Automatic extraction of infant vocalizations from one year home audio recordings

Guillem Bonafos, Clara Bourot, Pierre Pudlo, Jean-Marc Freyermuth, Laurence Reboul, Thierry Legou, Samuel Tronçon, Raoul Desbrières, Jean-Michel Bartoli & Arnaud Rey

I2M, Aix Marseille Univ, CNRS; LPC, Aix Marseille Univ, CNRS; LPL, Aix Marseille Univ, CNRS;ILCB, Aix Marseille Univ; Résurgences R&D, Arles; Department of Obstetrics and Gynecology, Saint Joseph Hospital, Marseille.

We introduce an innovative methodology for the automated extraction of infant vocalizations from extended home audio recordings. Our approach relies on a neural network trained with 72 minutes of meticulously curated infant vocalization examples (babblecor corpus) and soundscape segments derived from the original audio recordings. The trained model is then employed to construct a novel database exclusively dedicated to detect infant vocalizations. Over the course of one year, from birth to first birthday, we conducted extensive audio recordings capturing the vocalizations of 15 children within the confines of their homes, amassing a data set of 3,214 hours, or 2,1 To of data, recorded with a stereophonic microphone, with a 44,100 Hz sampling rate. Employing our neural network, we systematically extract vocalization segments from these extensive recordings, providing a detailed corpus of each child's vocal productions throughout the year. This methodology not only showcases the efficacy of our approach but also highlights the potential of the resulting database for advancing research in the field of infant vocal productions.