A Comprehensive Methodology for Voice Segmentation of Large Sets of Speech Files Recorded in Naturalistic Environments

Authors : Ana Londral, Burcu Demiray, Marcus Cheetham

Abstract : Speech recording is a methodology used in many different studies related to cognitive and behaviour research. Modern advances in digital equipment brought the possibility of continuously recording hours of speech in naturalistic environments and building rich sets of sound files. Speech analysis can then extract from these files multiple features for different scopes of research in Language and Communication. However, tools for analysing a large set of sound files and automatically extract relevant features from these files are often inaccessible to researchers that are not familiar with programming languages. Manual analysis is a common alternative, with a high time and efficiency cost. In the analysis of long sound files, the first step is the voice segmentation, i.e. to detect and label segments containing speech. We present a comprehensive methodology aiming to support researchers on voice segmentation, as the first step for data analysis of a big set of sound files. Praat, an open source software, is suggested as a tool to run a voice detection algorithm, label segments and files and extract other quantitative features on a structure of folders containing a large number of sound files. We present the validation of our methodology with a set of 5000 sound files that were collected in the daily life of a group of voluntary participants with age over 65. A smartphone device was used to collect sound using the Electronically Activated Recorder (EAR): an app programmed to record 30-second sound samples that were randomly distributed throughout the day. Results demonstrated that automatic segmentation and labelling of files containing speech segments was 74% faster when compared to a manual analysis performed with two independent coders. Furthermore, the methodology presented allows manual adjustments of voiced segments with visualisation of the sound signal and the automatic extraction of quantitative information on speech. In conclusion, we propose a comprehensive methodology for voice segmentation, to be used by researchers that have to work with large sets of sound files and are not familiar with programming tools.

Keywords : automatic speech analysis, behavior analysis, naturalistic environments, voice segmentation

Conference Title : ICSST 2017 : International Conference on Speech Science and Technology

Conference Location : Barcelona, Spain

Conference Dates : February 26-27, 2017