

Speak and unSpeak with PRAAT

By Paul Boersma and Vincent van Heuven

Introduction

By the Goodies Editor, Rob Goedemans

Many linguists use recorded speech in their research. In descriptive work, visual representations of such recordings (mostly oscillograms) are often annotated with IPA symbols and other labels, and then used to illustrate a phenomenon or defend a certain position regarding the nature of some phonetic or phonological property of the language in question. In phonetic and psychophysical research some parameter of the recorded speech (like tempo or intensity) is often altered, after which the new sound thus obtained is used in an experiment to test the sensitivity of the human ear, or brain, to certain speech properties.

The introduction of the computer has brought about a virtual revolution in the linguistic sciences with respect to the usage of speech recordings. A lab full of cumbersome machinery has now been replaced by one PC, Mac or workstation, on which anyone who puts his mind to it can record, annotate and modify speech with some simple commands or a few mouseclicks. Even the calculation of some speech parameters that were rather complicated to obtain in the past (like pitch and spectral analysis) but frequently used in phonetic research nonetheless, are now often just one or two mouseclicks away.

As a result, a growing number of colleagues find use for files with speech sounds in their linguistic explorations. The needs of this group are served by a rather small number of software packages designed for the representation, annotation and analysis of speech (and much more in many cases). In my opinion, one of these stands out in many ways. It is called "PRAAT"; the imperative form of *to speak* in Dutch. Since this package is rapidly gaining in popularity, we have decided to devote some attention to it in this issue. First, one of the authors, Paul Boersma, introduces the package and outlines its impressive functionality. Then, an experienced user, Vincent van Heuven, highlights some of the advantages and disadvantages of using PRAAT in everyday phonetic research.

It is my sincere hope that these two goodies will convince even more linguists to download PRAAT and experiment with it a little. They will see that incorporating, for example, some oscillograms of minimal pairs in their work is as easy as ABC. Their publications will undoubtedly be the better (and the livelier) for it.

PRAAT is a computer program for analysing, synthesizing, and manipulating speech. It has been developed since 1992 by Paul Boersma and David Weenink at the Institute of Phonetic Sciences of the University of Amsterdam. There are versions for most of the common operating systems: Macintosh, Windows, Linux, and several Unix workstations (Solaris, Silicon Graphics, Hewlett-Packard). By September 2001, there were more than 5,000 registered users in 99 countries.

PRAAT, a system for doing phonetics by computer

By Paul Boersma

1. Analysing speech with PRAAT

PRAAT allows you to record a sound with your microphone or any other audio input device, or to read a sound from a sound file on disk. You will then be able to have a look 'inside' this sound. The upper half of the sound window (see figure 1) will show you a visible representation of the sound (the *wave form*). The lower half will show you several acoustic analyses: the *spectrogram* (a representation of the amount of high and low frequencies available in the signal) is painted in shades of grey; the *pitch contour* (the frequency of periodicity) is drawn as a cyan curve; and *formant contours* (the main constituents of the spectrogram) are plotted as red dots.

PRAAT is most often used with speech sounds, in which case the pitch contour is associated with the vibration of the vocal folds and the formant contours are associated with resonances in the vocal tract. But the use of PRAAT is certainly not limited to speech sounds: musicians and bio-acousticians use it for the analysis of sounds produced by flutes, drums, crickets, or whales, and the interpretation of the three analyses will change accordingly.

