

# Voxalys – a Pedagogical Praat Plugin for Voice Analysis

Jonas Lindh

Department of Linguistics, Göteborg University

## **Abstract**

*Praat is an excellent free software for voice analysis. However, the flexibility of the program and its many possibilities also means that it is not particularly user friendly. Voxalys is a plugin for Praat containing the most basic functions for voice analysis. The plugin automatically creates and saves pictures, reports and sound files in a pre-defined folder structure to facilitate analysis and handling of files for direct feedback to clients or analyst and documentation. The advantage with the plugin is its flexibility and the fact that a complete analysis can be produced by just a few clicks.*

## **Introduction**

Teaching computer based voice analysis for a few years using Praat has been very rewarding. The possibilities the Praat program offers are almost inexhaustible. The price the user has to pay for this enormous flexibility is the fact that the program is not particularly user-friendly, especially if one wants to make more complicated analyses. Students often get confused by the multitude of options that they have to choose between. For some, the many degrees of freedom is an almost insurmountable obstacle. Praat has a built-in scripting language that make it possible for anyone with some knowledge of programming to create scripts and run them with Praat. To facilitate basic relevant voice analyses often used by speech therapists, the Voxalys Praat plugin was created with two sets of functions. The first set consists of recording functions to facilitate recording, saving and documenting voices and utterances automatically. The second set contains six analysis functions that are applied to the recorded sound objects. The analysis functions perform analyses, print images and save pictures in a predetermined folder structure to facilitate the handling of sound, picture or report files.

## **Recording Functions**

The recording functions are implemented as scripts to facilitate quick recording and saving. Starting the function first presents options to the user to record with or without a recording interface and options to save as a file. An option to automatically remove silence in the beginning

and end of the recording is also given. Filename is always based on the user's input of the client name plus date and time.

### *Recording with interface*

This option will present Praat's built-in SoundRecorder window together with instructions on how to use it. In the SoundRecorder the user can record multiple times until satisfied. After that, an object can be saved in the object window and a prompting window allows the user to continue. If the user chooses to save the recording the recording is automatically saved according to above and the object is named after the client name given by the user.

### *Recording without interface*

If the user is not in the need of a SoundRecorder window, the option of just using the default microphone settings and choosing recording time in seconds can be used. If the saving option is chosen the recording will be automatically saved as a file named after the client name the user has given plus date and time.

## **Analysis Functions**

When a sound has been recorded and is marked in Praat there are several options to choose from. After installing Voxalys there will be a menu with the same name for sound objects. The menu contains options for analysis of a recorded sound.

### *To F0-analysis... (Intonation)*

This function gives the user the opportunity to visualize intonation on a semitone scale. The

only input the script needs from the user is a floor and a ceiling for the F0 extraction algorithm based on Boersma (1993). The resulting image is accompanied by relevant numbers for date, duration, mean, median and standard deviation for F0.

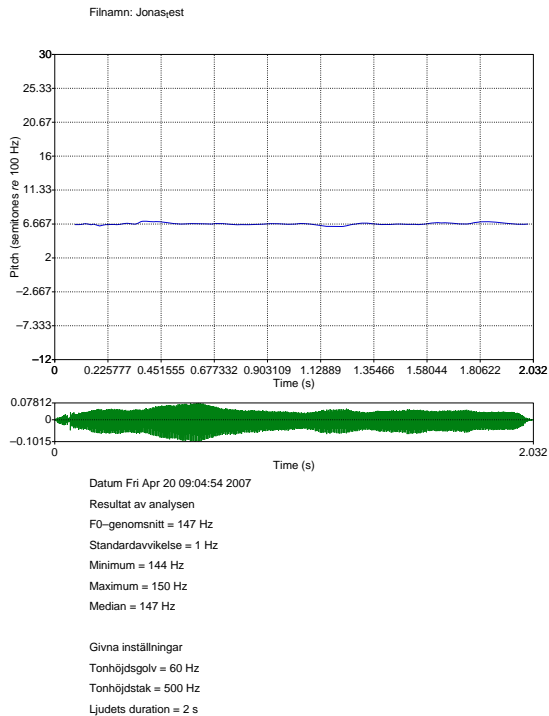


Figure 1. F0-analysis example picture from the author uttering a vowel (text in Swedish).

The picture can be used as feedback for the client's intonation of a certain word or phrase or just a visualization of the stability in fundamental frequency. If the saving option is chosen a picture file is saved in the folder structure under the client's name and picture folder.

### To F0-histogram...

An easily interpretable way to demonstrate the degree of stability and variation of F0 is to draw a histogram. The user supplies the script with floor and ceiling for the F0 extraction, the frequency range of the histogram, bin width (default 1 Hz), colour, line width and font size.

