Practical Acoustics in Clinical Voice Assessment: A Praat Primer

Youri Maryn

Department of Otorhinolaryngology and Head & Neck Surgery, Sint-Augustinus General Hospital, European Institute of ORL-HNS
Antwerp, Belgium
Department of Speech-Language Pathology and Audiology, University College Ghent
Ghent, Belgium
Faculty of Medicine and Health Sciences, University of Antwerp
Antwerp, Belgium

Disclosures
Financial: Youri Maryn has a financial interest in a website-based company that sells information described in this manuscript.
Nonfinancial: Youri Maryn has no relevant nonfinancial interests to disclose.

Abstract

Praat is a freely available software package for formatting and analyzing sound signals. It carries many features relevant to voice signals and it has the possibility to run customized scripts. Praat can therefore be employed for acoustic voice analysis for clinical, educational, and research purposes; however, under the condition that voice signals are sampled for sufficient recording quality. This text first addresses recording-related issues and subsequently describes how Praat can be operated to yield numerical and graphical information regarding, for example, fundamental frequency, sound level, spectrum, spectrogram, cepstrum, and cepstrogram.

Introduction: Prerequisites for Acoustic Voice Analysis

In the realm of laryngeal pathology and voice disorders, clinicians typically evaluate and document the following facets of the audible voice signal: pitch, loudness, quality, stability, and underlying tension or muscle tonus. Furthermore, they sometimes request patients to use their voice in a more challenging way (i.e., to phonate as long, high, low, loud, or soft as possible) in order to expose the voice's capacity. All these aspects are in one way or another related/relevant to vocal status and physiology, and because they are acoustically-auditorily emitted, auditory-perceptual as well as acoustic methods are logically and routinely applied in clinical voice assessment. This paper centers on acoustic analysis of air-borne voice signals, explicitly with the computer program Praat (Paul Boersma and David Weenink, University of Amsterdam, The Netherlands).

Analyzing microphone recordings of speech and voice signals (i.e., doing acoustic voice analysis) is common practice in clinical voice management (e.g., Roy et al., 2013) and particularly attractive because of obvious reasons: quantitative method, relatively objective, relatively low cost, noninvasiveness, ease of application, and its relation to underlying voice physiology, skill, and capacity. However, prior to undertaking high-standard acoustic voice analyses, clinicians have to make well-considered choices in all of the following items. First, *room acoustics and ambient noise* can interfere with the integrity of the direct voice signal and its analyses (Howard & Murphy, 2008). This implies that clinicians are to record voice signals in as quiet as possible conditions and to strive for environments minimally contaminated by relevant reverberations, airborne noise (e.g., computer ventilation or airconditioning) and/or structure-borne noise (e.g., closing of doors or foot steps in corridor). Second, recording equipment is crucial for acquiring high-quality voice recordings. Especially *type and placement of microphone* affect the voice signal and its acoustic markers (Parsa, Jamieson, & Pretty, 2001; Titze & Winholtz, 1993; Winholtz & Titze, 1997b), and the clinician should check at least the following microphone characteristics in its manual (Švec & Granqvist, 2010): transducer type, frequency range, frequency response, directionality, and dynamic range. Preference for a specific microphone finally depends on its purpose: for example, recording only comfortable/habitual vocalization or complete phonatory range (which stipulates dynamic range requirements), and applying it for only voice assessment or complete speech assessment (which determines frequency range and response prerequisites). Third, for an eventual *microphone preamplifier,* at least the following elements should be controlled by the clinician (Švec & Granqvist, 2010): input impedance, dynamic range, frequency range and response, and powering. Fourth, the clinician should also consider the following facets of the *digital audio capturing device* (Švec & Granqvist, 2010): bit resolution (Ternström & Granqvist, 2010), maximum input level, and sampling rate (Deliyski, Shaw, & Evans, 2005b). Table 1 provides an overview with some voice research-based recommendations pertaining to sound recording environment and equipment for clinical measurement purposes. For explanation of most of the items in Table 1, the reader is referred to Howard & Murphy (2008). Fifth, for sound level or intensity measures to be valid and accurate, their *calibration* according to Winholtz and Titze (1997a) or Maryn and Zarowski (2015) is indispensable and should be accomplished prior to recording and analyzing voice signals.

Table 1. Parameters in High-Quality Configuration for Clinical Sound Recording.

Parameter	**Argument**	**Source**
	Room acoustics	
Reverberation time	Sound laboratorium: 0.1–0.3 s Sound recording: 0.4–1.0 s	Alton Everest, 1989; Howard & Murphy, 2008
	Ambient noise	
Noise level	≥15 dB below signal level, or ≤ 30 dB_A	Švec & Granqvist, 2010
	(For perturbation measures:) ≥42 dB below signal level (recommended), or ≥ 30 dB below signal level (acceptable)	Deliyski et al., 2005a
	Microphone characteristics	
Transducer type	Electret Condenser	Titze & Winholtz, 1993; Švec & Granqvist, 2010
Frequency range	10–20000 Hz	Švec & Granqvist, 2010
Frequency response	50–8000 Hz (flat, within 2 dB)	Švec & Granqvist, 2010
Directionality	Cardioid	Titze & Winholtz, 1993; Švec & Granqvist, 2010
Dynamic range	≈115 dB	Švec & Granqvist, 2010
	Microphone position relative to the sound source	
Wearing	Head-mounted (invariant source-to-microphone distance)	Winholtz & Titze, 1997b
Direction	Toward sound source	Švec & Granqvist, 2010
Angle	45–90°	Titze & Winholtz, 1993
Distance	4–10 cm	Titze & Winholtz, 1993; Švec & Granqvist, 2010

(continued)

	Preamplifier	
Input impedance (Z_{in})	≤ Z_{in} of microphone	Švec & Granqvist, 2010
Dynamic range	≈115 dB	Švec & Granqvist, 2010
Gain	invariant (once set to avoid clipping of loud vocalizations and hyposensitivity to soft phonations)	
Frequency range and response	≈ microphone	Švec & Granqvist, 2010
Power supply	48 V phantom power supply (in case of condenser or electret transducer)	Švec & Granqvist, 2010
	Digital audio capturing device	
Bit resolution	≥24 bits	Švec & Granqvist, 2010; Ternström & Granqvist, 2010
Sampling rate	≥26000 Hz	Deliyski et al., 2005b
	≥2 * Nyquist frequency, or 44100 Hz (standard in audio recorders)	Švec & Granqvist, 2010
	Overall signal-to-noise ratio	
Signal-to-noise ratio for perturbation measures	≥42 dB (recommended), or ≥ 30 dB (acceptable)	Deliyski et al., 2005a

What it finally comes down to, is that all relevant vocalizations and speech tokens are sampled as pure as possible (i.e., as least as possible polluted by recording-related noise). Once possible influence of these methodological concerns and technical limitations have been dealt with and thus minimized, voice clinicians can relatively safely start recording and consequently analyzing whatever phonatory sound is needed to be assessed. There are many acoustic markers in vocal signals relevant to clinical assessment of voice and its disorders. Readers are advised to consult the tabulations of Buder (2000) to have an overview on the numerous acoustic voice measures. The following paragraphs provide an overview of some of these markers and describe how to determine them with the program Praat.