The Sound window allows you to zoom in for more detail, to scroll to the places that you are

Paul Boersma, Institute of Phonetic Sciences, University of Amsterdam, Herengracht 338, 1016 CG Amsterdam, The Netherlands, Paul.Boersma@hum.uva.nl

Vincent van Heuven, Universiteit Leiden Centre for Linguistics (ULCL), P.O. Box 9515, 2300 RA Leiden, The Netherlands, V.J.J.P.van.Heuven@let.leidenuniv.nl

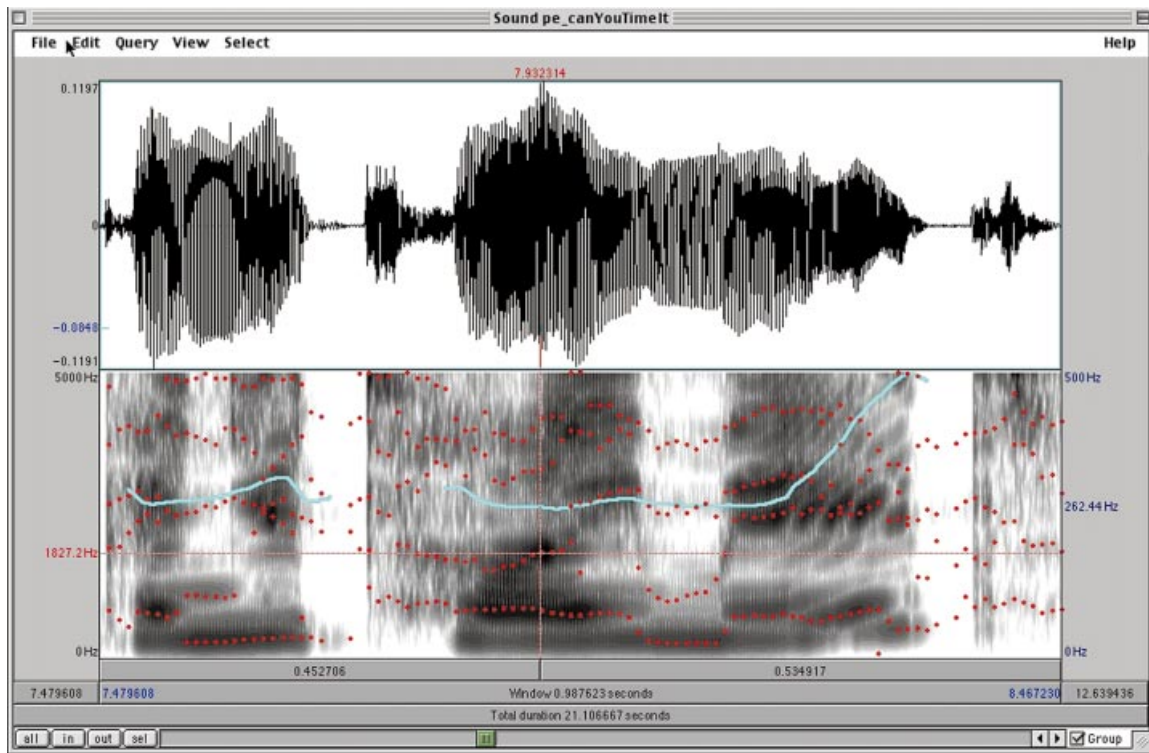


Figure 1. Praat's sound window.

interested in, to set a time cursor or select a time stretch, and to listen to the parts of the sound that you are viewing or selecting. You can easily query all the important properties of the analyses, e.g. obtain the average pitch value inside the selected time stretch. You can turn the analyses into separate objects (independent from the original sound), which is handy for further processing, e.g. it allows the

pitch contour to be saved, printed, or converted into something else.

2. Annotating speech with PRAAT

PRAAT is used by many linguists (phoneticians, phonologists, syntacticians) to label and segment their speech recordings. You can make transcriptions and