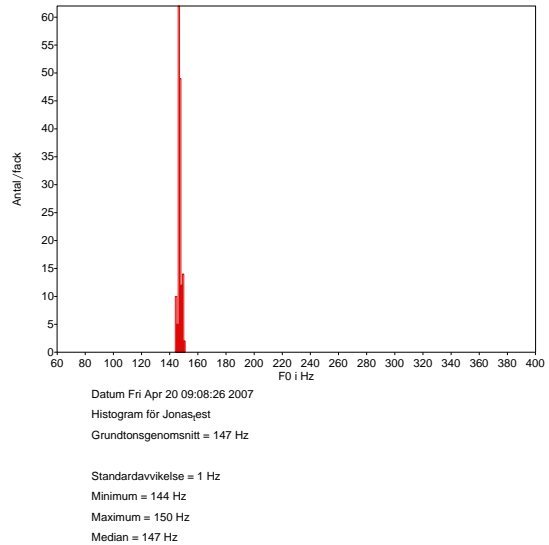


Figure 2. F0 histogram example picture from the author uttering a vowel (text in Swedish).

A sustained utterance of a single vowel will produce a very dominated bar at the centre fundamental frequency mode (Rose, 2002). An unstable voice will instead show bars distributed over a wider range or clusters of bars in the case of creakiness. Saving is processed in the same way as mentioned above.

### To Phonetogram...

This function gives the user the opportunity to combine the analyses of variation in vocal effort (expressed as SPL in dB) and fundamental frequency. SPL is plotted on the y-axis and frequency on the x-axis. The user provides floor and ceiling for both frequency (in Hz) and SPL (in dB), values that will be used for both the analysis and the display. It is also possible to change the colour of the plotting for example to illustrate the difference between multiple samples.

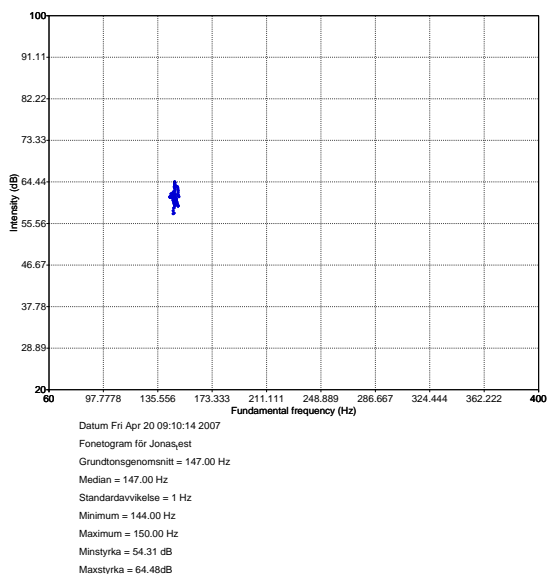


Figure 3. Phonetogram example picture from the author uttering a vowel (text in Swedish).

A phonetogram can partly be used to illustrate the instability of the voice by the spread of the data points. It can also be used for training, where a client can try to expand the range of their voicing abilities. This can also be done interactively in a so called Voice Range Profile (see below).

### To Perturbation analysis...

The output of the perturbation analysis can be a text file, an html-file or just a prompt with a detailed report (in a preferred language) on the voice. The report has format similar to Praat's built-in function producing a voice report (<http://www.fon.hum.uva.nl/praat/manual/Voice.html>). The report gives standard measures for F0 like mean, median and standard deviation. Below that follow measures of voice breaks, defined as fractions of locally unvoiced F0 frames. The number of voice breaks is given by the distances between consecutive pulses that are longer than 1.25 divided by the floor of the F0 analysis. The degree of voice breaks is then the total duration of the breaks between voiced parts

([http://www.fon.hum.uva.nl/praat/manual/Voice\\_1\\_Voice\\_breaks.html](http://www.fon.hum.uva.nl/praat/manual/Voice_1_Voice_breaks.html)). This is followed by the results of five different jitter and six shimmer measures and thresholds for pathology for those. Finally, a HNR (harmonic to noise ratio) is presented. All measures are printed with explanations and thresholds. An illustration of the voice breaks is presented in the picture

window. The measures and thresholds are partly based on results from Deliyski (1993) and used in the commercial software MDVP (Boersma, 2004). The purpose is to give the user (and possibly the client) a detailed report of the degree of instability in a voice.

### To Complete Voice Report... (.html)

This function exists for the user to be able to quickly receive an interpretable report on a voice for printout to for example a client. The function presents the options to the user to tick the preferred analyses presented above and give the frequency range for the F0 extraction. The script will then print and save all those analyses and present a report as a web page with images and figures and report one jitter and one shimmer measure and the comparison to the threshold for a pathological voice.