Using Praat for Acoustic Voice Assessment in the Clinic

There is a multitude of commercially as well as freely available computer programs for clinical voice acoustics. Examples of commercial voice analysis programs are: Multi-Dimensional Voice Program (MDVP; Pentax Medical), Voice Range Profile (VRP; Pentax Medical), Analysis of Dysphonia in Speech and Voice (ADSV; Pentax Medical), LingWAVES (Wevosys), DiVAS (Xion), TF32 (Paul Milenkovic), Dr. Speech (Tiger DRS), and Vox Metria (CTS Informática). Examples of gratis voice analysis software (some of which can be downloaded at the internet) are: Praat, Speech Filing System (Mark Huckvale,), Audacity (The Audacity Team), and Sky Application (Norma Antonanzas-Barroso).

Because of its combination of the following advantages, however, the program Praat (literally "Speak" in Dutch) is especially appealing for clinical voice analysis purposes. First, it can be downloaded for free (www.praat.org). Second, it offers packages for the most popular and common computer operating systems (i.e., Windows, Macintosh, Linux, and so forth) and can therefore be applied regardless operating platforms used by the voice and speech clinician. Third, the operator controls many of the analyses' parameters, and therefore isn't totally dependent on unmodifiable settings as pre-defined by the manufacturer. This increases the clinician's freedom and control, and offers the possibility to experiment with parameter arguments. Fourth, many vocologically relevant markers (i.e., acoustic voice measures related to fundamental frequency, sound level, formant, perturbation, spectral configuration, cepstral configuration, etc.) are readily available in Praat. Fifth, clinicians with questions and doubts can consult Praat's extensive manual and help function as well as its online discussion forum (https://uk.groups.yahoo.com/neo/groups/praatusers/conversations/topics/5431). Sixth, Praat can be applied to edit (i.e., to select, extract, zoom in and out, reverse, modify intensity and/or fundamental frequency, resample, annotate, filter, concatenate, etc.) voice signals and thus prepare them for whatever clinical or scientific purpose needed. Seventh, it can be employed to draw all kinds of graphs (e.g., oscillogram, spectrum, spectrogram, cepstrum, and cepstrogram) and statistics relevant to voice assessment and research. Eighth, there are numerous Praat scripts for automated analysis of the acoustic voice signal, providing single-button prompts to sometimes particularly complex procedures that otherwise would be far too time- and labor-consuming to apply in the clinic. Such macros can be found across the internet, or can be written and customized by the Praat user himself or herself. Furthermore, applying a script increases Praat's user-friendliness and induces standardization and consistency in analysis methods.

Praat also comes with some disadvantages or caveats. First, there is no service contract and when confronted with a problem, the user has to find a solution on his or her own or has to consult someone else (i.e., forum discussant, scriptwriter, etc.) for support without contractual obligations. Second, Praat's code and commands are updated regularly, which may cause a script to get jammed when not upgraded for the latest Praat versions. Third, unlike other programs, Praat cannot provide information in real time (i.e., during recording and while phonating). It therefore cannot be used as a biofeedback tool (e.g., in voice therapy). Fourth, it has a less intuitive graphical user interface compared to other programs with more familiar layouts, styles, and operation dynamics. However, once acquainted with its visual style and different functionalities, voice clinicians can rapidly become proficient in using Praat for analyzing voice signals, especially when scripts have been implemented (e.g., following Maryn, 2016). This text aims at lowering the threshold to apply Praat. The following paragraphs are meant as a brief introduction to working with Praat, with only some of many relevant elements described and explained in tables for practical reasons and brevity. For Praat's extensive manual, readers are referred to its website or help function.

Some General Functions in Praat

When Praat is launched, two windows automatically appear: Praat Objects and Praat Picture (see Figure 1). There are generally two strategies to use in Praat for analyzing voice signals: with pre-set buttons and functions or with self-entered buttons initializing scripts. Before exploring both ways, however, it is important to record and save a new voice sound or open an earlier recorded voice sound. Such general functions occur in the Praat Objects window. Table 2 provides a functional overview of some of Praat's general functions. The Praat Picture window is where graphs and statistics are drawn (e.g. for post-hoc printing or saving in PDF format).

Figure 1. The Two Windows That Appear When Praat is Opened: Praat Objects and Praat Picture.

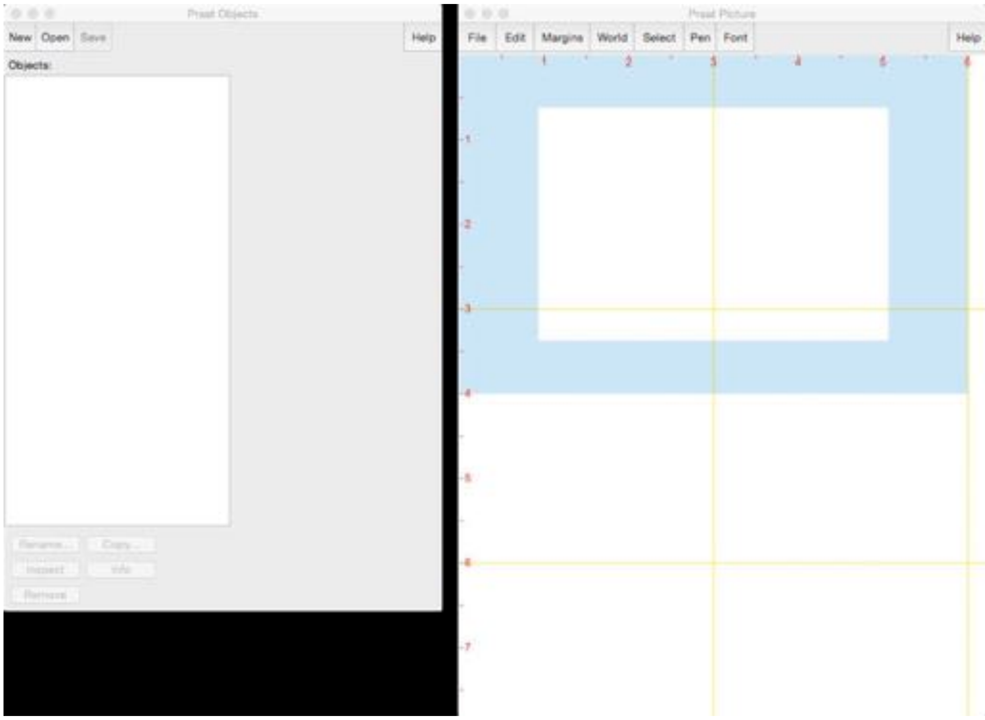


Table 2. Functional Overview of Some General Functions in the Praat Objects Window.

Function	**Window**	**Operation***
To record a sound	Praat Objects	New > Record mono Sound…
	SoundRecorder	Choose number of Channels, Input source and Sampling frequency
	SoundRecorder	Record > Stop
	SoundRecorder	Save to list (transfers the signal to Objects: list), or Save to list & Close (transfers the signal to the Objects: list in Praat Objects and shuts SoundRecorder)
To save a recording	Praat Objects	Save > Save as WAV file…
To open an existing recording	Praat Objects	Open > Read from file…
To scan a recording	Praat Objects	View & Edit
To rename or copy an object	Praat Objects	Rename… or Copy… (and type a new name in the line of small window) > OK
To inspect an object	Praat Objects	Info

**The symbol '>' means 'next operation'.*

Some Pre-Set Acoustic Voice Markers in Praat

Praat offers by default the following analyses of the acoustic voice signal: (a) fundamental frequency (i.e., f_o) as estimate of vocal pitch; (b) sound intensity as estimate of vocal loudness; (c) jitter, shimmer, harmonic-to-noise ratio, and smoothed cepstral peak prominence as estimates of vocal sound quality; (d) number and degree of voice breaks as estimates of voice break intervals; and (e) central frequency and bandwidth of the first four formants (i.e., F_1-F_4 and B_{F1}-B_{F4}, respectively) as estimates of articulation physiology.[1] There are two strategies to reach these data. First, the clinician can choose an *analysis in Praat's Object window* (e.g., with To Pitch…, and than query the resulting object [in this example it will be a Pitch object]). For f_o and its interesting statistics, this is illustrated in Table 3, provided that the clinician has clicked on a sound object in the list. Similar routines (i.e., query following analysis) can be followed for other acoustic markers. Depending on estimate, speech task, and purpose, arguments can be adjusted.[2] Second, the clinician can initiate *Praat's sound viewer and editor* with the "View & Edit" function. This function opens the Praat Editor window to graphically show analyses and to offer various relevant queries. Figure 2 displays an example of a sound signal with all analyses activated. Table 4 provides an overview of relevant functions in the Praat Editor window.