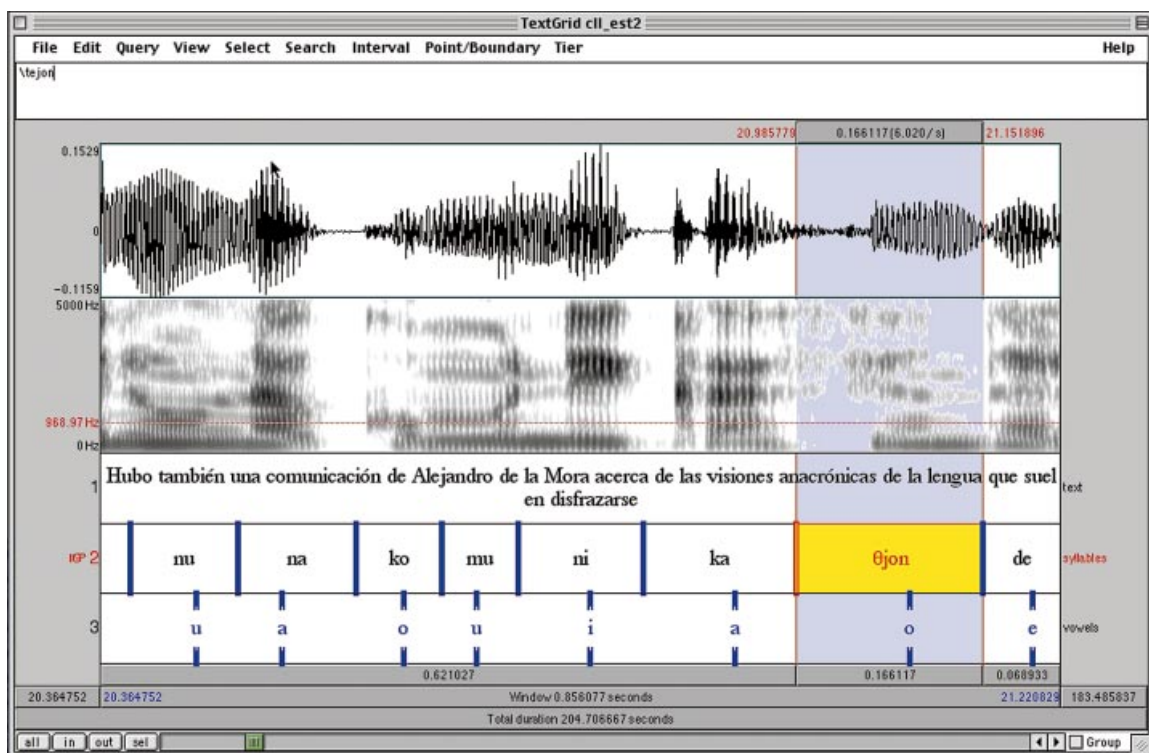


Figure 2. Praat's annotation window.

annotations on multiple levels simultaneously (see the three levels in figure 2), in a window that typically also shows visible representations of the sound, the spectrogram, and perhaps the pitch contour. PRAAT supports an easy use of special symbols in annotations, including nearly all symbols defined by the International Phonetic Association (such as the θ symbol, typed as “\te”, in figure 2).

3. Synthesizing speech with PRAAT

PRAAT is *not* a text-to-speech system: you cannot type in an English sentence and have the program read it aloud. But you can generate many types of sounds with PRAAT. First, you can use formulas to generate simple sounds like sine waves or white noise from scratch, or to generate more complicated sounds from other sounds. Second, you can create sounds from other types of data, e.g. you can turn a pitch contour in a pulse train. Third, you can do *source-filter synthesis*: from stylized pitch, intensity, and formant contours that you can build from scratch, you can create speech-like sounds. Fourth, you can perform *articulatory synthesis*: from a specification of timed muscle contractions, Praat will compute the resulting sound. Fifth, you can create sounds from other sounds by a variety of filtering and enhancement techniques.

4. Manipulating speech with PRAAT

A specialized manipulation window allows you to stylize and modify the pitch contour of an utterance.

In figure 3, the impatient-sounding question “can you time it?”, with an original final high rise, has been converted into a slightly whining command. The same window allows you to modify relative durations within this utterance. In this way, you can change the intonation and stress patterns of the utterance, which is useful when creating stimuli for research into the perception of prosody.

