

Stress, Affective State and Voice

Analysis of Speaking Behavior and Voice Sound Characteristics

User Manual (Android)



Why You Should Have a Closer Look at it

Lifestyle: Getting to Know Yourself Better

- 1) Get involved and learn more about the complex interplay between body and mind under high levels of stress caused, for example, by modern work conditions or by tight schedules and frequent examinations in college/university.
- 2) Learn more about your affects, those evolutionarily very old processes deep inside your body. Affects such as aggression, fear, anger, happiness, sadness, grief, or love, can be triggered through internal and external events, and are often accompanied by bodily reactions, such as sweat, rapidly increased blood pressure, increased heart rate, dizziness, and others.
- 3) Learn how regular voice recordings can improve your coping behavior under stress, so that you feel more comfortable, for example, when speaking in public – so that mean vocal pitch doesn't go up by a halftone anymore.

Subjects at Risk of Mood Disturbances

- 4) Get involved and learn how low mood reduces the dynamic expressiveness of your voice and how your voice sounds when it regains energy, vivacity and spiritedness.
- 5) Find out how helpful it is to go for a 20-minute walk, to ride a bicycle, or to do sports.
- 6) And best of all: getting involved on a routine basis is a great remedy to better cope with the ups and downs of real life.

Philanthropy

You may want to upload your voice recordings to our server in a strictly anonymous way and to share your data with the research community, thus helping to further improve the «voxApp» application.

Table of Contents

Scientific background	4
Principal operation	5
Initial setup	6
Modifying setup parameters	7
Voice recordings: Overview	8
Voice recordings: System initialization	9
Voice recordings: System test	10
Voice recordings: Counting out loud from 1-30	11
Voice recordings: Reading out loud standard text	12
Analysis: Custom versus automatic	13
Analysis: Select recording for analysis	14
Analysis: Inspecting results	15
Analysis: Inspecting distribution curves and spectra	17
Results: Interpretation of distribution plots	18
Results: Typical examples of distribution plots	19
Results: Interpretation of spectrum plots	20
How to delete recordings	21
Longitudinal analysis: Time development	22
Longitudinal analysis: Inspecting results	24
Longitudinal Analysis: Time course plots	25
Longitudinal analysis: Interpretation of time course plots	26
Longitudinal analysis: Typical examples of time course plots	27
Standard text to be read out aloud (English)	28
Texte standard à lire à voix haute (Français)	29
Standardtext zum lauten Vorlesen (Deutsch)	30
Testo standard da leggere ad alta voce (Italiano)	31
Texto estándar para leer en voz alta (Español)	32
References	33
External microphones	34

Scientific Background

Chronic stress can lead to serious health problems and can affect nearly every system of the human body, as suggested by physical, cognitive, affective and behavioral symptoms. Indeed, for a certain percentage of the general population, chronic stress raises blood pressure, increases the risk of heart attack and stroke, suppresses the immune system, and increases the vulnerability to psychiatric disorders. Health surveys indicate that the stress-induced burden is closely related to a pronounced lack of coping skills which obviously can let things escalating on the long run.

Speech is greatly influenced by the speakers' affective state which reflects feelings like sadness, happiness, grief, guilt, fear, anger, aggression, faintheartedness, shame, love, stress, or doziness – and, occasionally, depressive or psychotic symptoms. For example, hectic and abrupt, or delayed and monotonous speaking behavior can be indicators of stress-related or affective problems if such behavior persists over a longer time period.

Voice analysis has its focus on speaking behavior and voice sound characteristics. Speaking behavior relates to characteristics like “speech flow”, “loudness”, and “intonation”, while voice sound characteristics relate to the distribution and intensity of “overtones” that make up a speaker’s individual vocal “timbre” (e.g., bright or dark color).

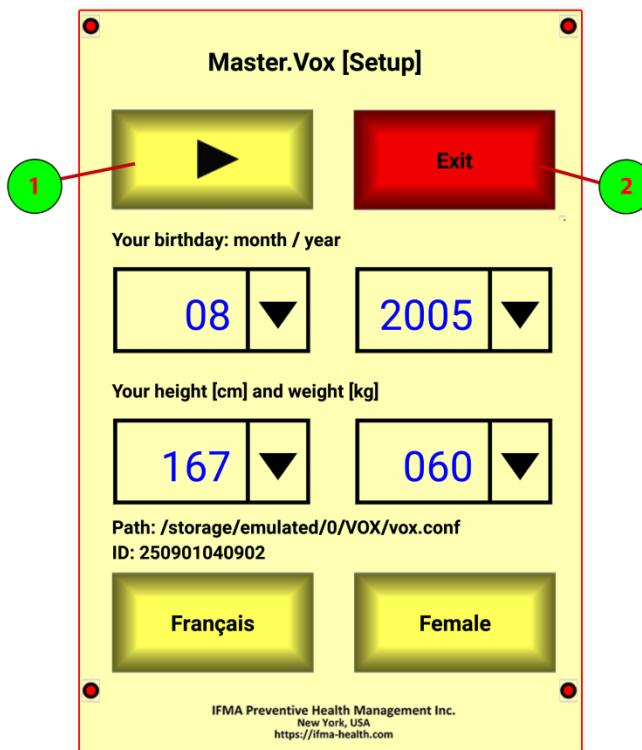
Based on normative data from our previous speech studies we have developed the «voxApp» application program for the analysis of speaking behavior and voice sound characteristics in a language-independent way and as a function of gender and age. It is available for Laptops, Tablets, and Smartphones under Windows and Android (iPhone version in preparation).

Unlike methods that yield abstract classifiers with unknown (“hidden”) factors in the background, the «voxApp» application gives immediate feedback to users in form of directly interpretable results regarding fluctuations of affective state over time (“biofeedback”).

The «voxApp» application allows users to monitor affect- and stress-induced behavior over time by means of repeated 2-minute voice recordings at 1-day intervals. Specifically, the «voxApp» application looks for deviations from each subject’s individual baseline which exceed language-, gender-, and age-specific thresholds. However, short-term deviations from “normal” values in the range of 1-2 days are constituents of human life and do not require specific action.

