Experimental evidence on computational mechanisms of concurrent temporal channels for auditory processing

Xiangbin Teng (xiangbin.teng@gmail.com)
Max-Planck-Institute for Empirical Aesthetics, Grüneburgweg 14
60322 Frankfurt am Main, Germany

David Poeppel (david.poeppel@gmail.com)
Max-Planck-Institute for Empirical Aesthetics, Grüneburgweg 14
60322 Frankfurt am Main, Germany
Department of Psychology, New York University, 6 Washington Place
New York, NY, USA 10003

Abstract:

Natural sounds convey perceptually relevant information over multiple timescales, and the necessary extraction of multi-timescale information requires the human auditory system to work over distinct ranges. Here, we show behavioral and neural evidence that acoustic information at two discrete timescales (~ 30 ms and ~ 200 ms) is preferably coded and that the theta and gamma neural bands of the auditory cortical system correlate with temporal coding of acoustic information. We then propose an computational approach to investigate how the cortical auditory system implements canonical computations at the two prominent timescales – the auditory system constructs a multi-timescale feature space to achieve sound recognition.

Keywords: multiplexing; temporal discretization; auditory perception; behavior; neuroimaging

Instruction

Correctly perceiving behaviorally significant sounds—speech, music, and the acoustic environment—requires integrating acoustic information over time to extract relevant regularities. A fundamental question about this process is: How does the auditory brain integrate information of continuously varying sounds, typical of many natural auditory signals? To derive the appropriate perceptual representations, the auditory system must extract rapidly varying information on a scale of milliseconds (approximately 10–50 ms), operating with high temporal resolution, and concurrently analyze more slowly varying signal attributes on a scale of hundreds of milliseconds (about 150–300 ms), enabling sufficient spectral resolution (Poeppel, 2003).

Here, we first present behavioral evidence that the human auditory system extracts fine-grained acoustic details at the timescale of ~ 30 ms and integrates auditory information over a timescale of ~ 200 ms to abstract global acoustic features (Teng, Tian, & Poeppel, 2016). We then looked for neural correlates for the behavioral finding by entraining the human auditory system using sounds with acoustic modulations across a wide-ranging timescales. Indeed, we found robust neural coding in the theta (4-7 Hz) and gamma (30 – 45 Hz) bands, but in between (8 – 30 Hz) (Teng & Poeppel, 2019; Teng, Tian, Rowland, & Poeppel, 2017).

Behavioral Evidence

We employed a same-different paradigm and tested 10 participants on extraction of local and global acoustic features. In the local task, two sounds were presented, which only differed in the local details; in the global task, the global pitch contours of the sounds were different. The stimulus length was varied. The participants' performance was best below 50 ms in the local task but above 200 ms in the global task.

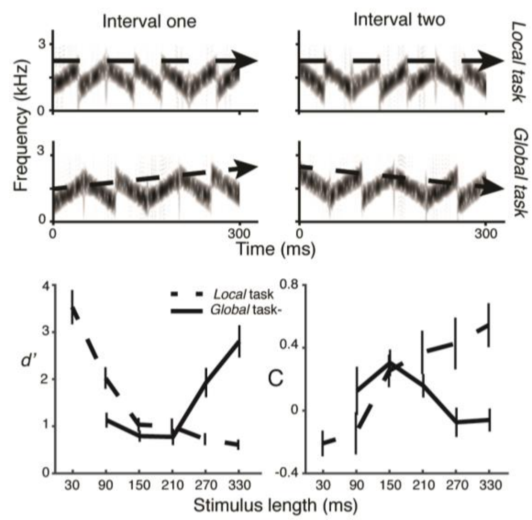

Figure 1: the upper panel shows an example of spectrograms for the 'different' pair of stimuli is plotted separately for the *local task* and the *global task*. The lower panel shows the results using segment duration

of 30 ms, which demonstrate the opposite effect of the stimulus length on the local and global tasks.

Neural Evidence

We generated sounds with temporal modulations across a wide temporal range: 5, 10, 17, 25, 37 Hz and presented those sounds to 15 participants while measuring their neural signals using Magnetoencephalography. We found robust neural entrainment in the theta and gamma bands.

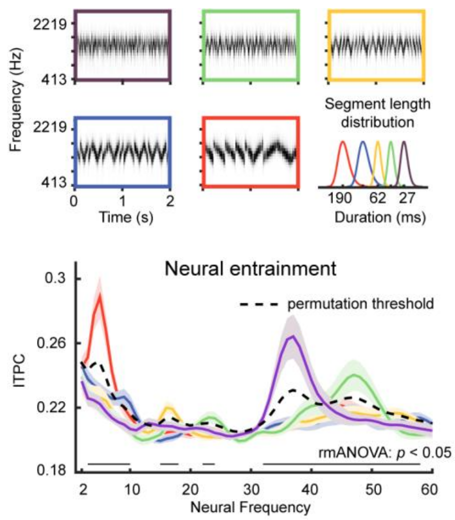

Figure 2: the upper panel shows an example of spectrograms for sounds with modulations in various temporal ranges. The lower panel shows the finding of neural entrainment.

Computational Approach

The behavioral and neural evidence awaits an computational explanation to why the human auditory system evolves to primarily work on two discrete timescales and how this specific coding scheme aids sound recognition. We propose that, as acoustic signals in natural sounds are temporally correlated within specific timescales (Rosen, 1992; Singh & Theunissen, 2003) and the same auditory object can be represented in sounds scaled to different temporal ranges (i.e. stretched or compressed speech), the auditory system may sparsely extract a set of auditory features at the two discrete timescales and use the features to construct a multi-dimensional/timescale feature space, which is robust to temporal scaling. We train LSTM models with different combinations of time constants to recognize natural sounds (i.e. speech and environmental sounds) and extract features of corresponding timescales. Such features are used to construct a multi-timescale feature space to investigate how different combinations of timescales maximize both feature separation and performance on sound recognition.

Acknowledgments

The studies were supported by NIH 5R01DC005660 to DP and the Max-Planck-Society.

References

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as "asymmetric sampling in time." Speech Commun., 41(1), 245–255. http://doi.org/10.1016/S0167-6393(02)00107-3

Rosen, S. (1992). Temporal Information in Speech: Acoustic, Auditory and Linguistic Aspects. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 336(1278), 367–373. http://doi.org/10.1098/rstb.1992.0070

Singh, N. C., & Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. The Journal of the Acoustical Society of America, 114(6), 3394–3411. http://doi.org/10.1121/1.1624067

Teng, X., & Poeppel, D. (2019). Theta and Gamma Bands Encode Acoustic Dynamics over Wide-ranging Timescales. bioRxiv, 547125. http://doi.org/10.1101/547125

Teng, X., Tian, X., & Poeppel, D. (2016). Testing multi-scale processing in the auditory system, 6, 34390. http://doi.org/10.1038/srep34390

Teng, X., Tian, X., Rowland, J., & Poeppel, D. (2017). Concurrent temporal channels for auditory processing: Oscillatory neural entrainment reveals segregation of function at different scales. PLOS Biology, 15(11), e2000812. http://doi.org/10.1371/journal.pbio.2000812