5. Graphical capabilities of PRAAT

PRAAT comes with a separate Picture window into which you can draw your sounds, pitch contours, spectrograms, and any other data types. You can add text (several fonts, many special symbols, several sizes, any rotation), lines (several colours, any widths, several styles), circles/ellipses/rectangles (filled or outlined), and several types of markers along and inside your drawings. Figure 4, for instance, shows the modified pitch contour of figure 3, with appropriate vertical text to its left, and a comment added above it.

The Picture window is designed for producing publication-quality graphics for your articles and dissertations. From this window, you can print to any printer (PostScript, Macintosh, Windows) and save your drawings as EPS files (best quality, but works with PostScript printers and PDF creators only), WMF files (Windows), or PICT files (Macintosh). All of these can be easily imported into your word processor. The Macintosh and Windows versions support the graphical clipboard as well, so that you can use simple copy-and-paste to move PRAAT

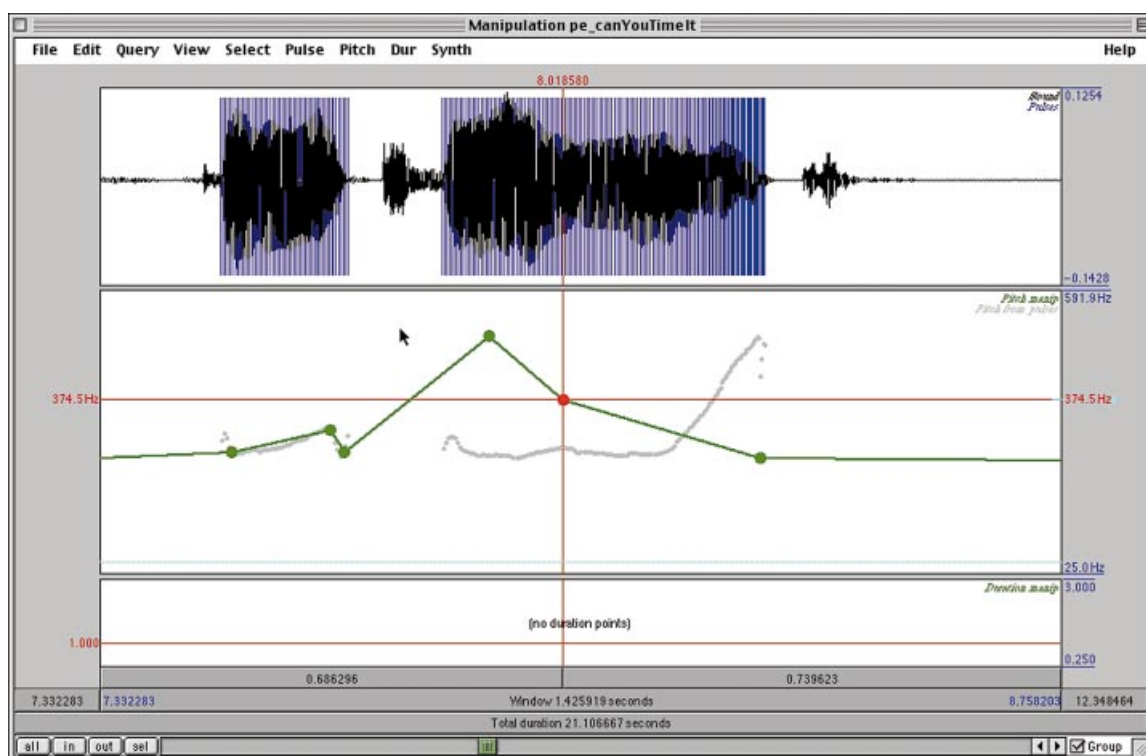


Figure 3. Praat's manipulation window.

pictures to your word processor, if you have no use for PostScript quality.