Principal Operation

The «voxApp» application is an easy-to-use tool for monitoring the effects of chronic stress on body and mind through voice analysis with repeated assessments over time. It has been specifically designed for self-assessments in the home environment. The «voxApp» application has a modular structure and is comprised of four different sections: «Setup» [initial setup], «Recording» [standardized voice recordings], «Analysis» [detailed analysis of single assessments], and «Longitudinal Evaluation» [monitoring the effects of chronic stress over time]. The button “►” [“advance”] let users browse through the four sections. The “Exit” button terminates the program. The «voxApp» application provides directly interpretable results (“biofeedback”). This type of feedback can strengthen resilience and can induce positive behavioral changes. Emphasis lies on “immediate” where results are presented in such a way that they are directly understandable by users and can help to better cope with the ups and downs of the daily grind. In fact, getting involved and doing something about it is the most important step for people with an elevated risk of stress-induced health problems.

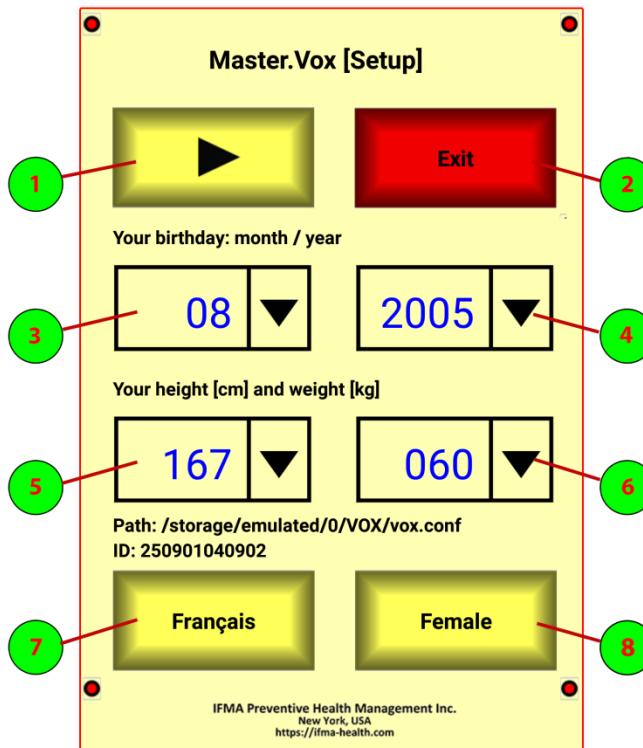


① Advance: <Setup | Recording | Analysis | Longitudinal > browses through sections.

② Exit: terminates the «voxApp» application.

Initial Setup

Reliable assessments of speaking behavior and voice sound characteristics critically depend on gender, age, and spoken language. Available Languages are: English, French, German, Italian, and Spanish. Users must specify the correct data to make sure that the right normative data are used as reference for data analyses.

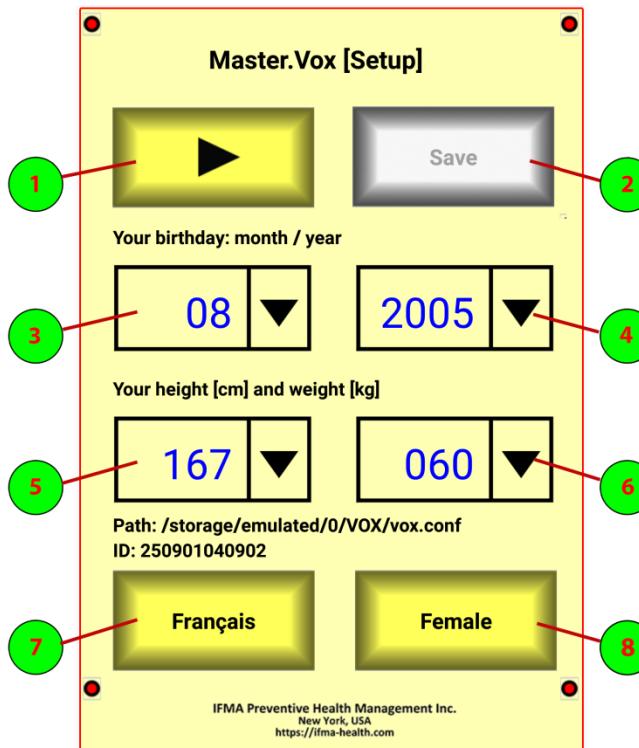


- ① Advance <Setup | Recording | Analysis | Longitudinal>: browses through sections.
- ② Exit: terminates the «voxApp» application.
- ③ Month [birthday]: selects month of birthday.
- ④ Year [birthday]: selects year of birthday.
- ⑤ Height [cm]: selects body height in [cm].
- ⑥ Weight [kg]: selects body weight in [kg].
- ⑦ Language <English | French | German | Italian | Spanish>: selects language.
- ⑧ Gender <Female | Male>: selects gender.

Data Privacy: All information is processed in a strictly anonymous way and used only for the purpose of the «voxApp» project based on statistical (epidemiological) analyses.

Modifying Setup Parameters

Parameters specified during initial setup can always be modified by returning to the setup screen again. Modifications must be saved explicitly through the “Save” button, otherwise modifications are ignored. **Please note:** if key parameters (language, gender, age) are modified new recordings will not be compatible with previous ones.



- ① Advance <Setup | Recording | Analysis | Longitudinal>: browses through sections.
- ② Save: saves modifications (activated when data have been modified).
- ③ Month [birthday]: selects month of birthday.
- ④ Year [birthday]: selects year of birthday.
- ⑤ Height [cm]: selects body height in [cm].
- ⑥ Weight [kg]: selects body weight in [kg].
- ⑦ Language <English | French | German | Italian | Spanish>: selects language.
- ⑧ Gender <Female | Male>: selects gender.

Data Privacy: All information is processed in a strictly anonymous way and used only for the purpose of the «voxApp» project based on statistical (epidemiological) analyses.

Voice Recordings: Overview

We analyze two pieces of spoken text when assessing the speakers' affective state through speaking behavior and voice sound characteristics: (1) Counting out loud from 1-30; and (2) Reading out loud a standard text. The speakers' affective state is influenced by feelings like sadness, happiness, grief, guilt, fear, anger, aggression, shame, love, doziness, or stress – and, occasionally, by depressive or psychotic symptoms.

Speech characteristics can be described by a few major features: speech flow, loudness, and intonation, along with the distribution and intensity of overtones which constitute the speakers' vocal "timbre". Speech flow relates to the speed at which utterances are produced, as well as to number and duration of temporary breaks in speaking. Loudness reflects the amount of energy associated with the articulation of utterances and, when regarded as a time-varying quantity, the speaker's dynamic expressiveness. Intonation is the manner of producing utterances with respect to rise and fall in pitch. Intonation leads to tonal shifts in either direction of the speaker's mean vocal pitch.

