.podwinska@edu.salford.ac.uk