6. The PRAAT scripting language

In most parts of the world, slavery was abolished in the 19th century, but it is not unusual to see phoneticians measure the pitch values of 1500 vowels by hand. You would probably want to replace such work by an automated procedure that, say, loops over all the sound files that reside in a certain directory or over all the segments marked “u” in a vowel annotation. Such things can easily be performed by the PRAAT scripting language, which is a general-purpose programming language with special capabilities for simulating menu choices and button presses in the PRAAT program. Many people use this language for all their analyses, tabulations, statistics (there are special functions for computing levels of significance in *t*, χ^2 , or *F* tests), and complicated pictures. In fact, you can use PRAAT as a general drawing program: figure 4 shows a PRAAT script that draws the complicated figure at the top of the Picture window.

7. Other features of PRAAT

The PRAAT program contains several possibilities in areas that are only remotely connected to phonetics. Phonologists and syntacticians like its implementation of Optimality-Theoretic learning (constraint demotion, gradual learning algorithm, robust interpretive parsing), which you can apply to your own cases. Other possibilities include neural-net

modelling and extensive high-level statistics (principal-component analysis, discriminant analysis, multidimensional scaling).

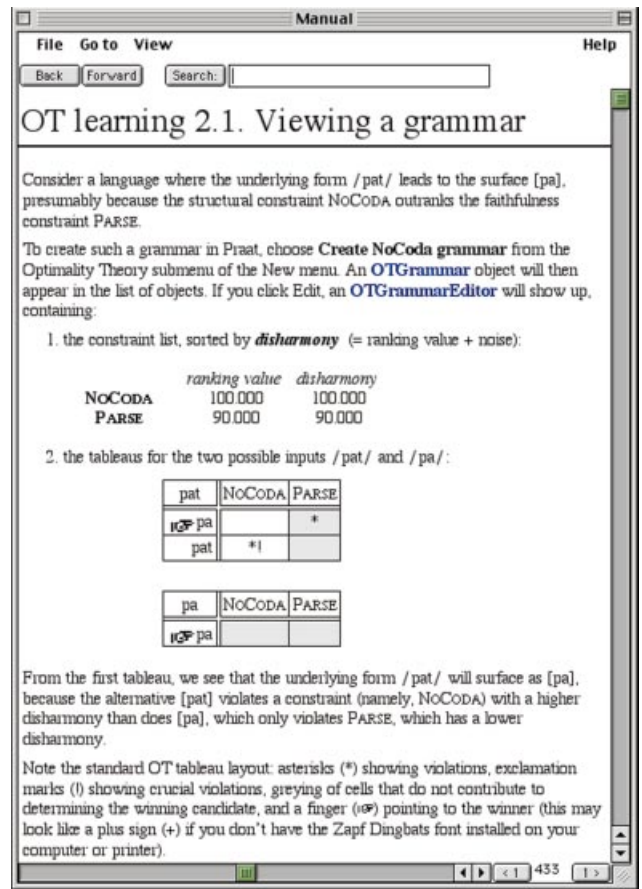


Figure 5. Praat’s manual window.

