Feasibility of acoustic testing with fMRI for speaker recognition experiments

Almut Braun[1], *Jens Sommer*[2], *and Andreas Jansen*[2]
[1]*Department of Phonetics, University of Marburg, Germany*
almut.braun@staff.uni-marburg.de
[2]*Section of BrainImaging, Department of Psychiatry and Psychotherapy, University of Marburg, Germany*.
{jens.sommer|andreas.jansen}@med.uni-marburg.de

Introduction:

The present study aims to investigate human speaker recognition ability while listeners are undergoing a functional MRI scan. Central questions are: To what extent - if at all - is it possible to do a more complex speaker recognition experiment within an MR scanner? If so, could different patterns of BOLD (blood-oxygen-level-dependent) activation and deactivation be linked to a listeners‘ performance in a speaker recognition task? Do familiar voices evoke BOLD activations in other areas of the brain than voices which have just been heard one time before?

In previous studies, voice coding has been associated with the superior temporal sulci (STS) and the inferior frontal cortex (Andics et al. 2013), and voice recognition was associated with the middle and posterior STS, the right ventrolateral prefrontal regions and the insular cortex, the anterior temporal pole (Andics et al. 2010).

This is a feasibility study. It will be tested whether it is generally possible to do a speaker recognition experiment within the noisy environment of a 3-tesla MR scanner. Different settings of the scanner as well as different types of headphones (electroacoustic/pneumatic) have been tested to reduce the subjective noise level. It was reported that the best noise reduction could be obtained when the listener was wearing electroacoustic headphones (mr-confon). Additionally, the acoustic condition was improved by wrapping the participant’s head with special foam material inside the head coil. Further improvements could be achieved by separating the noise and voice frequencies by adjusting the scanning parameters (e.g. echo time, repetition time, field of view, matrix)
If the feasibility study reveals no weaknesses in the local setup, a follow-up study with blind and sighted listeners will be carried out. Gougoux et al. 2009 found different activation patterns for blind and sighted listeners in a voice discrimination task (same/different speaker).

Method:

The experiment consists of two parts. In the first part, a sound file with 15 spontaneous voice samples of different male speakers is played to the listeners while lying inside the MR scanner. Ten of these voice samples come from famous speakers which are supposed to be recognized easily, five samples are voices which the listeners have never heard before. Each voice sample (duration: 30 seconds) is followed by a silent interval of 10 seconds to ensure that a baseline can be established. The famous speakers are selected according to a prior experiment with other participants. Famous striking voices are included in the experiment to provoke extreme reactions to estimate a general effect size for the given task.
In a second part, the participants undergo a second fMRI scan. This time, they have to listen to another sound file. This sound file consists of voice samples of 3 non-famous speakers they had heard in the first part of the experiment and 6 new (unknown) voices. Participants are

asked to indicate which of the voices they have heard before.

References

Andics, A., McQueen, J.M., Petersson, K.M.,Gal, V., Rudas, G., and Vidnyanszky, Z. (2010). Neural mechanisms for voice recognition.*Neuroimage* **52**, 1528–1540.

Andics, A. , McQueen,J.M., Petersson, K.M.. (2013) Mean-based neural coding of voices. *Neuroimage* **79**, 351-360

Gougoux, F.; Belin, P.; Voss, P.; Lepore, F.; Lassonde, M.; Zatorre, R.J. (2009): Voice perception in blind persons: a functional magnetic resonance imaging study. *Neuropsychologia* **47** (13), 2967-2974.