fMRI encoding and decoding of natural sounds in the aged human auditory cortex

Julia Erb[1] (julia.erb@uni-luebeck.de)
Lea-Maria Schmitt[1] (l.schmitt@uni-luebeck.de)
Jonas Obleser[1] (jonas.obleser@uni-luebeck.de)

[1]Institute of Psychology, University of Lübeck, Maria-Goeppert-Str. 9A
23562 Lübeck, Germany

Current models of auditory cortical processing describe the neural response to natural sounds as a function of complex acoustic features. Here, we ask whether and how cortical responses to those acoustic features change over the life span. We acquired 3T-fMRI data while young and older adults listened to a story presented against a competing stream of resynthesized natural sounds. We modelled the fMRI responses in auditory cortex as a function of the spectro-temporal modulations contained in the sound mixture, and derived single- and multi-voxel modulation transfer functions (MTFs; fMRI encoding and decoding). We found that topographical best feature maps were preserved in the aged auditory cortex: Tonotopic maps showed the typical mirror-symmetric frequency gradients along Heschl's gyrus in both age groups. Decoding yielded highest accuracies at low frequencies of 230–580 Hz, coarse scales of 0.25 cyc/oct, and slow rates of 4–8 Hz, irrespective of age. However, we observed age differences in tuning to temporal rate, with signatures of broadened rate tuning in older participants. These results indicate that the specificity of temporal representation declines in the aged auditory cortex.

Keywords: fMRI; spectro-temporal modulations; life span

Introduction

Recent advances in computational modelling of fMRI data have furthered our understanding of how natural sounds become represented in the auditory cortex. A series of studies in human and non-human primates have shown that cortical processing of sounds is optimized for the spectro-temporal modulations which are typically present in conspecific vocalizations such as speech or animal calls (Santoro et al., 2014, Hullett et al., 2016, Santoro et al., 2017, Erb et al., 2018). How those representations change in the aged auditory cortex is not well understood. Human psychoacoustic studies suggest that the ability to process temporal cues declines with age (e.g., Gordon-Salant et al., 2006). Recent findings in macaque monkeys indicate that the temporal fidelity of cortical responses decreases and neurons become more broadly-tuned to temporal

modulations in the aged auditory cortex (Chi-Wing and Recanzone, 2017). The present study combines fMRI and computational modeling to examine auditory cortical responses to natural sounds in young and older humans. We focus on the age-group comparison at the level of representation of fundamental acoustic features (spectro-temporal modulations).

Materials and Methods

Participants and data acquisition. Participants were young (n = 33; aged 18–32) and older adults (n = 26; aged 51–74). We acquired continuous fMRI at 3T in eight runs of eight minutes each (2.5 mm isotropic voxels, TR = 947 ms, TE = 28 ms, number of slices = 52, whole-brain coverage).

Stimuli and task. Participants listened to continuous speech (a story) presented against a competing stream of resynthesized natural sounds ("sound textures"; McDermott and Simoncelli, 2001) at 0 dB SNR. Textures were synthesized from the spectro-temporal modulation content of a large set of real-life sounds (n = 192), including speech and vocal samples, music pieces, animal cries, scenes from nature and tool sounds. Participants were asked to listen to the story and answer three multiple-choice questions on the semantic content after each run.

Computational modeling. fMRI responses to sound mixtures were modelled using both model-based univariate encoding (Santoro et al., 2014) and multivariate decoding (Santoro et al., 2017). In brief, sound mixtures were decomposed into acoustic features, that is, the frequency-dependent spectro-temporal modulation content (Chi et al., 2005). Analyses were restricted to voxels in an anatomic mask of the auditory cortex. In a fourfold cross-validation procedure, we used ridge regression to derive single- and multi-voxel MTFs. Decoders were trained on a subset of the data (training) for each acoustic feature

separately (Figure 1). Decoding accuracies were obtained as Pearson's *r* between predicted and actual acoustic features in the testing data. Results were compared across age groups using Wilcoxon rank-sum tests.
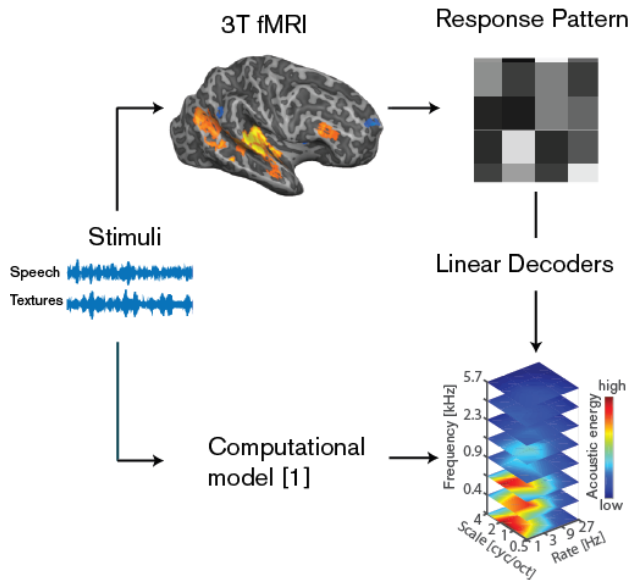

Figure 1: Model-based decoding. Participants listened to a story embedded in sound textures while we acquired 3T-fMRI data. The sound mixture was decomposed into its modulation spectrum (Chi et al., 2005) [1]. Linear decoders were trained on the response patterns in auditory cortex for each feature of the modulation representation. Predictions were then tested on a left-out testing data set.