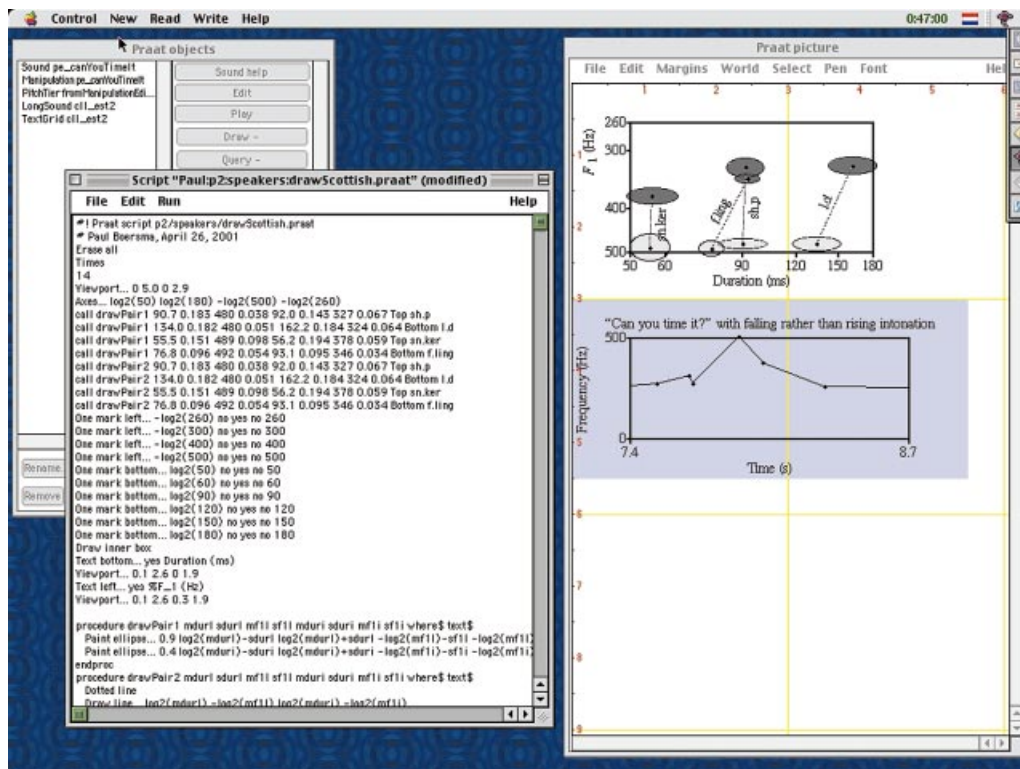


Figure 4. Praat’s script and picture windows.

8. The PRAAT manual

PRAAT comes with an extensive tutorial, which you can start by choosing "Introduction to PRAAT" from the "Help" menu. The entire reference manual is contained in the program as well and consists of about 800 pages that are connected via hyperlinks (see figure 5). Help buttons are available in most windows and dialog boxes, and clicking them will take you into the part of the manual that is most appropriate in the current context.

9. Why PRAAT?

You will want to choose PRAAT for most of your phonetic research not only because it is the most complete program available (it contains *much* more than could be discussed here), or because it is distributed for free, but also because it comes with the finest algorithms. The pitch analysis algorithm is the most accurate in the world; the articulatory synthesis is the only one that can handle dynamic length changes (ejectives), non-glottal myo-elasticities (trills), and sucking effects (clicks, implosives); and the gradual learning algorithm is the only linguistically-oriented learning algorithm that can handle free variation. But of course, there will always be things related to phonetics that other programs are better at. For your convenience, PRAAT has therefore been designed to interface reasonably well with Matlab, SPSS, Excel, and the Klatt synthesizer.

10. How to get the PRAAT program

You can get the PRAAT program through its web site, www.praat.org. By writing an e-mail message to the first author, you obtain a free licence to download all current and future versions of the program, install as many copies as you like on as many computers as you like, and use the program for any legal purpose at your work, at home, and in the field. You will also be informed about major updates of the program, which appear approximately twice a year. The source code of the PRAAT program is distributed under the General Public Licence.

A user's comments on PRAAT

By Vincent J. van Heuven

Introduction

PRAAT is probably the most comprehensive toolbox for phonetic research available worldwide, and it is certainly the most affordable; it actually costs no money at all. In fact, it is so diverse that I have never met anyone – apart from its authors – who could claim to have experience with all the modules that the program contains. I for one will have to limit the present appraisal to just those few modules that my co-workers and I have used in our laboratory. Moreover, PRAAT rejuvenates at an alarming rate. The

release that I am currently using is version 3.9.36 running on the Windows NT platform.