Overtones are the higher tones which faintly accompany a fundamental tone, thus being responsible for the tonal diversity of sounds. Distribution and intensity of overtones make up the subjects' individual voice sound characteristics, influenced in a characteristic way by positive and negative feelings.

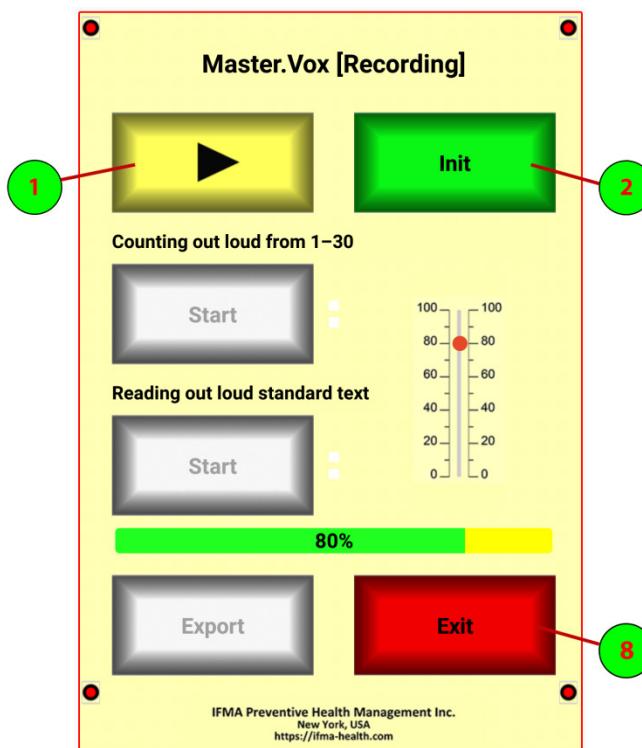
Users should find a quiet place for voice recordings which have to be carried out under comparable experimental conditions in repeated assessments. Users can either use the smartphone's internal microphone, or an external microphone attached with its small clip to a suitable location of the clothing (similar location in repeated assessments).

Audio signals are recorded with a sampling rate of 48 kHz and at a 16-bit resolution (stereo). The resulting quality depends on the audio equipment in use. External microphones are more versatile and in many cases better suited for general-purpose use than the smartphones' or tablets' internal devices. In particular, dynamic "Lavalier" microphones offer good response characteristics over the frequency range of 50 Hz - 18 kHz, are small in size, fit neatly near the neck of a speaker, and clip freely to almost any piece of clothing. Thus, lavalier microphones provide a hassle-free technique of recording an audio source at high quality.

Voice Recordings: Initialization

Voice recordings include two pieces of spoken text: (1) Counting out loud from 1-30; and (2) Reading out loud a standard text. The entire recording procedure takes about 2-3 minutes. To start a recording session, the «voxApp» recording system must be initialized by means of the “Init” button. This, in turn, will activate the “Start” button for the “Counting out loud” task. Once started, the “Stop” button is activated which is used to terminate the recording session. Upon completion, the “Start” button of the “Reading out loud” task becomes active (again terminated through the “Stop” button).

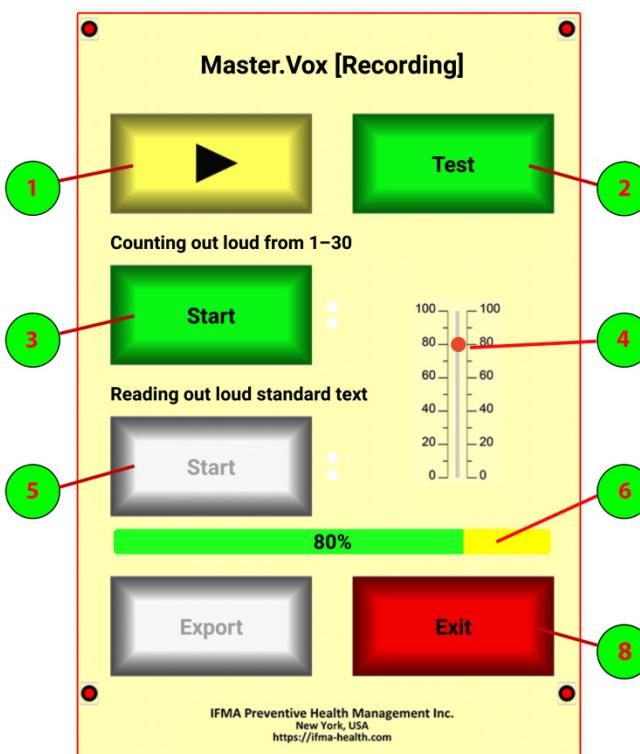
Users should present the two texts with their “normal” voice and sufficient loudness (intensity of voice). The “peak meter” should not exceed 90% during a session. If necessary, the volume can be adjusted by means of the “volume control” on the right side. Each recording task can be repeated by clicking the respective “Start” button again (sessions can be repeated several times). A standard text is given in the Appendix.



- ① Advance <Setup | Recording | Analysis | Longitudinal>: browses through sections.
- ② Init: initializes the «voxApp» recording system.
- ③ Exit: terminates the recording session.

Voice Recordings: System Test

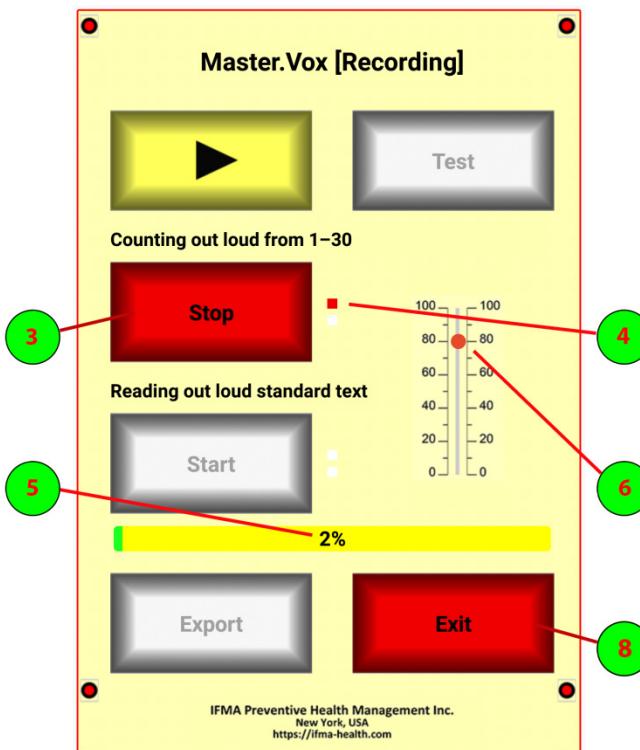
The spoken text used for voice analysis is critically important for reliable assessments of speaking behavior and voice sound characteristics. The «voxApp» application relies on two types of speech: (1) Counting out loud from 1-30 [automatic speech]; and (2) Reading out loud a standard text [emotionally neutral]. Users should present texts with their normal voice and sufficient loudness (intensity of voice).