[1]Symbolic notation style of frequencies, harmonics and formants are in accordance with the consensus report of Titze et al. (2015).

[2]For example, for the calculation of the Smoothed Cepstral Peak Prominence according to Maryn & Weenink (2015), the arguments of the following parameters should be changed: Subtract tilt before smoothing = no, Time averaging window=0.01 s, Quefrency averaging window=0.001 s, and Line type = Straight.

Table 3. Functional Overview of Relevant f_o *Markers in the Praat Objects Window When a Sound Object is Selected in the List.*

Marker	**Window**	**Operation***
	Praat Objects	Analyze Periodicity > To Pitch…
	Sound: To Pitch	Keep the default arguments, or adjust them according to phonatory task (e.g., phonating at highest possible pitch) > OK (a Pitch object has been added in the Objects list)
Mean f_o	Praat Objects	Query > Get mean…
	Pitch: Get mean	Keep or adjust arguments > OK (output appears in Praat info window)
Median f_o[a]	Praat Objects	Query > Get quantile…
	Pitch: Get quantile	Keep or adjust arguments > OK (output appears in Praat info window)
Minimum f_o	Praat Objects	Query > Get minimum…
	Pitch: Get minimum	Keep or adjust arguments > OK (output appears in Praat info window)
Maximum f_o	Praat Objects	Query > Get maximum…
	Pitch: Get maximum	Keep or adjust arguments > OK (output appears in Praat info window)
Standard deviation of f_o	Praat Objects	Query > Get standard deviation…
	Pitch: Get standard deviation	Keep or adjust arguments > OK (output appears in Praat info window)

**The symbol '>' means 'next operation'.*

[a] Median is the preferred estimate, because it is not influenced by extreme or outlying data in contrast to mean/average.

Figure 2. Illustration of the Praat Editor Window. In the upper part it shows the sound signal of somebody saying "Perspectives in voice and voice disorders," with vertical blue lines superposed to indicate vocal cycle/pulse boundaries and thus voiced segments. In the lower part of the window it shows the time-aligned spectrogram (grey, 0–4000 Hz), f_o *contour (blue, 50–500 Hz), intensity contour (yellow, 30–100 Hz), and contours of* F_1, F_2, F_3 *and* F_4 *(red, speckled, 0–4000 Hz). With the cursor randomly clicked at 2.670199 s and 754.9 Hz in the spectrogram, this window immediately shows that* f_o*=119.3 Hz and intensity=65.37 dB at this time.*

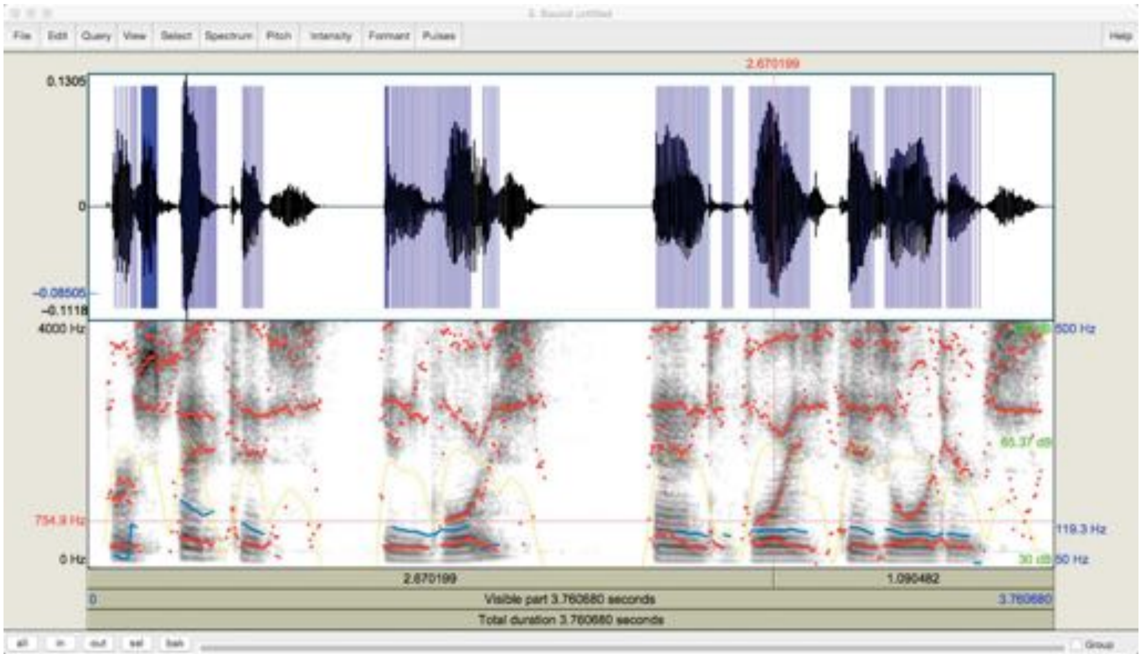


Table 4. Functional Overview of Relevant Markers in the Praat Editor Window After the View & Edit Button Was Hit for a Sound Object in the Praat Objects Window. The output counts for either the interval that has been selected (e.g., by scrolling with the mouse) or only the place of the cursor.

Marker	Window	Operation*
	Praat Editor	Spectrum > Show spectrogram (a spectrogram is shown in the lower part)[a]
	Praat Editor	Pitch > Show pitch (a blue pitch contour superposes the spectrogram)
Mean f_o	Praat Editor	Pitch > Get pitch
Minimum f_o	Praat Editor	Pitch > Get minimum pitch
Maximum f_o	Praat Editor	Pitch > Get maximum pitch
	Praat Editor	Intensity > Show intensity (a yellow intensity contour superposes the spectrogram)
Mean intensity	Praat Editor	Intensity > Get intensity
Minimum intensity	Praat Editor	Intensity > Get minimum intensity
Maximum intensity	Praat Editor	Intensity > Get maximum intensity
	Praat Editor	Formant > Show formants (red speckled formant contours superpose the spectrogram)
F_1	Praat Editor	Formant > Get first formant
B_{F1}	Praat Editor	Formant > Get first bandwidth
F_2	Praat Editor	Formant > Get second formant
B_{F2}	Praat Editor	Formant > Get second bandwidth
F_3	Praat Editor	Formant > Get third formant
B_{F3}	Praat Editor	Formant > Get third bandwidth
F_n	Praat Editor	Formant > Get formant…
B_{Fn}	Praat Editor	Formant > Get bandwidth…
	Praat Editor	Pulses > Show pulses (blue vertical stripes superpose the oscillogram in the upper pane)