PRAAT started out as a collection of programs that were specifically designed to produce top-quality graphic representations of speech, i.e. oscillograms, spectra, spectrograms, fundamental frequency and intensity plots, etc. However, the flexible and well-planned structure of the program allowed its maker(s) to extend PRAAT's functionality almost indefinitely. Often, the same tasks can be done by PRAAT using different modules with different algorithms. Pitch extraction, for example, can be done with the aid of at least four different algorithms: autocorrelation, cross-correlation, SPINET, and subharmonic summation. Help files are available for each of the algorithms, explaining the meaning of the many parameter values that can be specified in for each algorithm and providing references to the literature. Each algorithm comes with a set of default parameter settings that can be overridden by the user. Also, there is an unmarked algorithm (which turns out to be an autocorrelation technique) that allows no special tuning.

In all, there would seem to very little that PRAAT cannot do for you. However, some things can be done instantaneously, other tasks can be performed only in non-obvious ways that the novice user will never discover by himself. Fortunately, the makers of PRAAT take great pride in their product, and are willing to answer queries from the floor 24 hours a day, or so it seems, again at no cost.

It should be pointed out that PRAAT is not a self-study course in experimental/instrumental phonetics. To be true, a detailed on-line technical reference manual is included with the program, but it generally does not discuss the pros and cons of alternative approaches/solutions to speech analysis problems. The user must decide on his own which algorithm will suit his purposes best. In this respect PRAAT is not unlike the magic broom that takes off with the sorcerer's apprentice. The general advice would be: do not try this at home, and always consult your local phonetician.

Multipanel editors

A recent development seems to have been toward providing smorgasbord-like complex presentations which display speech parameters as a function of time in multiple synchronized panels. Two such complex editors are provided.

1. The first is the basic waveform editor (which is invoked by a Sound object), which can be tailored to the user's taste. It allows for simultaneous display of the waveform, spectrogram, formant tracks (in red), a pitch curve (blue) and an intensity curve (yellow), all superimposed on the spectrogram. Each of the five displays can be switched on/off, scales can be adjusted for optimal visual resolution, there is a (limited) choice of algorithms that can be invoked for each display, and parameter settings can be chosen

independently for each display. Values can be eyeballed and read out under cursor control; digital readouts can be obtained through data queries. The edit functions allow cut, copy and paste, zero, and time-reverse. The parameter tracks can be extracted from each display and stored separately.

2. The second is the editor that is used for Manipulation objects. The waveform is displayed together with a pitch track (default pitch determination algorithm) and a relative duration parameter. In the waveform the moments of glottal closure are indicated by vertical blue lines. The corresponding pitch-synchronous frequency value is displayed in light gray in the pitch manipulation display. Presence/absence and location of glottal pulses can be manipulated. Also the user can stylize the pitch curve and/or change the pitch curve in any way he wants. Similarly, time intervals can be selected and given different relative durations. This allows portions of the utterance to be stretched or compressed in time. After manipulation the sound can be resynthesized using two different analysis-resynthesis schemes:

a. PSOLA resynthesis: a relatively simple waveform manipulation technique that affords the manipulation of pitch and duration but detracts very little from the original sound quality.

b. LPC resynthesis: a statistical data reduction technique that generally leads to considerable loss of sound quality but affords – in principle – the manipulation not only of prosodic parameters (pitch and duration) but also of spectral parameters (sound quality or timbre). Unfortunately, the display and manipulation (smoothing, stylization, frequency shift) of spectral parameters (formant tracks) is not implemented in the manipulation editor, nor are these functions easily available elsewhere in the package.

It should be doable, in principle, to reduce the two editors to just one generalized editor that allows the display, interactive measurement and manipulation of all the relevant properties of the speech signal. The manipulable properties should include the intensity curve. This parameter is currently displayed in the waveform editor (optionally) but cannot be manipulated.