- ① Advance <Setup | Recording | Analysis | Longitudinal>: browses through sections.
- ② **Test:** verifies activation of the recording system.
- ③ **Start:** starts recording of “counting out loud” task and activates “Stop” button, used to finish the recording session.
- ④ **Volume Control:** adjusts amplification if necessary (a factor of “80%” works perfectly fine in most cases).
- ⑤ **Start:** starts recording of “reading out loud” task and activates “Stop” button, used to finish the recording session.
- ⑥ **Peak Meter:** should not exceed 90% during a recording session.
- ⑦ **Exit:** terminates the recording session.

Recordings: Counting out loud from 1-30

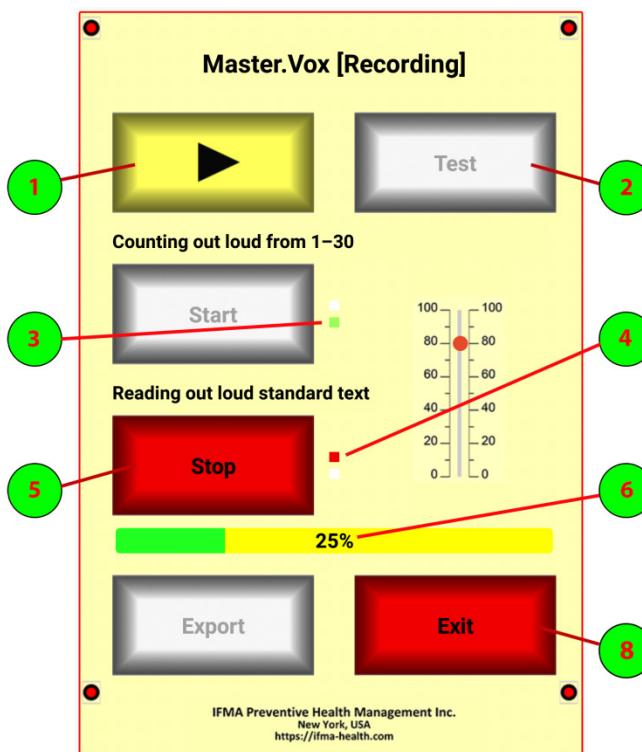
The spoken text used for voice analysis is critically important for reliable assessments of speaking behavior and voice sound characteristics. This is particularly true when assessing the time development of stress-related affective states. The «voxApp» application relies on two types of speech: (1) Counting out loud from 1-30 [automatic speech]; and (2) Reading out loud a standard text [emotionally neutral]. Users should present texts with their normal voice and sufficient loudness (intensity of voice).



- ③ **Stop:** upon start of the “counting out loud” task the “Stop” button is activated which finishes the recording session.
- ④ **Activity LED:** indicates that the microphone is active. It turns to green when clicking the “Stop” button and the recording was successful.
- ⑤ **Peak Meter:** should not exceed 90% during a recording session.
- ⑥ **Volume Control:** adjusts amplification if necessary (a factor of “80%” works perfectly fine in most cases).
- ⑧ **Exit:** terminates the recording session.

Recordings: Reading out loud Text

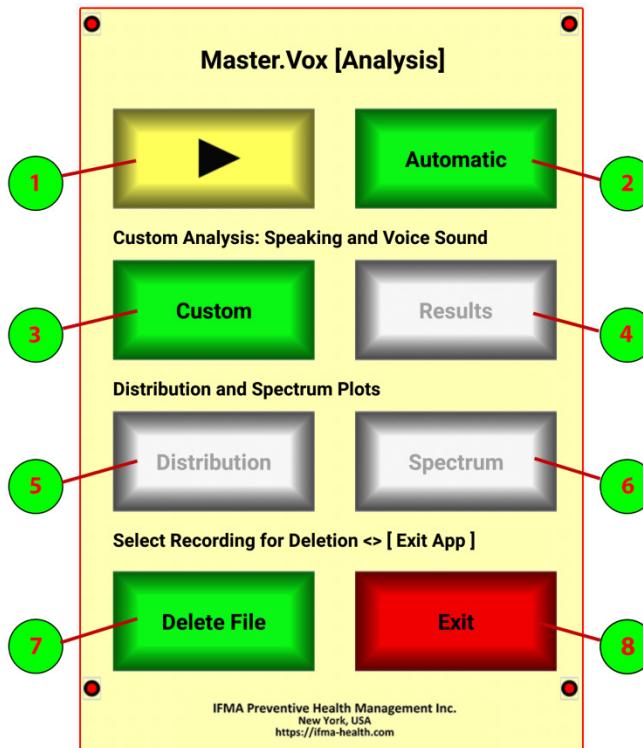
The «voxApp» application relies on two types of speech: (1) Counting out loud from 1-30 [automatic speech]; and (2) Reading out loud standard text [emotionally neutral]. Principally, any emotionally neutral text of 2 minutes length can be used – the same, however, through all repeated assessments. The standard text given in the Appendix has proven its suitability in 5 different languages. Users should print it out for their recordings. Texts have to be presented with normal voice and sufficient loudness (intensity of voice).



- ① Advance <Setup | Recording | Analysis | Longitudinal>: browses through sections.
- ② Test: used to verify activation of the recording system.
- ③ Completion LED: indicates successful completion of the “counting out loud” task.
- ④ Activity LED: indicates that the microphone is active for the “reading out loud standard text” task. It turns to green when clicking the “Stop” button and the recording was successful.
- ⑤ Stop: finishes the “reading out loud” task and activates the green “Completion LED”.
- ⑥ Peak Meter: should not exceed 90% during a recording session.
- ⑦ Exit: terminates the recording session.

Analysis: Automatic versus Customized

The analysis of single assessments can be carried out either automatically (“Automatic”) which is in most cases the preferred method, or in a customized way (“Custom”) which yields more detailed insights into speaking behavior and voice sound characteristics.

