voice: easier audio analysis for digital phenotyping

Filipe J. Zabala[1]
Graduate Program of Psychiatry and Behavioral Sciences, UFRGS, Porto Alegre - RS - Brazil
Giovanni A. Salum[2]
Graduate Program of Psychiatry and Behavioral Sciences, UFRGS, Porto Alegre - RS - Brazil

ABSTRACT

A significant barrier to progress in psychiatric research has been the difficulty to accurately and reliably quantifying relevant phenotypes to mental disorders. Current practice mainly relies on questionnaires and interviews, but recent interest has been generated by the use of information present in our daily lives such as voice analysis. Under this perspective `voice` package was developed, aiming to make audio analysis easier for digital phenotyping. We present new functionalities and provide some initial validation for the prediction of sex. This non-invasive biometrical tool may also be used to automatically detect and split spoken parts on recorded audio files, as well as detect speechers. The package may help to easily extract and summarize features endorsed by the specialized literature, like fundamental frequency (F0), formants (F1 to F8), Mel Frequency Cepstral Coefficients (MFCC), Reflection Coefficients (RFC) and Zero-Crossing Rate (ZCR), to name a few.

Keywords: voice; biometrics; digital phenotyping; speecher recognition.