Additional displays

Cochleagrams. Hidden further down the hierarchy of PRAAT functions are the possibilities to create auditory spectrograms (or cochleagrams). As an option with the cochleagram the loudness (expressed in Sones) of a time-slice can be queried. It is not possible, in its present state, to instruct PRAAT to produce a loudness trace as a function of time (although the user can generate and print such a contour, using the built-in programming language).

Vowel diagrams. It is also possible to plot a vowel, or even a series of vowels, as points in a vowel diagram, i.e. a two-dimensional graph plotting the

first formant frequency F1 against the second formant frequency F2. Optionally, dispersion ellipses can be drawn around the scatter clouds of vowel points in the F1-by-F2 display. Such plotting facilities are also provided by the – expensive – Kay Computerized Speech Lab (CSL) package. Using the annotation tools incorporated in PRAAT, beautiful print-quality vowel diagrams can be produced. For teaching purposes it would be attractive if this display could also be used as part of a user interface to generate vowel sounds by moving the cursor around in the display (interactively or from pre-defined custom-made trajectories), using LPC synthesis. The authors at one time promised that this facility would be made available but I have not seen it (yet). As far as I know, there is no interactive software around that can do this sort of vowel synthesis (although the Vowel Hunter program developed at the Phonetics Laboratory of Bonn University, Germany, comes close). Also, there is the talking vowel diagram provided on the Speech Production and Perception I CD-ROM issued by Sensimetrics. However, this product does not provide for on-line vowel synthesis; it just plays a fairly small number of pre-stored vowel waveforms.

Scripting language

PRAAT comes with a full programming language which can be used to create script that can be run in batch mode, allowing the user to analyze large quantities of data automatically – with or without user intervention, and to store measurements in a database for off-line statistical data analysis using such packages as SPSS. PRAAT scripts can be programmed from scratch or the user can build upon a basic script that is generated by the PRAAT macro-recorder. PRAAT keeps a log of any button pressed or keystroke entered during the interactive session. At any moment the session's history can be loaded into a text editor and used as a starting point for a program.

Using the programming tool, the user can extend PRAAT any way he likes, defining new functions and making these easily accessible in the PRAAT user interface as optional buttons. Any user with a basic grasp of computer programming will be able to construct PRAAT scripts. The PRAAT interactive manual provides lots of sample scripts to give the novice a basic feel of how to go about generating scripts.

PRAAT as a sound generator and teaching tool

For the teaching of basic acoustics – often a tough subject for undergraduate language students with a non-technical background – PRAAT provides a complex tone generator with very limited possibilities. To be true, PRAAT also allows the user to define any waveform by typing in and/or editing full formulae such as:

$$1/2 * \sin(2 * \pi * 377 * x) + \text{randomGauss}(0, 0.1),$$

which generates a 377-Hz sine wave with some white noise superimposed, but this is not an option for the beginner. It would therefore be more fun if the simple tone generator could be extended such that the user could interactively set and adjust the fundamental and the intensities of a number of harmonics in a spectral display, observe the effects of the spectral adjustments in a waveform display and listen to it, all at the same time. Conversely, it would be ideal if some pre-stored or external sound (either from a tape recording or from a live microphone) could be simultaneously displayed in an on-line fashion as a waveform and as a spectrum. Older (UNIX) versions of the GIPOS speech processing package developed at the former Institute of Perception Research at the Technical University of Eindhoven, The Netherlands,

contained such a facility as a goody, but it is no longer included with the Windows edition of GIPOS.

Especially amusing and instructive is the function for generating Shepard tone spirals. This is a complex tone signal with a pitch that seems to be continually rising, without getting anywhere.

Conclusion

In summary, PRAAT is a formidable research and teaching tool for phonetics. This report has not done justice to its makers in two respects: first, it singled out only a small part of PRAAT's many possibilities, and second, it put undue emphasis on things PRAAT cannot (yet) do. I end this review by reinstating that PRAAT is unrivalled as a general purpose speech analysis tool.