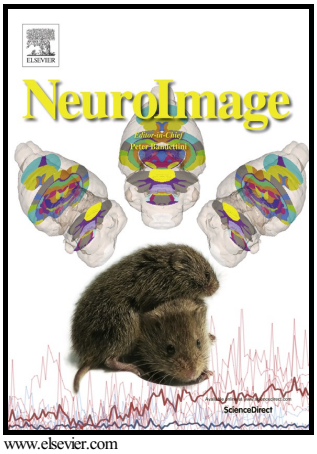
Noise-robust cortical tracking of attended speech in real-world acoustic scenes

Søren Asp Fuglsang, Torsten Dau, Jens Hjortkjær



PII: S1053-8119(17)30318-X
DOI: http://dx.doi.org/10.1016/j.neuroimage.2017.04.026
Reference: YNIMG13967

To appear in: *NeuroImage*

Received date: 10 December 2016
Accepted date: 10 April 2017

Cite this article as: Søren Asp Fuglsang, Torsten Dau and Jens Hjortkjær, Noise-robust cortical tracking of attended speech in real-world acoustic scenes *NeuroImage,* http://dx.doi.org/10.1016/j.neuroimage.2017.04.026

Noise-robust cortical tracking of attended speech in real-world acoustic scenes

Søren Asp Fuglsang[1,*], Torsten Dau[1], Jens Hjortkjær[1,2,*]

[1]Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark, Ørsteds Plads, Building 352, 2800 Kgs. Lyngby, Denmark

[2]Danish Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, Kettegaard Allé 30, 2650 Hvidovre, Denmark

soerenf@elektro.dtu.dk
jhjort@elektro.dtu.dk
[*]**Corresponding author at**: Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark, Ørsteds Plads, Building 352, 2800 Kgs. Lyngby, Denmark. Tel (+45) 45253950; Fax (+45) 45880577.

Abstract

Selectively attending to one speaker in a multi-speaker scenario is thought to synchronize low-frequency cortical activity to the attended speech signal. In recent studies, reconstruction of speech from single-trial electroencephalogram (EEG) data has been used to decode which talker a listener is attending to in a two-talker situation. It is currently unclear how this generalizes to more complex sound environments. Behaviorally, speech perception is robust to the acoustic distortions that listeners typically encounter in everyday life, but it is unknown whether this is mirrored by a noise-robust neural tracking of attended speech. Here we used advanced acoustic simulations to recreate real-world acoustic scenes in the laboratory. In virtual acoustic realities with varying amounts of reverberation and number of interfering talkers, listeners selectively attended to the speech stream of a particular talker. Across the different listening environments, we found that the attended talker could be accurately decoded from single-trial EEG data irrespective of the different distortions in the