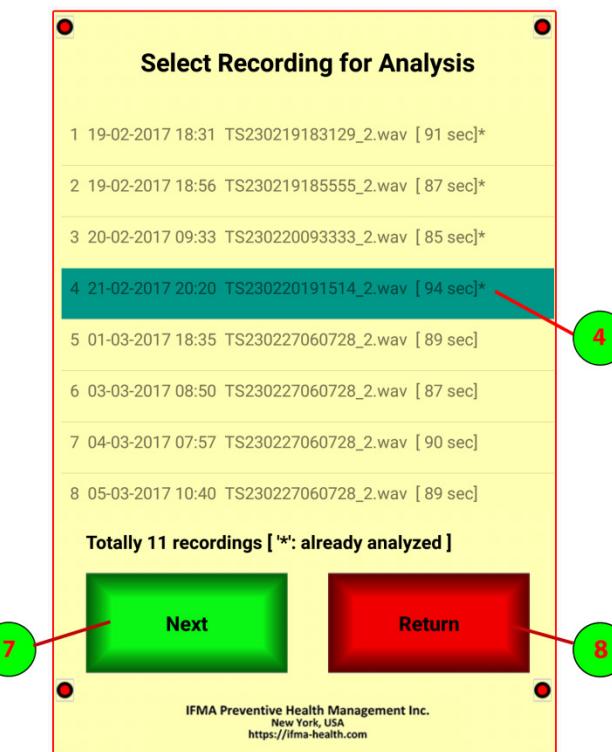


- ① Advance <Setup | Recording | Analysis | Longitudinal>: browses through sections.
- ② Automatic: analyzes all, not yet analyzed recordings without user intervention.
- ③ Custom: analyzes one single, user selectable recording with user intervention.
- ④ Results: the “Results” button is activated upon completion of the analysis; it lets users inspect the outcome of each analysis.
- ⑤ Distribution: the “Distribution” button is activated upon completion of the analysis; it lets users browse through the distribution plots: “Pause Duration”, “Utterance Duration”, “Loudness (Energy)”, “Vocal Pitch (Intonation)”, “F0-Amplitude”, and “55-440 Hz Power”.
- ⑥ Spectrum: the “Spectrum” button is activated upon completion of the analysis; it lets users browse through the spectra computed for consecutive 2-sec epochs.
- ⑦ Delete File: lets users delete one or more recordings.
- ⑧ Exit: exits the «voxApp» application.

Select Recording for Analysis

Selects the recording to be analyzed by highlighting the respective file name (in the example: file name: TS230220191514_2.wav; recording date February 21st, 2017 at 8:20 pm).

Recordings that have previously been analyzed are marked by “*” yet can be re-analyzed at any time. The analysis is started by clicking the “Next” button. Typically, the analysis takes some 20-30 seconds to complete.

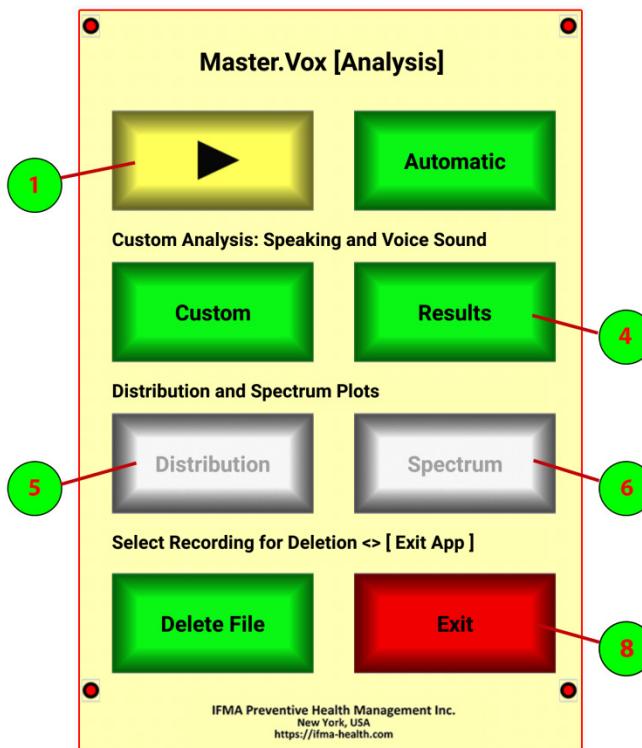


- ④ **File name:** selects recording for analysis by highlighting the respective file name.
- ⑦ **Next:** starts analysis of the chosen recording.
- ⑧ **Return:** goes back to previous section.

Please note: Recordings are stored as standard *.wav files (48 kHz sampling rate, 16 Bit resolution, stereo) in Folder “VOX/myWavFiles”. Thus, files can easily be copied to external devices (“backup”) and played back with standard playback software.

Analysis: Inspecting Results

Data analyses yield directly interpretable results regarding “speech flow”, “loudness”, “intonation” [speaking behavior], along with the distribution and intensity of “overtones” which make up a speaker’s individual vocal “timbre” [voice sound characteristics].

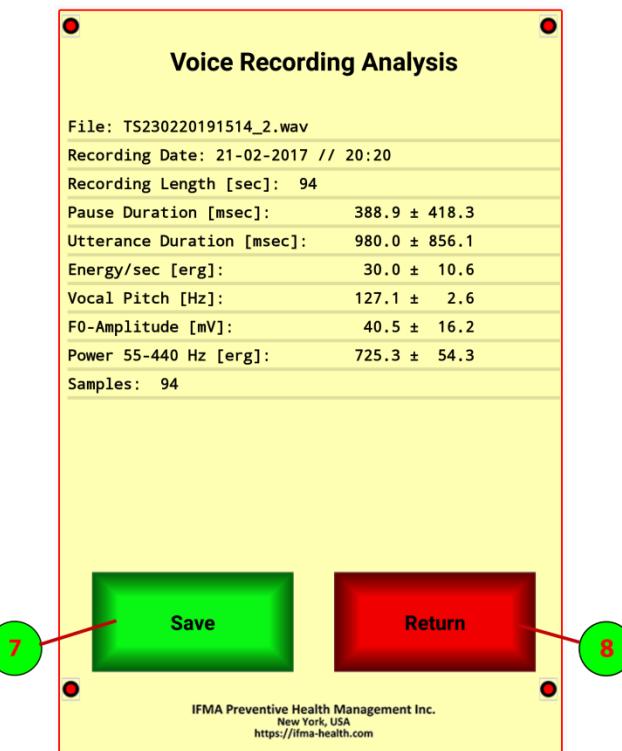


- ① Advance <Setup | Recording | Analysis | Longitudinal>: browses through sections.
- ② Automatic: analyzes all, not yet analyzed recordings without user intervention.
- ③ Custom: analyzes one single, user selectable recording with user intervention.
- ④ Results: the “Results” button is activated upon completion of the analysis; it lets users inspect the outcome of each analysis.
- ⑤ Distribution: the “Distribution” button is activated upon completion of the analysis; it lets users browse through the distribution plots: “Pause Duration”, “Utterance Duration”, “Loudness (Energy)”, “Vocal Pitch (Intonation)”, “F0-Amplitude”, and “55-440 Hz Power”.
- ⑥ Spectrum: the “Spectrum” button is activated upon completion of the analysis; it lets users browse through the spectra computed for consecutive 2-sec epochs.
- ⑦ Delete File: lets users delete one or more recordings.
- ⑧ Exit: exits the «voxApp» application.