(continued)

Mean f_o, median f_o, standard deviation of f_o, minimum f_o, maximum f_o, number of voice breaks, degree of voice breaks, various jitters[b], various shimmers[c], and harmonics-to-noise ratio	Praat Editor	Pulses > Voice report

**The symbol '>' means 'next operation'.*

[a]Via Spectrum > Spectrogram setting… the type of spectrogram can be determined: Window length=0.03 s gives a narrowband spectrogram, Window length=0.005 s results in a broadband spectrogram. A narrowband spectrogram is favored for voice analysis.[b]Praat's Voice report offers five shimmer measures: Jitter (local; a.k.a. percent jitter), Jitter (local absolute; a.k.a. absolute jitter), Jitter (rap; a.k.a. three-point relative average perturbation), Jitter (ppq5; a.k.a. 5-point pitch perturbation quotient), and Jitter (ddp).[c]Praat's Voice report provides five jitter measures: Shimmer (local; a.k.a. percent shimmer), Shimmer (local, dB; a.k.a. shimmer in dB), Shimmer (apq3; a.k.a. 3-point amplitude perturbation quotient), Shimmer (apq5; a.k.a. 5-point amplitude perturbation quotient), Shimmer (apq11; a.k.a. 11-point amplitude perturbation quotient), and Shimmer (dda).

Complex Voice Analyses With Customized Praat Scripts

Above-mentioned simple procedures already illustrate Praat's applicability as an acoustic voice analysis program. However, often clinical or scientific voice evaluations require much more complex methods due to number of operations, combination of acoustic markers, underlying calculations, iterative/looping analyses, conditional statements, graphical outputs, or other factors. Phonetography, for example, involves (a) extraction of voiced segments in often relatively long signals, (b) determination of f_o and calibrated sound level as covariants per short time window, and finally (c) drawing these covariants as coordinates in a f_o-intensity-plot. It would be very time- and labor-demanding to do this by hand. With a so-called script (i.e., a macro or an executable text consisting of menu and action commands), on the other hand, all these operations can run automatically in Praat by hitting only one button. For example, Maryn & Weenink (2015) have published a Praat script to determine the Acoustic Voice Quality Index (AVQI; a measure related to dysphonia severity)—originally from Maryn, Corthals, Van Cauwenberge, Roy, and De Bodt (2010)—with a single mouse click after appropriate recordings are made (see Figure 3). Similarly, Maryn, Morsomme, and De Bodt (in press) developed a procedure to automate determination of the Dysphonia Severity Index (DSI; a measure related to vocal function)—originally from Wuyts et al. (2000)—in Praat (see Figure 4). The author of the present text also developed Praat scripts to obtain various relevant values and graphs within a complete voice assessment: f_o (see Figure 5), vocal intensity (see Figure 6), spectrographic analysis with time-domain voice markers (see Figure 7), cepstrographic analysis with quefrency-domain voice markers (including the Smoothed Cepstral Peak Prominence or CPPS; see Figure 8), phonatory range estimation (see Figure 9), and vocal tremor modulation (see Figure 10). Having coupled the scripts behind these figures to a button in the dynamic menu of the Praat Objects window, they can easily be prompted within an efficient voice assessment procedure.

Figure 3. Output of the Praat Script for Determination of the AVQI (Version 02.02) According to Maryn & Weenink (2015).

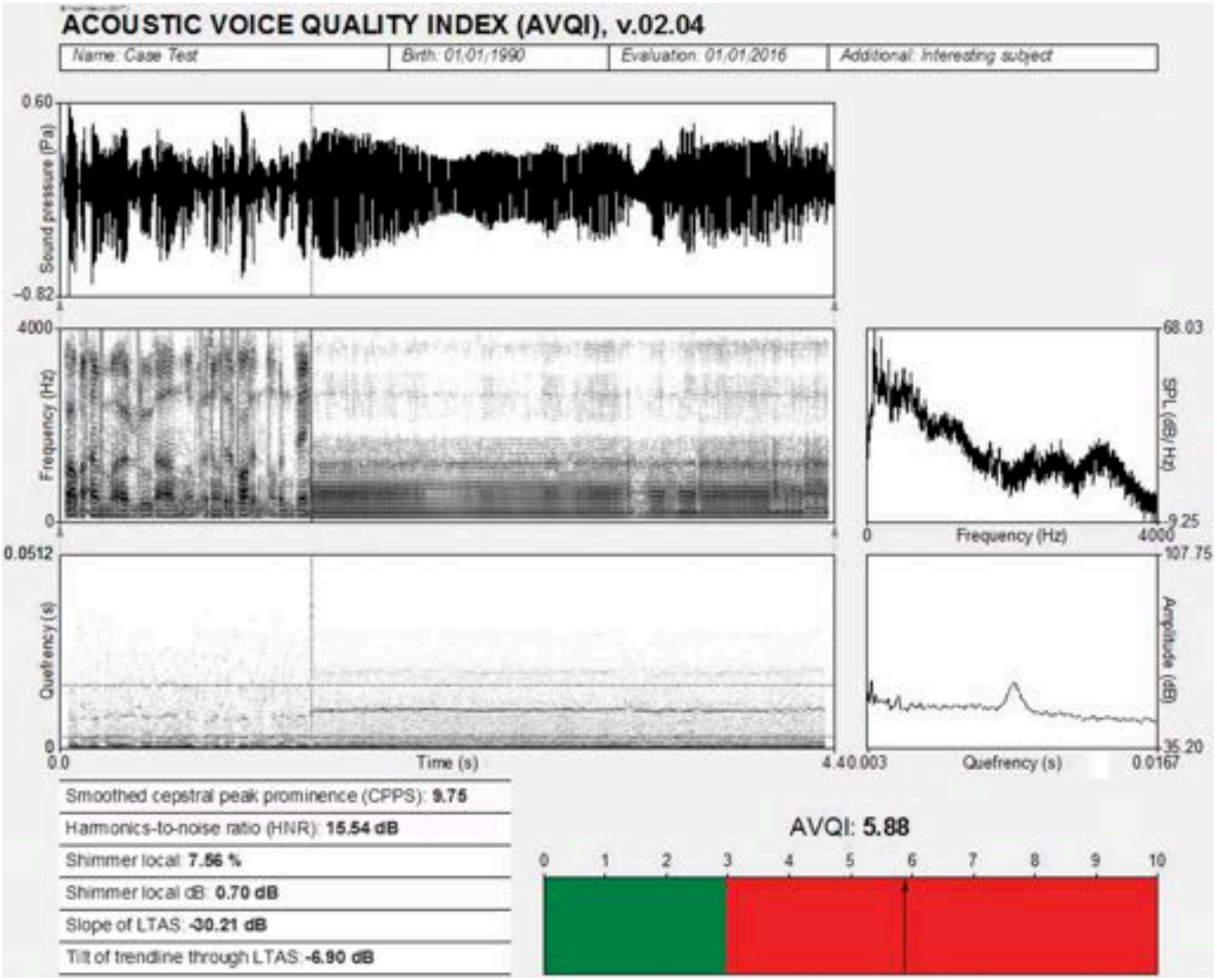


Figure 4. Output of the Praat Script for Determination of the DSI (Version 02.01) According to Maryn et al. (in press).

