Stefan Schadwinkel
Dr. sc. hum.

Investigation of Auditory Stream Segregation based on Spatial Cues using Magnetoencephalography and Functional Magnetic Resonance Imaging

Promotionsfach: Neurologie
Doktorvater: Prof. Dr. med. Dr. h.c. Dipl. Psych. Werner Hacke

In complex auditory environments like a crowded city or a cocktail party, multiple sources emit sound in parallel, but hit our ears as a single convolute. It is then on our brain to reconstruct the original sound sources from this mixture. One important process that subserves this auditory scene analysis is stream segregation (or streaming), which separates and assigns interleaved sound sequences into distinct perceptual entities.

The aim of this thesis was to investigate the neural correlates of auditory stream segregation, especially with the focus on a spatial streaming cue, namely interaural time differences (ITD). Interaural time differences refer to the difference in the arrival time at both ears of the sound waves emitted by a single sound source. Its effect is a perceptual lateralization of the sound in the horizontal plane, where a difference of 0 µs corresponds to a sound straight ahead and the maximum difference of about 690 µs corresponds to maximal lateralization, i.e. a sound directly to left or the right.

In a first experiment, known streaming-associated neural correlates, observed for streaming elicited by pitch-based cues, were compared to streaming based on ITD. Streaming can be elicited when subsequent tones differ in ITD, i.e. when the tones are perceived to originate from different spatial positions. The difference between the ITDs of these different subsequent tones is referred to as ΔITD. For this first experiment, twelve listeners were investigated using magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI) to evaluate if the previously reported correlates of streaming in auditory cortex represent a general mechanism that is independent of the stimulus cue. The MEG results showed enhancement of the P_1m after 60-90 ms that was similar during streaming based on ITD and pitch. Sustained fMRI activity was enhanced at identical sites in Heschl's gyrus and the Planum temporale for both cues; no topographical specificity for space or pitch was found for the streaming-associated enhancement.

A second experiment investigated how interaural time differences modulate the streaming-associated fMRI activation with an emphasis on the interactions between physical cue strength (ΔITD) and the prevalence and salience of actual streaming perception. Eight listeners were investigated by fMRI while being presented with five different ΔITD conditions out of an active baseline with zero ΔITD. Results showed a dissociation between transient and sustained components of the fMRI activation: transient activity after stimulus onset increased in amplitude with ΔITD in a non-linear fashion; sustained activity increased during streaming, with no further modulation by perceptual salience, prevalence, or ΔITD.

The modulation of the transient activity in the second experiment could be explained in part by the jitter of the perceptual reversals in relation to the stimulus. These perceptual reversals relate to auditory stream segregation as a multistable perceptual phenomenon. Streaming perception can spontaneously alternate between one integrated and two segregated streams,

similarly to visual illusions like the Necker Cube, when the difference between the two streams remains at an ambiguous level. To study these reversals in more detail, a third experiment was conducted: twelve listeners were investigated by fMRI using a semi-ambiguous ITD-based streaming paradigm, which aimed at dissociating the perceptual reversals from the physical stimulus properties, while at the same time withholding the next reversal long enough to allow imaging using fMRI. Results showed distinct fMRI activity synchronized to the perceptual reversals and not to the physical stimulus properties. This activity was observed in and around the auditory cortex as well as in a subcortical auditory processing stage, the inferior colliculus. Activity in auditory cortex additionally showed enhanced sustained fMRI activation during those time intervals, where segregated streams were possible to be perceived.

In a fourth step, the data which were acquired during the three experiments were investigated using a machine-learning paradigm (support vector machines) in order to predict the actual percept based on the acquired data. This analysis produced meaningful results which were generally consistent with the results gained by the more standardized analysis methods used in the individual experiments, but found no additional insights.

Transient and sustained signal components of BOLD activity in the human auditory cortex represent different processes in relation to streaming, they are found throughout auditory cortex and are not modulated by the actual streaming cue. The onset of the sustained component appears not to be locked to the onset of streaming perception and is not able to fully dissociate epochs with segregated streaming perception from epochs with integrated streaming perception. The sustained activity is connected to stimulus-synchronized neural activity, which can be measured by MEG and is not modulated by the prevalence of perceived streaming nor by the salience of the streaming percept. The process behind this activity appears to present a functional relationship to feature-selective adaptation, which is relevant to streaming but does not completely determine the percept. On the other hand, transient activity reflecting the perceptual streaming reversal was observed in cortical and subcortical auditory processing stages. The actual streaming perception is thus likely to emerge from an interaction of peripherally-based bottom-up processing and cortically-based top-down modulation.