Analysis: Inspecting Results

Single assessment data analysis yields directly interpretable quantitative results in terms of “Pause Duration”, “Utterance Duration”, “Energy/sec”, “Vocal Pitch”, “F0-Amplitude”, and “Power 55-440 Hz”. The within-subject variation of these quantities is highly informative in the sense of “biofeedback” regarding the speaker’s current affective state and stress response. These variations can be displayed by means of the separate buttons (“Distribution” and “Spectrum”).

Please note: results of single assessment analyses must be saved explicitly by clicking the “Save” button so that they can be used in subsequent longitudinal analyses.



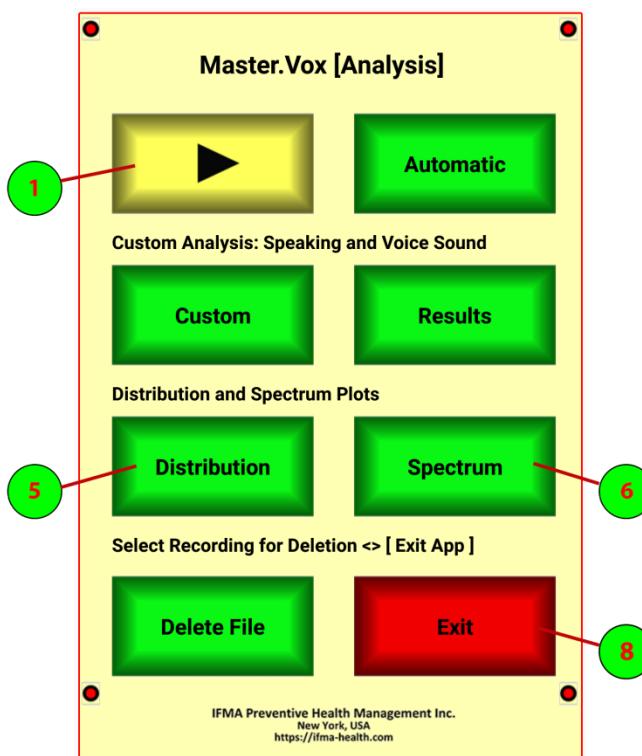
7 Save: saves results for longitudinal analyses.

8 Return: goes back to previous section.

Please note: Recordings are stored as standard *.wav files (48 kHz sampling rate, 16 Bit resolution, stereo) in Folder “VOX/myWavFiles”. Thus, files can easily be copied to external devices (backup) and played back with standard playback software.

Inspecting Distribution Curves and Spectra

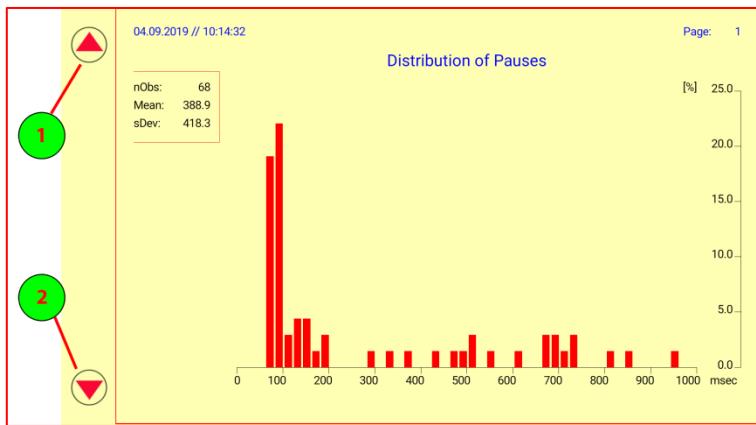
Distribution plots provide directly interpretable results (“biofeedback”) in terms of “Pause Duration”, “Utterance Duration”, “Energy/sec”, “Vocal Pitch”, “F0-Amplitude”, and “Power 55-440 Hz”. The within-subject variation of these quantities is highly informative and reflects the speaker’s current affective state and stress response. These variations can be displayed by means of the button “Distribution”. Additionally, the button “Spectrum” generates plots regarding combination and intensity of tones that constitute the speakers’ vocal sound characteristics.



- ① Advance <Setup | Recording | Analysis | Longitudinal>: browses through sections.
- ⑤ Distribution: the “Distribution” button is activated upon completion of the analysis; it lets users browse through the distribution plots: “Pause Duration”, “Utterance Duration”, “Loudness (Energy)”, “Vocal Pitch (Intonation)”, “F0-Amplitude”, and “55-440 Hz Power”.
- ⑥ Spectrum: the “Spectrum” button is activated upon completion of the analysis; it lets users browse through the spectra computed for consecutive 2-sec epochs.
- ⑧ Exit: exits the «voxApp» application.

Results: Interpretation of Distribution Plots

Distribution plots provide directly interpretable results (“biofeedback”) for a piece of spoken text. The «voxApp» application relies on a standard text of about two minutes length in order to analyze the variation of a speaker’s “Pause Duration”, “Utterance Duration”, “Energy/sec”, “Vocal Pitch”, “F0-Amplitude”, and “Power 55-440 Hz”.



1 Backward: goes back to the previous distribution plot.

2 Forward: advances to the subsequent distribution plot.

The buttons «▲» (backward) and «▼» (forward) let users browse through the distribution pages: (1) Variation of Pause Duration; (2) Variation of Utterance Duration; (3) Variation of Loudness (Energy); (4) Variation of Vocal Pitch (Intonation); (5) Variation of F0-Amplitude; and (6) Variation of 55-440 Hz Power.