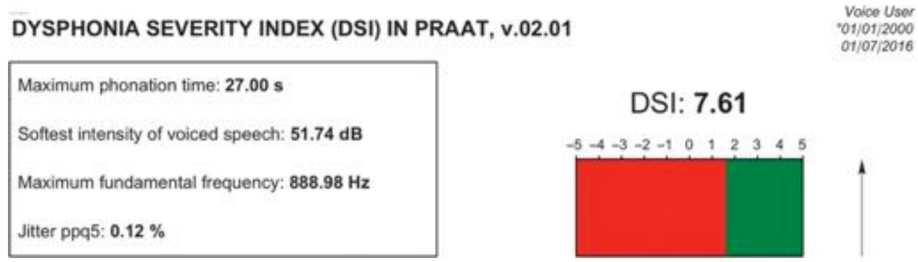
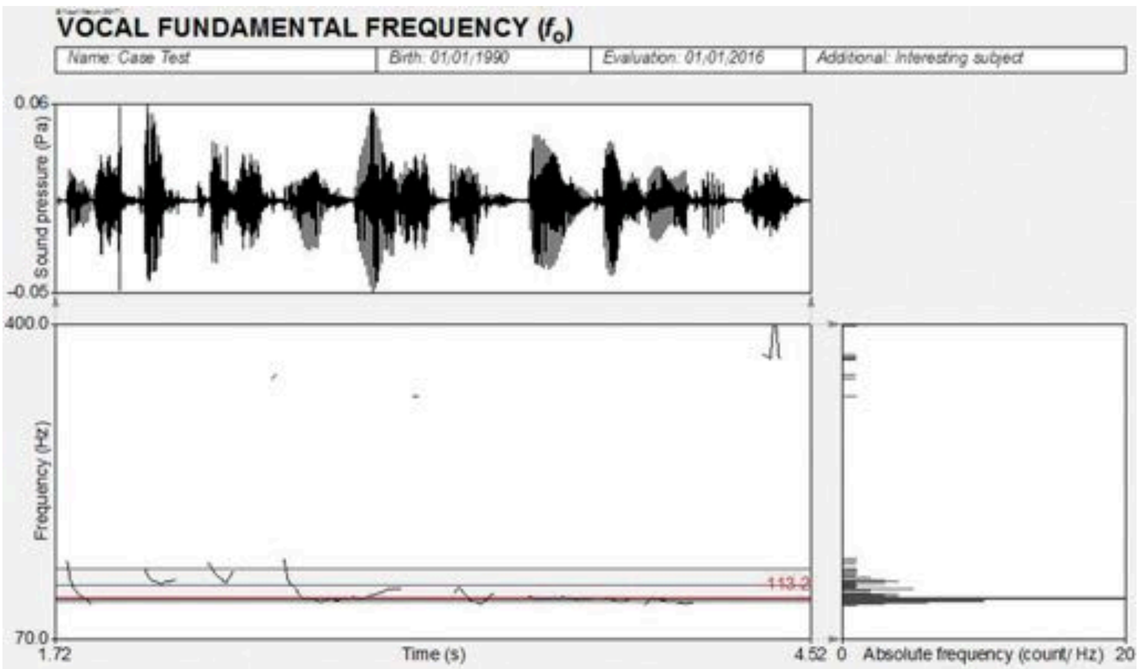
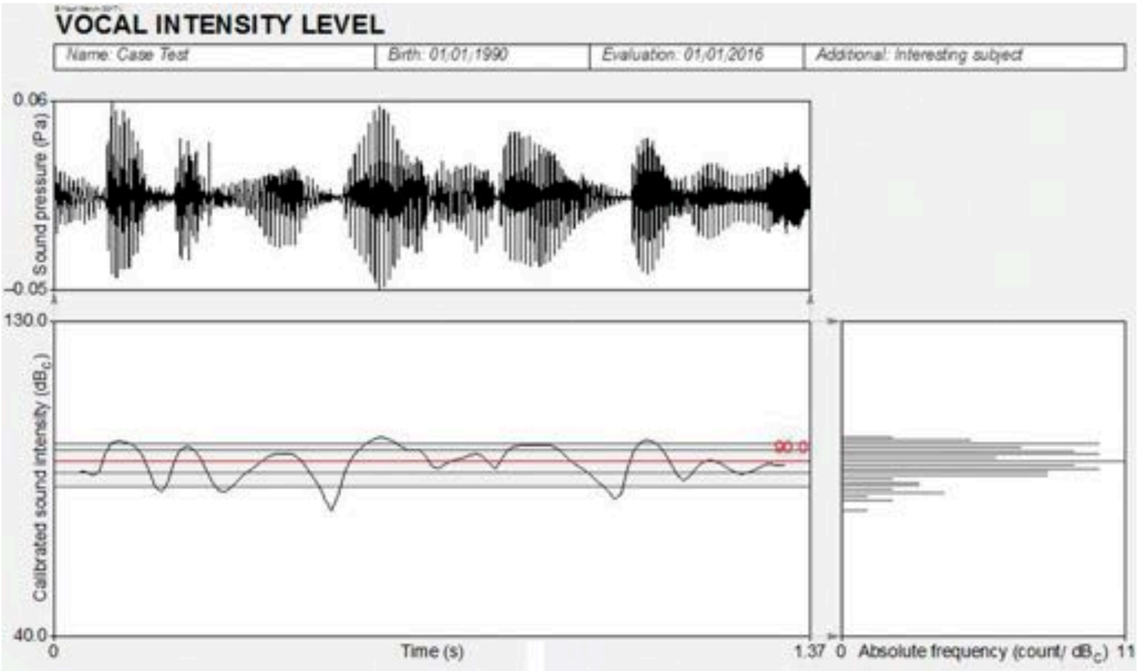


Figure 5. Output of the Praat Script for Analysis of f_o*, for a Similar Sample as in Figure 2.*



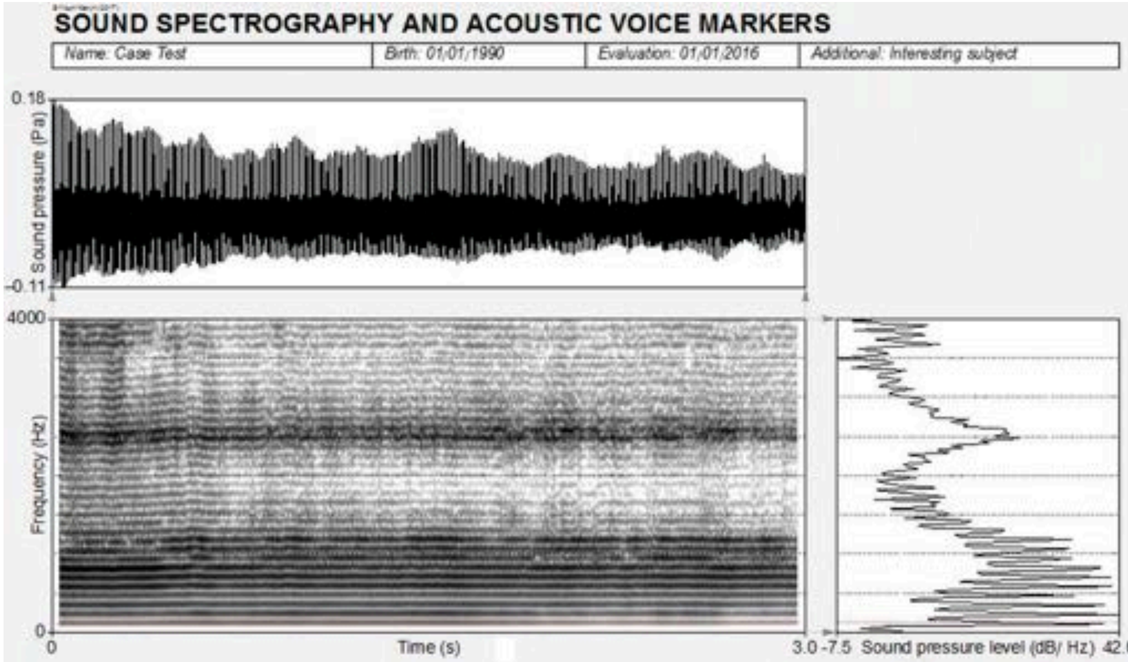
Measures of vocal f_o	in Hertz (Hz)	in semitones (relative to 1 Hz)
Minimum	106.32	80.79
10th percentile (P10)	109.33	81.27
25th percentile (P25) - 1st quartile (Q1)	111.53	81.62
50th percentile (P50) - 2nd quartile (Q2) - median	113.18	81.87
Mean	134.75	83.84
75th percentile (P75) - 3rd quartile (Q3)	126.91	83.85
90th percentile (P90)	143.64	86.00
Maximum	399.84	103.72
Standard deviation	**63.50**	**5.18**
Interquartile range (IQR) - Q3-Q1	**15.38**	**2.24**
Range between 10th and 90th percentiles - P90-P10	**34.31**	**4.73**

