2014

Probing auditory scene analysis.

Deike, S

http://hdl.handle.net/10026.1/9353

10.3389/fnins.2014.00293
Front Neurosci

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.
Probing auditory scene analysis

Susann Deike1*, Susan L. Denham2,3 and Elyse Sussman4,5

1 Special Lab Non-Invasive Brain Imaging, Leibniz Institute for Neurobiology, Magdeburg, Germany
2 Cognition Institute, University of Plymouth, Plymouth, UK
3 School of Psychology, University of Plymouth, Plymouth, UK
4 Department of Neurosciences, Albert Einstein College of Medicine of Yeshiva University, Bronx, NY, USA
5 Department of Otolaryngology-Head and Neck Surgery, Albert Einstein College of Medicine of Yeshiva University, Bronx, NY, USA

*Correspondence: sdeike@ln-magdeburg.de

Edited and reviewed by:
Isabelle Peretz, Université de Montréal, Canada

Keywords: auditory scene analysis, multistable perception, ambiguity, realistic auditory scenes, stream segregation

In natural environments, the auditory system is typically confronted with a mixture of sounds originating from different sound sources. The sounds emanating from different sources can overlap each other in time and feature space. Thus, the auditory system has to continuously decompose competing sounds into distinct meaningful auditory objects or “auditory streams” associated with the possible sound sources. This decomposition of the sounds, termed “Auditory scene analysis” (ASA) by Bregman (1990), involves two kinds of grouping. Grouping based on simultaneous cues (e.g., harmony) and on sequential cues (e.g., similarity of acoustic features over time). Understanding how the brain solves these tasks is a fundamental challenge facing auditory scientists. In recent years, the topic of ASA was broadly investigated in different fields of auditory research using a wide range of methods, including studies in different species (Hulse et al., 1997; Fay, 2000; Fishman et al., 2001; Moss and Surlykke, 2001), and computer modeling of ASA (for recent reviews see, Winkler et al., 2012; Gutchalk and Dykstra, 2014). Despite advances in understanding ASA, it still proves to be a major challenge for auditory research, especially in verifying whether experimental findings are transferable to more realistic auditory scenes. This special issue is a collection of 10 research papers and one review paper providing a snapshot of current ASA research. The research paper on visual perception provides a comparative view of modality specific as well as general characteristics of perception.

One approach for understanding ASA in real auditory scenes is the use of stimulus parameters that produce an ambiguous percept (cf. Pressnitzer et al., 2011). The advantage of such an approach is that different perceptual organizations can be studied without varying physical stimulus parameters. Using a visual ambiguous stimulus and combining real-time functional magnetic resonance imaging and machine learning techniques, Reichert et al. (2014) showed that it is possible to determine the momentary state of a subject’s conscious percept from time resolved BOLD-activity. The high classification accuracy of this data-driven classification approach may be particularly useful for auditory research investigating perception in continuous, ecologically-relevant sound scenes.

A second advantage in using ambiguous stimuli in experiments on ASA is that perception of them can be influenced by intention or task (Moore and Gockel, 2002). By manipulating task requirements one can mirror real hearing situations where listeners often need to identify and localize sound sources. The studies by Shestopalova et al. (2014) and Kondo et al. (2014) examined the influence of motion on stream segregation. In general, and corresponding to earlier findings, both of these studies found that sound source separation in space promoted segregation. Surprisingly, however, the effect of spatial separation on stream segregation was found to be temporally limited and affected by volitional head motion (Kondo et al., 2014), but unaffected by movement of sound sources or by the presentation of movement-congruent visual cues (Shestopalova et al., 2014). Another study, by Sussman-Fort and Sussman (2014), investigated the influence of stimulus context on the buildup of stream segregation. They found that the build-up of stream segregation was context-dependent, occurring faster under constant than varying stimulus conditions. Based on these findings the authors suggested that the auditory system maintains a representation of the environment that is only updated when new information indicates that reanalyzing the scene is necessary. Two further studies examined the influence of attention on stream segregation. Nie et al. (2014) found that in conditions of weak spectral contrast, attention facilitated stream segregation. Shuai and Elhilali (2014) found that different forms of attention, both stimulus-driven and top-down attentional processes, modulated the response to a salient event detected within a sound stream.

The special issue also includes two research papers that extend current views on multistability and perceptual ambiguity. The psychophysical study by Denham et al. (2014) showed that streaming sequences could be perceived in many more ways than in the traditionally assumed (Integrated vs. Segregated organizations) and that the different interpretations continuously compete for dominance. Moreover, despite being highly stochastic, the switching patterns of individual participants could be distinguished from those of others. Hence, perceptual multistability can be used to characterize both general mechanisms and individual differences in human perception. By comparing stimulus conditions that promote one perceptual organization with those causing an ambiguous percept Dollezal et al. (2014) found specific BOLD responses for the
ambiguity, these factors may at least partially explain the well
required a 15 and 20 dB larger S/N ratio than younger listen-
reset to noise (S/N) ratio at which normal hearing young and
both groups. This similarity may suggest that stream segre-
gation was similar in quality between
whether the time course of auditory streaming differs between
normal-hearing listeners and CI users and found that the percep-
tion of streaming sequences was similar in quality between
both groups. This similarity may suggest that stream segrega-
tion is not solely determined by frequency discrimination, and
that CI users do not simply respond to differences between A
and B sounds but actually experience the phenomenon of stream
segregation.

The review by Bendixen (2014) suggests predictability as a
cue for source sound decomposition. Bendixen collected empir-
ical evidence spanning issues of predictive auditory processing,
predictive processing in ASA, and methodological aspects of mea-
suring ASA. As a result, and as a theoretical framework, an anal-
ogy with the old-plus-new heuristic for grouping simultaneous
acoustic signals was proposed.

Taken together, this special issue provides a comprehensive
summary of current research in ASA, relating the approaches and
experimental findings to natural listening conditions. It
would be highly desirable in future research on ASA to use more
natural stimuli and to test the ecological validity of these find-
ings. With this special issue we hope to raise awareness of this
issue.

ACKNOWLEDGMENT
We thank the authors for their contributions and the reviewers
for their useful comments. Susann Deike was funded by the National
Institutes of Health (DC004263).

REFERENCES


Rieger, J. W. (2014). Online tracking of the contents of conscious percep-


Steshopulova, L., Bohm, T. M., Bendixen, A., Andreou, A. G., Georgiou, J., Garreau, G., et al. (2014). Do audio-visual motion cues promote segre-


**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 18 August 2014; accepted: 27 August 2014; published online: 12 September 2014.


This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience.

Copyright © 2014 Deike, Denham and Sussman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.