The characteristics of each of these distributions provide valuable information about the speaker’s current affective state and stress response. For example:

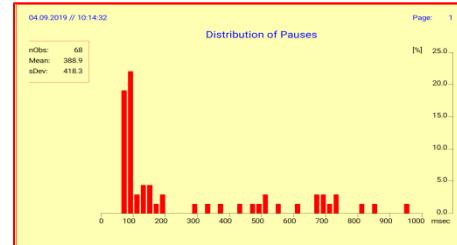
- people under chronic stress tend to present a text in a more monotone, automatized way that lacks variation;
- relaxed speakers typically vary the speed by which they produce utterances in order to gain more attention;
- the same is true for the variation of loudness (“dynamic expressiveness”) which is often completely absent in a state of low mood;
- a pronounced lack of variation in mean vocal pitch (“intonation”) or in F0-amplitude typically means deficiency of emotions and empathetic feelings.

Typical Examples of Distribution Plots

Distribution plots provide directly interpretable results (“biofeedback”) rather than abstract classifiers with unknown (“hidden”) factors in the background.

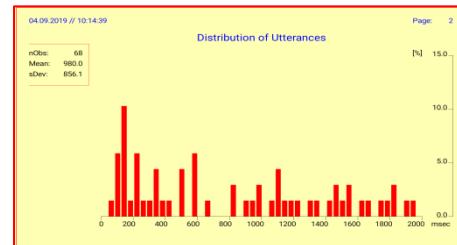
Variation of Pause Duration

When presenting a text, relaxed speakers produce a variety of short pauses and separate main text sections by longer pauses. This is in contrast to speakers under chronic stress who tend to present a text in a more monotone, automatized way that lacks variation.



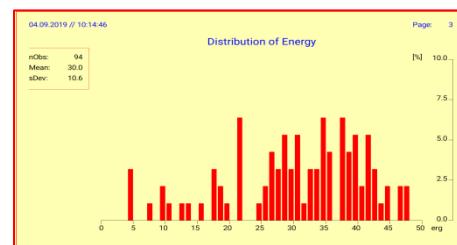
Variation of Utterance Duration

When presenting a text, relaxed speakers typically vary the speed by which they produce utterances in order to make the presentation more attractive and interesting. People under chronic stress wouldn't do this.



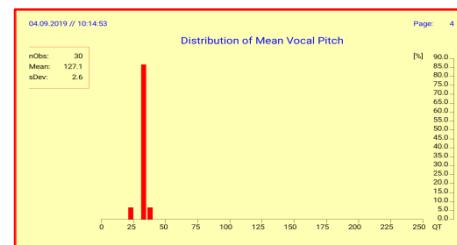
Variation of Loudness (Energy)

When presenting a text, relaxed speakers typically vary loudness (dynamic expressiveness) to make their presentation more attractive and interesting. The width of the distribution of loudness (energy) is a direct measure of dynamic expressiveness.



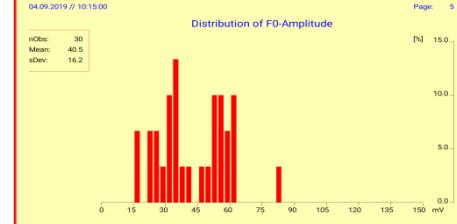
Variation of Vocal Pitch (Intonation)

Intonation is the manner of producing utterances with respect to rise and fall in pitch, and leads to tonal shifts in either direction of the speaker's mean vocal pitch. The “broader” the variation the “richer” the intonation (the example shows a more monotonous speech with a pronounced lack of intonation).



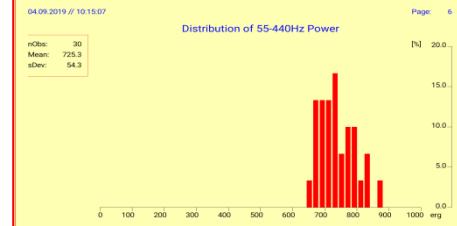
Variation of F0-Amplitude

The variation of “F0-Amplitude” is an indicator of the “richness” of a voice sound. A narrow distribution typically means deficiency of emotions and empathetic feelings. By contrast, a broad distribution suggests a lively and mindful person.



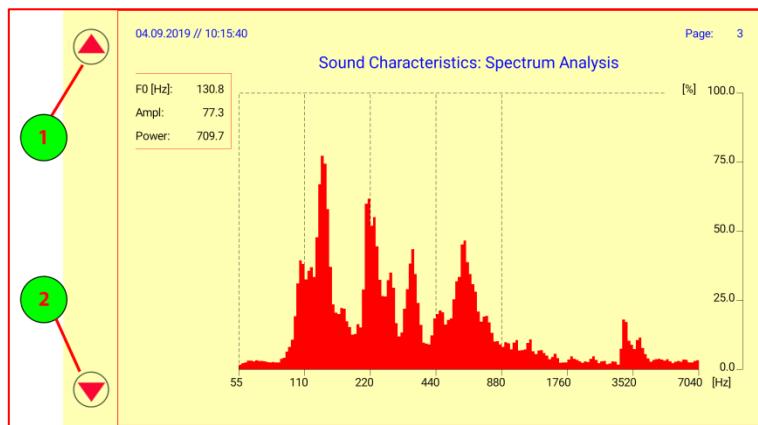
Variation of 55-440 Hz Power

The variation of “55-440 Hz Power” is another measure of the tonal “richness” of a speaker's voice. Reduced variation indicates a more monotonous presentation of utterances caused, for example, by sorrowfulness, weariness, or fatigue.



Results: Interpretation of Spectrum Plots

Spectrum plots reveal the combination and intensity of tones that constitute the sound of speech segments. The first maximum represents the speaker's mean vocal pitch F0, the other maxima show the higher harmonics of F0 (overtones) relevant for the vocal "timbre" of a voice. Mean vocal pitch of average male speakers lies around 110 Hz, that of average female speakers about one octave higher (220 Hz).



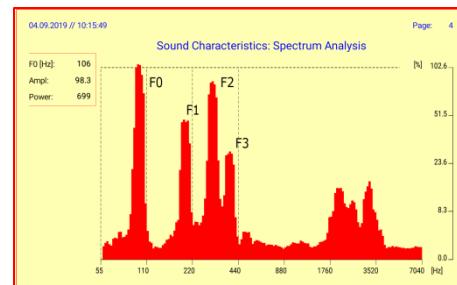
1 Backward: goes back to the previous spectrum plot.

2 Forward: advances to the subsequent spectrum plot.

The buttons «▲» (backward) and «▼» (forward) let users browse through the spectrum pages: spectra are computed for consecutive 2-sec epochs so that sound quality can be visualized as a function of time.