Figure 6. Output of the Praat Script for Analysis of Intensity, Only for the Voiced Segments of the Utterance From Figure 2.



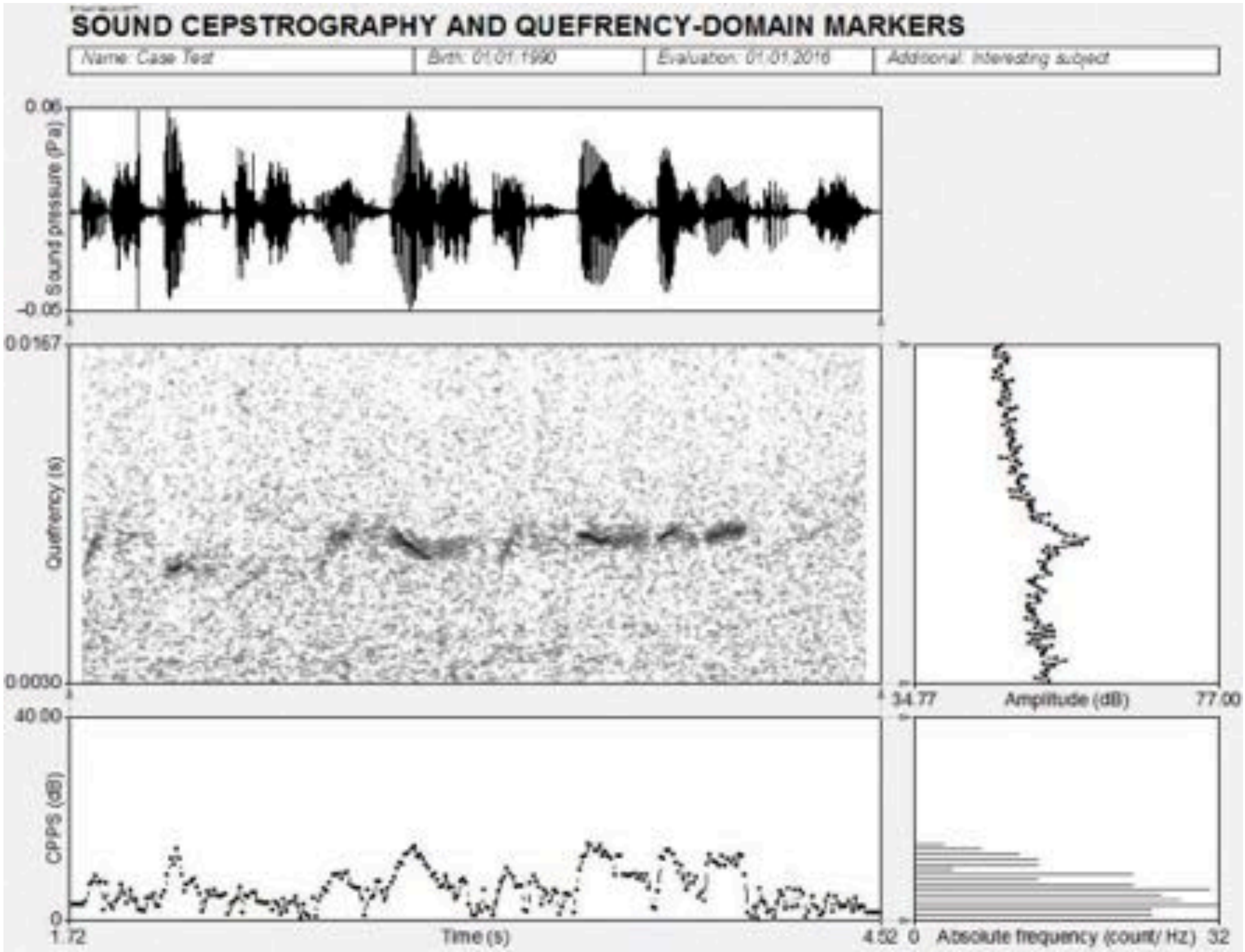
Measures of vocal intensity level (only voiced segments of speech signal)	in calibrated decibel (dB_C)
Minimum	75.79
10th percentile (P10)	82.80
25th percentile (P25) - 1st quartile (Q1)	86.85
50th percentile (P50) - 2nd quartile (Q2) - median	**89.98**
Mean	91.42
75th percentile (P75) - 3rd quartile (Q3)	93.38
90th percentile (P90)	95.02
Maximum	96.80
Standard deviation	**4.53**
Interquartile range (IQR) - Q3-Q1	**6.52**
Range between 10th and 90th percentiles - P90-P10	**12.22**

Figure 7. Output of the Praat Script for Spectrographic Analysis, for a Sustained Vowel [a:].