### Voice Range Profile (VRP)

To be able to practice the range of a voice for both pitch and loudness a script for performing a voice range profile is available in the Voxalys plugin. The user is prompted to choose an amount of time for recording and range for both F0 and amplitude within which the range is measured. VRP produces the same output as the phonetogram described above with the exception that it is interactive, i.e., is plotted while recorded. To be able to follow the voice production over time, the change of colour indicates the voicing development over time.

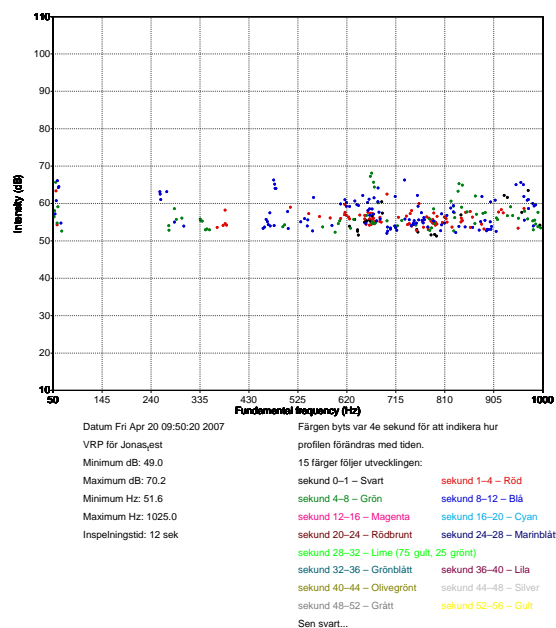


Figure 4. Voice Range Profile (VRP) example picture from the author uttering a vowel (text in Swedish) in a very noisy environment for 12 sec.

VRP can be used either as interactive training for the vocal range in vocal effort and pitch or as documentation and feedback for a client. After treatment a new profile can be done to visualize possible improvements. This kind of training and documentation might also be interesting for vocal performers (singers, actors, etc.).

### *A predefined folder structure*

All images, sound files and reports that the users choose to save are saved in a predefined folder structure. This function is implemented as a help for the user. The folder structure is built automatically in the Voxalys plugin directory when the users choose to save and it is created as shown in figure 5 below.

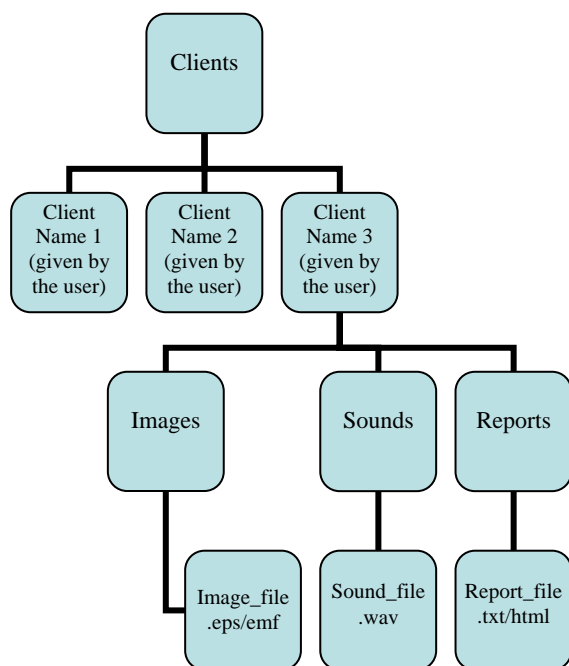


Figure 5. Voxalys predefined folder structure.

This might be considered as a constraint for the user. However, a user can either move files easily by drag and drop from the folder structure (accessible from the start menu in windows) or copy and paste the files to a different setting to their preference. A more advanced user can simply just adjust scripts to their personal preference of structure.

### **Future developments**

The possibility of developments for Voxalys is more or less inexhaustible in the same way as Praat. However, the current version is working well enough to be used for the basic voice analysis and recording functions described above. It is also possible to easily build

individual functions or make changes in existing scripts. The next step is to develop a version of Voxalys more directed towards singing. That version will include functions like hit-the-note and LTAS analysis for singing formant (Sundberg, 1977).

### **References**

- Boersma, P. & Weenink, D. (2007). Praat: doing phonetics by computer (Version 4.5.18) [Computer program]. Retrieved March 30, 2007, from <http://www.praat.org/>
- Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. Proceedings of the Institute of Phonetic Sciences 17: 97-110. University of Amsterdam.
- Boersma, P. (2004). Stemmen meten met Praat. Stem-, spraak-, en taalpathologie 12, 237-251.
- Deliyiski, D. (1993). Acoustic model and evaluation of pathological voice production. Proceedings Eurospeech '93, Berlin. 1969.1972.
- Rose, P. (2002). Forensic Speaker Identification. New York, Taylor & Francis.
- Sundberg, J. (1977). The Acoustics of the Singing Voice, Scientific American Mar 77, p82