# ON PRENATAL AUDITORY EXPERIENCE IN HUMANS AND ITS RELEVANCE FOR MACHINE HEARING

**Marin Vogelsang**[1*] **Lukas Vogelsang**[1,2*] **Sidney P. Diamond**[1] **Pawan Sinha**[1]

[1] Massachusetts Institute of Technology, Cambridge, MA, USA
[2] École Polytechnique Fédérale de Lausanne, Switzerland

## ABSTRACT

Given the markedly better generalization capabilities of the human perceptual system relative to computational models, a question naturally arises about the genesis of this disparity. Here, we propose that a key to robust human perception might lie in its developmental trajectory. Unlike standard computational training procedures, perceptual development in humans undergoes a stereotypical temporal progression in which sensory inputs are initially highly degraded and gain quality later on. We focus here on the auditory domain, in which this progression commences already before birth: A fetus' experience in the womb comprises low-pass filtered versions of voices and other sounds in the environment. Such degraded inputs may induce the acquisition of mechanisms capable of performing extended temporal integration, facilitating robust analysis of information carried by slow variations in the auditory stream, such as emotions or other prosodic content. To computationally test this proposal, we assessed the consequences of training with different temporal progressions of filtered audio signals on a deep convolutional neural network's internal representations and subsequent classification of emotional prosodic content. We found that training with an auditory trajectory approximately mimicking the pre-to-post-natal progression yielded best generalization performance; it significantly exceeded outcomes following exclusively full-frequency, exclusively low-frequency, or inverse-developmental training protocols. The developmentally-trained model further acquired temporally extended receptive fields in its first convolutional layer and, when tested with full-frequency inputs, exhibited the strongest resilience to the ablation of units tuned to high frequencies. These simulations suggest that the progression from low-to-full-frequency signals, rather than being an epiphenomenon, may be an enabling feature of perceptual development, conferring significant benefits to later auditory processing. The results also point to the utility of incorporating similar procedures into the training of computational model systems and, more generally, to the inspiration that human development may provide towards the goal of achieving more robust generalization.

[*] Equal contribution