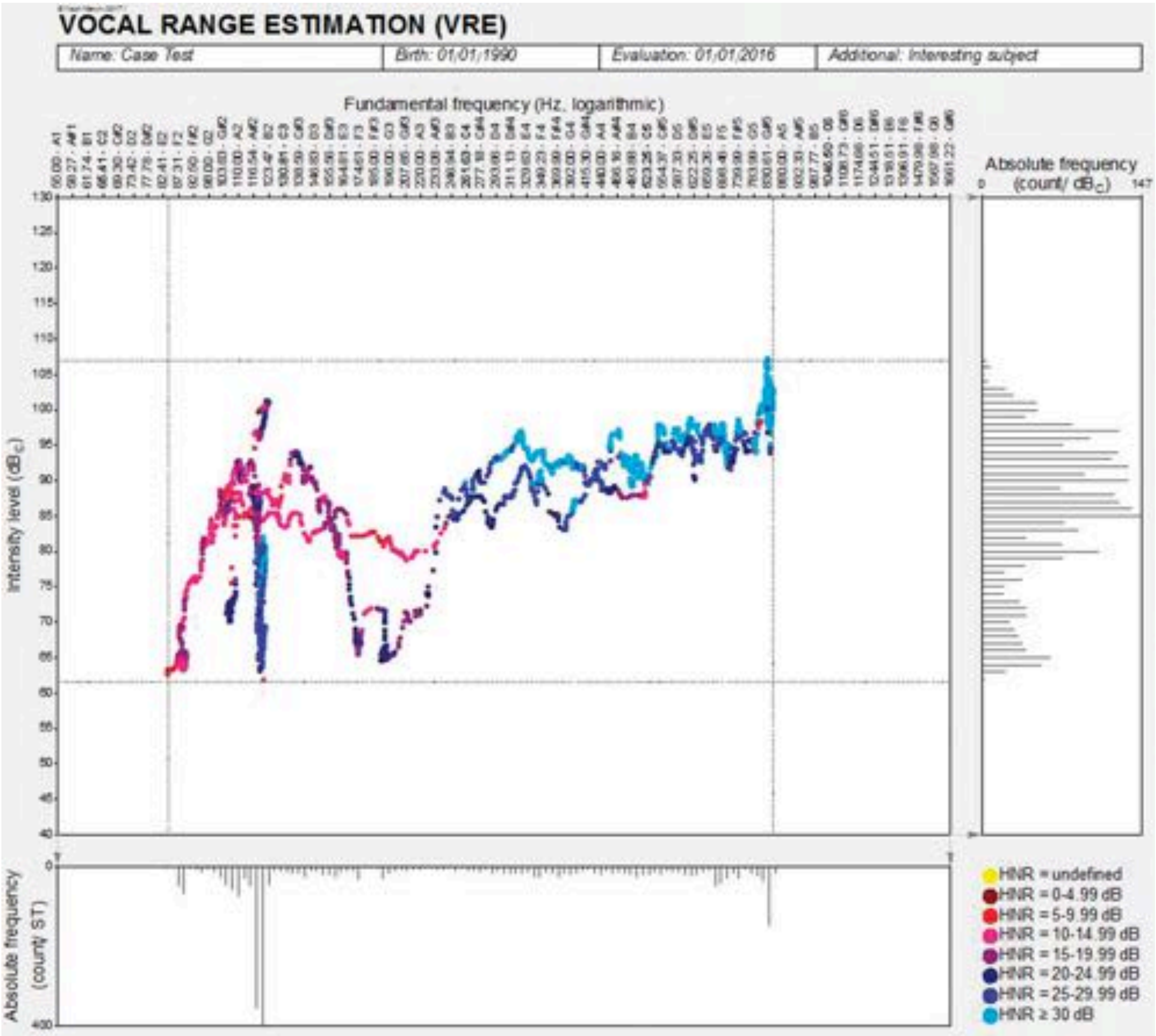
Fundamental frequency (f_o)		**Amplitude perturbation (shimmer)**	
Median f_o (Hz)	117.864	Shimmer local (%)	3.185
Calibrated sound level (SPL)		Shimmer local dB (dB)	0.289
Median SPL (dB_C)	99.482	Shimmer apq3 (%)	1.628
Fundamental frequency perturbation (jitter)		Shimmer apq5 (%)	1.811
Jitter local (%)	0.272	Shimmer apq11 (%)	2.846
Jitter local absolute (s)	0.00002	**Voicing and stability**	
Jitter rap (%)	0.155	Fraction of voiced frames (%)	0
Jitter ppq5 (%)	0.154	Number of voice breaks	0
Harmonicity of voiced parts		Degree of voice breaks (%)	0
Noise-to-harmonics ratio	0.020		
Harmonics-to-noise ratio (dB)	18.092		

Figure 8. Output of the Praat Script for Cepstrographic Analysis (i.e., Cepstral Information as Well as Smoothed Cepstral Peak Prominence in Time), for the Same Sample as in Figure 2.



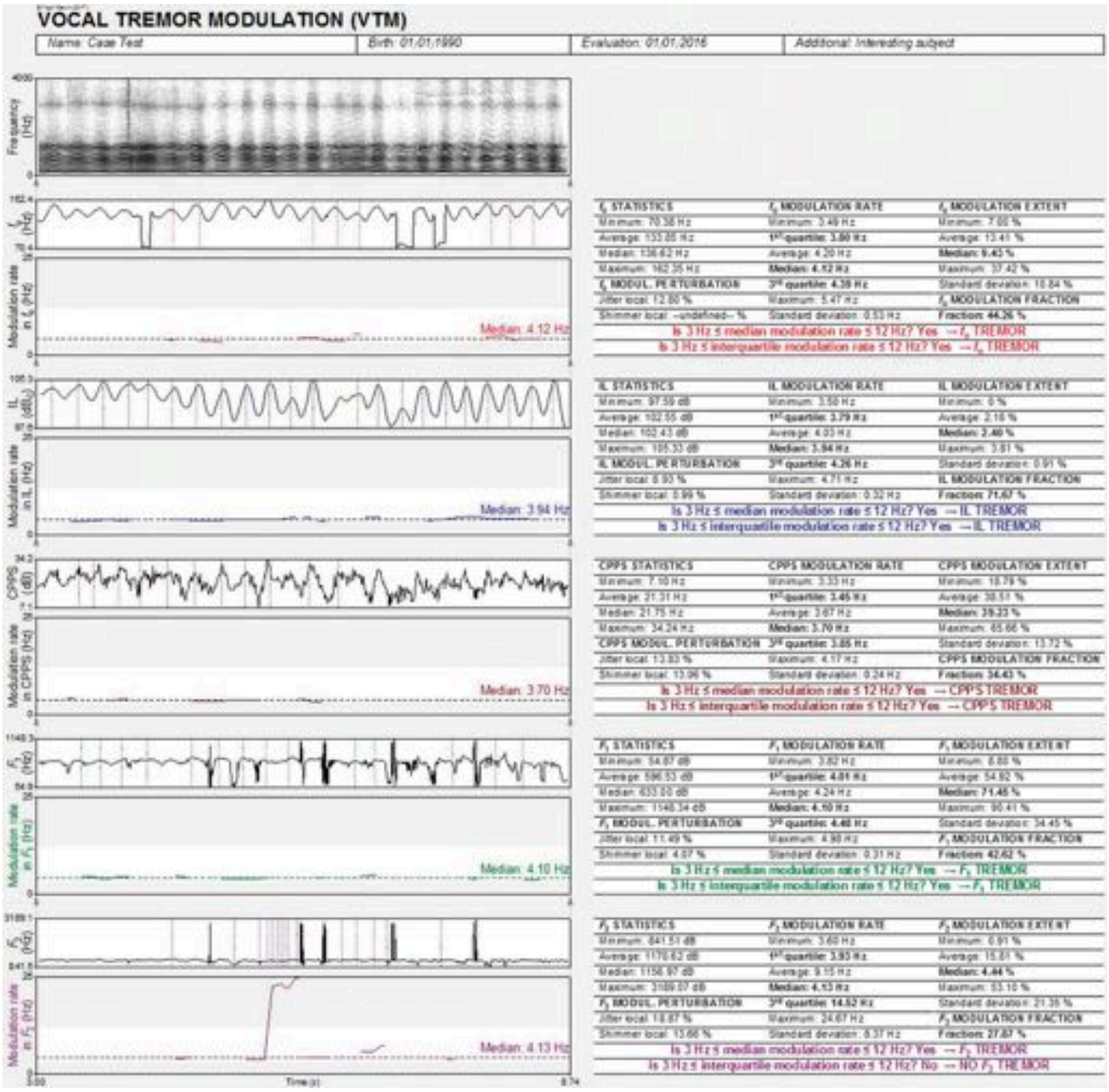
Measures of Smoothed Cepstral Peak Prominence (CPPS)	in decibel (dB)
Minimum	0.34
10th percentile (P10)	1.54
25th percentile (P25) - 1st quartile (Q1)	2.98
Overall CPPS	6.58
50th percentile (P50) - 2nd quartile (Q2) - Median	5.59
Mean	6.05
75th percentile (P25) - 3rd quartile (Q3)	8.67
90th percentile (P90)	12.12
Maximum	15.01
Standard deviation	3.73
Interquartile range (IQR) - Q3-Q1	**5.69**
Range between 10th and 90th percentiles - P90-P10	**10.58**

Figure 9. Output of the Praat Script for Phonatory Range Estimation (i.e., f_o and Calibrated Sound Level Coordinates, With Coloration Based on Voice Quality as Expressed in Harmonics-to-Noise Ratio) From Soft, Loud, Low and High Phonations.



