THE USE OF PRAAT FOR STUDYING THE VOICE

Paul Boersma
University of Amsterdam
paul.boersma@uva.nl

The computer program Praat is the most used software in the world when it comes to speech research. In this talk I illustrate that this software, which is freely available for Windows, Mac and Linux computers, can be employed by anybody who studies the voice, be it speech therapists, physicians, voice engineers, or singers.

Built into Praat are several standard measurements of voice quality. Thus, when you use your microphone to record a speech utterance into Praat and view the resulting sound's waveform, you will be able to select a part of this sound (by dragging with your mouse, similarly to how you select text in Microsoft Word) and specify several kinds of analyses. One of those analyses is simply called the "Voice Report", and choosing this analysis will give you information on various vocal aspects of the selected part of the utterance, such as its average vocal-fold vibration frequency, how much of the selected part is voiced, several measures of the irregularity in the vocal frequency (jitter), several measures of the irregularly in the strength of the vocal pulses (shimmer), and several measures of non-periodic additive noises (harmonicity).

For specialized applications for which Praat has no built-in buttons, you (or your programmer) can assemble your own analysis, combining Praat's building blocks to create your specialized analysis with the help of Praat's scripting language, as the Belgian speech pathology community can testify (Youri Maryn's AVQI script for measuring the severity of dysphonia).

Although Praat has existed since 1992, and some of its current algorithms, especially the ones that nobody has been able to improve upon, data from that time, the authors of Praat regularly include new types of analyses and improve existing analyses on the basis of the recent literature and their own research. A project running in 2017 involves the improvement of vocal frequency measurement algorithms by deep artificial neural networks that are trained for optimal performance on speech recordings with known vocal frequencies (as measured by electroglottography or as generated by simulation). In this way, Praat's voice quality measurements continue to approach the ground truth of the underlying vocal fold events, separating the various components of voice quality problems. In the talk I show some recent results in this area.