Results

Topographical best feature maps were obtained from univariate encoding models by marginalizing the single-voxel MTFs for the dimension of interest. In particular, best frequency maps were obtained by averaging over the rate and scale dimension, and assigning the frequency with the maximal response to a given voxel.
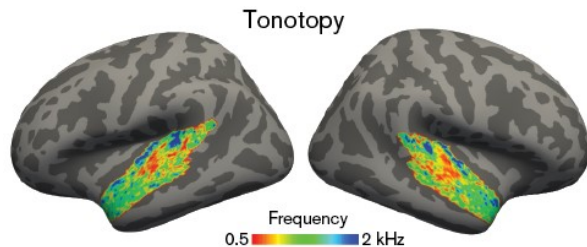

Figure 2: Tonotopy group map for older adults.

Tonotopic maps showed the typical mirror-symmetric frequency gradients along Heschl's gyrus in both young and aged auditory cortex (Figure 2), irrespective of age. Topographical best feature maps for rate and scale were more complex; but were highly correlated between age groups.

Cortical sensitivity to acoustic features was quantified using multivariate model-based decoding. Decoding yielded highest accuracies at frequencies of 230–580 Hz (median $r = 0.52$), scales of 0.25 cyc/oct ($r = 0.5$), and rates of 4–8 Hz ($r = 0.42$), irrespective of age (Figure 3a). However, we found age differences in specificity of tuning to temporal rate as quantified by the ratio of maximum to mean response: Specificity of temporal rate tuning was higher in young than in older participants. Lower specificity (expressed as the best temporal rate tuning divided by the average rate tuning in auditory cortex) is consistent with a broadening of the temporal tuning in the aged auditory cortex (Figure 3b). In the older group, specificity of rate tuning did not correlate with hearing loss (pure tone average) though ($rho = 0.36$, $p = 0.1$).
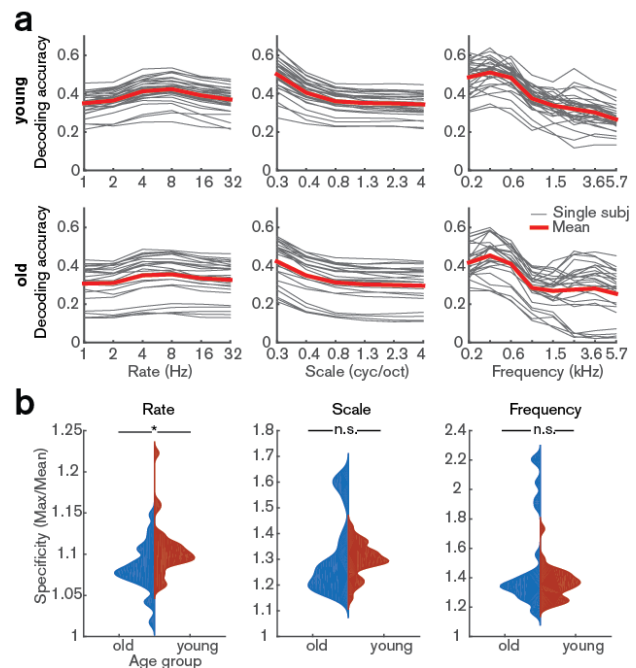

Figure 3: *a,* Marginal profiles of multi-voxel MTFs for temporal rate, spectral scale and frequency for auditory cortex of young and older subjects. *b,* Distribution of specificity of tuning (quantified by the ratio of maximum to mean decoding accuracy), compared between age groups. Specificity is higher in young than older adults for temporal rate but not spectral scale and frequency. *$p < 0.05$.

Conclusions

Although the large-scale topographic organization of acoustic features appears to be preserved in the aged auditory cortex, age-related differences in the marginal profiles of multi-voxel MTFs were evident. Tuning to temporal rate was sharper in young than in older participants. These results suggest that rate specificity declines in the healthy aging auditory cortex.

Acknowledgements

References

Chi-Wing N, Recanzone GH (2017) Age-Related Changes in Temporal Processing of Rapidly-Presented Sound Sequences in the Macaque Auditory Cortex. Cerebral cortex 1-22.

Chi T, Ru P, Shamma SA (2005) Multiresolution spectrotemporal analysis of complex sounds. The Journal of the Acoustical Society of America 118:887-906.

Erb J, Armendariz M, De Martino F, Goebel R, Vanduffel W, Formisano E (2018) Homology and Specificity of Natural Sound-Encoding in Human and Monkey Auditory Cortex. Cerebral cortex.

Gordon-Salant S, Yeni-Komshian GH, Fitzgibbons PJ, Barrett J (2006) Age-related differences in identification and discrimination of temporal cues in speech segments. The Journal of the Acoustical Society of America 119:2455-2466.

Hullett PW, Hamilton LS, Mesgarani N, Schreiner CE, Chang EF (2016) Human Superior Temporal Gyrus Organization of Spectrotemporal Modulation Tuning Derived from Speech Stimuli. The Journal of neuroscience : the official journal of the Society for Neuroscience 36:2014-2026.

McDermott JH, Simoncelli EP (2011) Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. Neuron 71:926-940.

Santoro R, Moerel M, De Martino F, Goebel R, Ugurbil K, Yacoub E, Formisano E (2014) Encoding of natural sounds at multiple spectral and temporal resolutions in the human auditory cortex. PLoS computational biology 10:e1003412.

Santoro R, Moerel M, De Martino F, Valente G, Ugurbil K, Yacoub E, Formisano E (2017) Reconstructing the spectrotemporal modulations of real-life sounds from fMRI response patterns. Proceedings of the National Academy of Sciences of the United States of America 114:4799-4804.