Vocal range estimates	f_o, in Hertz (Hz)	f_o, in semitones (st)	Intensity, in calibrated decibel (dB$_C$)
Minimum	83.805	76.668	61.645
Maximum	844.012	116.653	106.877
Range	760.207	39.986	45.232

Figure 10. Output of the Praat Script for Analysis of Rate and Perturbation of Vocal Tremor Modulations.



Discussion

High-quality sampling of voice signals is a conditio sine qua non for acoustic voice analysis methods. Table 1 offers a summary of recommendations before purchasing a recording system for clinical purposes. With recording-related noise kept to minimum, clinicians may want to yield various acoustic voice markers with the program Praat. Estimation of f_o, sound level, and time/frequency/quefrency-domain measures of voice sound quality can relatively easily be calculated and visualized with Praat's pre-set functions, as illustrated in Tables 3 or 4. Furthermore, when equipped with scripts, Praat becomes proficient in returning relevant statistics and graphics (see Figures 3–10) with minimum time- and labor-consumption of the clinician within a complete voice assessment.

The present text is not a comprehensive manual on Praat for clinical voice analysis. It merely serves as a primer, intended to lower thresholds before employing it for that purpose. When specific acoustic algorithms and methods are insufficiently clear, however, readers are recommended to consult more specialized literature (e.g., Baken & Orlikoff, 2000; Buder, 2000; Howard & Murphy, 2008; and Titze, 1995). Additional education focusing on speech and voice acoustics, audio recording technology, clinically valid voice markers and their theoretical and computational background, measures' specifications and parameters with their task-specific arguments, signal formatting and marker-specific preparation, exercises for efficient application of Praat's specific features, and other topics also may be essential. Finally, it should be stated that although Praat offers common analyses and/or graphs in the scope of modern voice analysis, there are interesting acoustic features that currently are not available, such as relative fundamental frequency (as an estimate of vocal tension or effort; e.g. Stepp, Sawin, & Eadie, 2012) or nonlinear dynamics measures (as estimates of signal dimensionality; e.g. Zhang, Jiang, Biazzo, & Jorgensen, 2005).

Conclusion

With sufficient background knowledge and recording quality, it is feasible to apply Praat for clinical assessment of acoustic voice properties, especially when employing customized scripts.

References

Alton Everest, F. (1989). *The master handbook of acoustics* (2nd ed.). Blue Ridge Summit, PA: Tab Books.

Baken, R., & Orlikoff, R. (2000). *Clinical measurement of speech and voice* (2nd ed.). San Diego, CA: Singular Thomson Learning.

Buder, E. H. (2000). Acoustic analysis of voice quality: A tabulation of algorithms 1902–1990. In R. D. Kent & M. J. Ball (Eds.), *Voice quality measurement* (pp. 119–244). San Diego, CA: Singular Publishing Group.

Deliyski, D. D., Shaw, H. S., & Evans, M. K. (2005a). Adverse effects of environmental noise on acoustic voice quality measurements. *Journal of Voice, 19*(1), 15–28.

Deliyski, D. D., Shaw, H. S., & Evans, M. K. (2005b). Influence of sampling rate on accuracy and reliability of acoustic voice analysis. *Logopedics, Phoniatrics, Vocology, 30*(2), 55–62.

Howard, D. M., & Murphy, D. T. (2008). *Voice science, acoustics and recording*. San Diego, CA: Plural Publishing Inc.

Maryn, Y. (2016). *Clinical voice acoustics with the program Praat [Intensive hands-on course]*. Zagreb, Croatia: SUVAG Policlinic.

Maryn, Y., Corthals, P., Van Cauwenberge, P., Roy, N., & De Bodt, M. (2010). Toward improved ecological validity in the acoustic measurement of overall voice quality: Combining continuous speech and sustained vowels. *Journal of Voice, 24,* 540–555.

Maryn, Y., Morsomme, D., & De Bodt, M. (in press). Measuring the Dysphonia Severity Index (DSI) in the program Praat. *Journal of Voice*.

Maryn, Y., & Weenink, D. (2015). Objective dysphonia measures in the program Praat: Smoothed cepstral peak prominence and acoustic voice quality index. *Journal of Voice, 29,* 35–43.

Maryn, Y., & Zarowski, A. (2015). Calibration of clinical audio recording and analysis systems for sound intensity measurement. *American Journal of Speech-Language Pathology, 24,* 608–618.

Parsa, V., Jamieson, D. G., & Pretty, B. R. (2001). Effects of microphone type on acoustic measures of voice. *Journal of Voice, 15,* 331–343.

Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M. P., Mehta, D., Paul, D., & Hillman, R. (2013). Evidence-based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology, 22,* 212–226.

Stepp, C., Sawin, D., & Eadie, T. (2012). The relationship between perception of vocal effort and relative fundamental frequency during voicing offset and onset. *Journal of Speech, Language, and Hearing Research, 55,* 1887–1896.

Švec, J. G., & Granqvist, S. (2010). Guidelines for selecting microphones for human voice production research. *American Journal of Speech-Language Pathology, 19,* 356–368.

Ternström, S., & Granqvist, S. (2010). Personal computers in the voice laboratory: Part two–audio devices. *Logopedics Phoniatrics Vocology, 35,* 98–102.

Titze, I. (1995). *Workshop on acoustic voice analysis: summary statement.* Iowa City, IA: National Center for Voice and Speech.

Titze, I., Baken, R., Bozeman, K., Granqvist, S., Henrich, N., Herbst, C., . . . Wolfe, J. (2015). Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization. *Journal of the Acoustical Society of America, 137,* 3005–3007.

Titze, I. R., & Winholtz, W. S. (1993). Effect of microphone type and placement on voice perturbation measurements. *Journal of Speech and Hearing Research, 36,* 1177–1190.

Winholtz, W. S., & Titze, I. R. (1997a). Conversion of a head-mounted microphone signal into calibrated SPL units. *Journal of Voice, 11,* 417–421.

Winholtz, W. S., & Titze, I. R. (1997b). Miniature head-mounted microphone for voice perturbation analysis. *Journal of Speech, Language, and Hearing Research, 40,* 894–899.

Wuyts, F. L., De Bodt, M. S., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., . . . Van de Heyning, P. H. (2000). The Dysphonia Severity Index: An objective measure of vocal quality based on a multiparameter approach. *Journal of Speech, Language, and Hearing Research, 43,* 769–809.

Zhang, Y., Jiang, J. J., Biazzo, L., & Jorgensen, M. (2005). Perturbation and nonlinear dynamic analyses of voices from patients with unilateral laryngeal paralysis. *Journal of Voice, 19,* 519–528.

History:
Received August 07, 2016
Revised September 10, 2016
Accepted September 16, 2016
https://doi.org/10.1044/persp2.SIG3.14