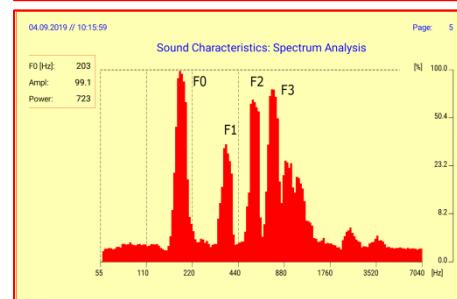
Spectra of Male Speakers

Spectra of male speakers typically display a mean vocal pitch of about 110 Hz. In the plots on the right the intensity of intrinsic tones is plotted along the y-axis for the frequency range of 55-7040 Hz (7 octaves). As to overtones: F1 is the octave above the fundamental tone F0; F2 the fifth above F1; and F3 the fourth above F2 (vowel "a").



Spectra of Female Speakers

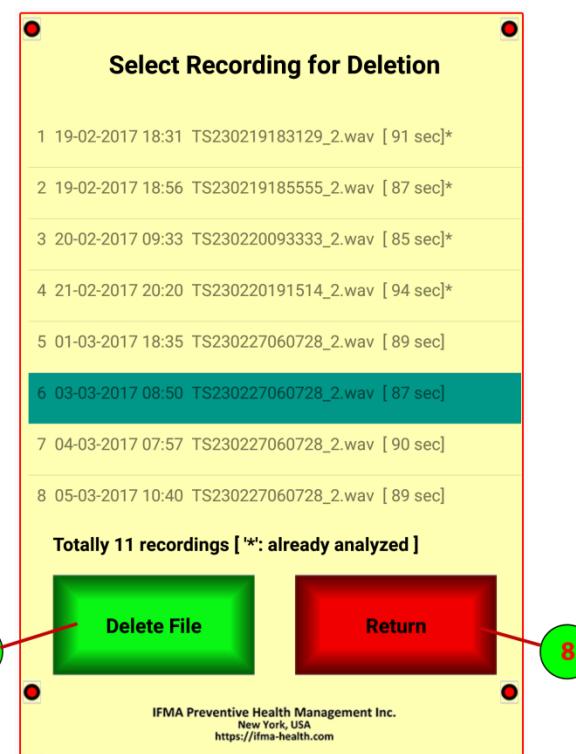
Spectra of typical female speakers display a mean vocal pitch of 220 Hz which is one octave above that of the average male speaker (110 Hz). As to overtones: the octave above F0 at 400 Hz, the fifth above F1 at 600 Hz, the forth above F2 at 800 Hz (vowel "a").



How to Delete Recordings

There are situations where users may want to delete a recording: (1) the recording is of insufficient quality; (2) the recording is outdated and no longer of interest.

Please note: once deleted, a recording cannot be recovered unless the user has made a backup on another device. Therefore, it is highly recommended to regularly backup important data. Select the recording that should be deleted by highlighting the respective file name (in the example: recording date March 3rd, 2017 at 8:50 am; file name: TS230227060782_2.wav). Then click the “Delete File” button and confirm deletion. Recordings that have already been analyzed are marked by “*”.

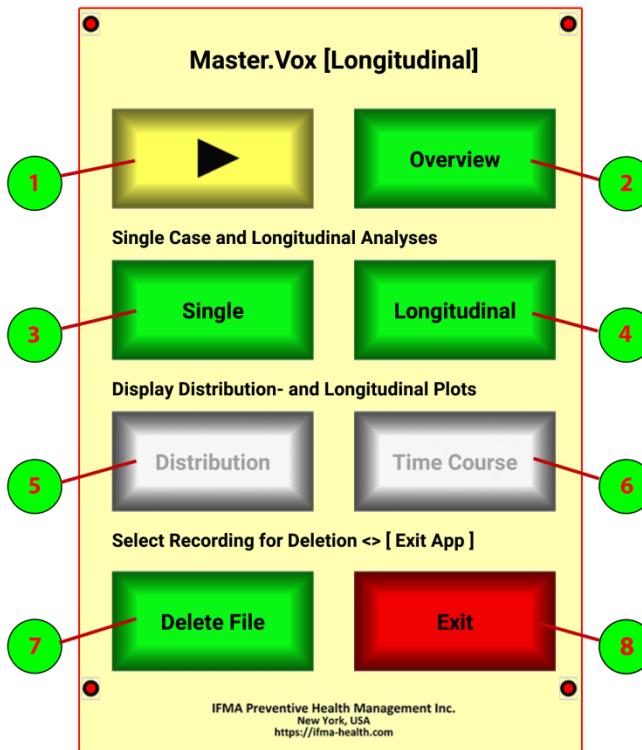


- 7 **Delete File:** deletes the highlighted recording from the «voxApp» system.
- 8 **Return:** goes back to the previous section.

Please note: Recordings are stored as standard *.wav files (48 kHz sampling rate, 16 Bit resolution, stereo) in Folder “VOX/myWavFiles”. Thus, files can easily be copied to external devices (backup) and played back with standard playback software.

Longitudinal Analysis: Time Development

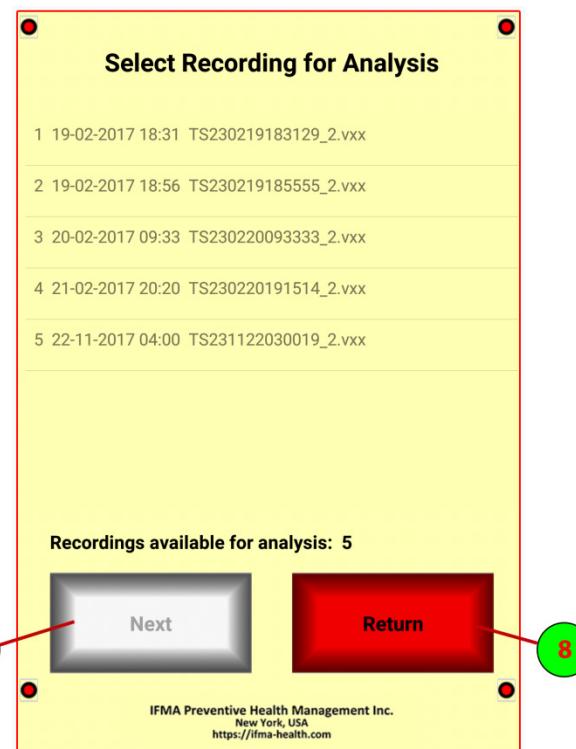
Voice analysis has its focus on within-subject –rather than between-subject– variations. That is, focus is laid on single case analysis with repeated assessments over a certain time period where each individual person serves as his/her own reference.



- ① Advance <Setup | Recording | Analysis | Longitudinal>: browses through sections.
- ② Overview: lists all recordings available for longitudinal analyses.
- ③ Single: starts user selected single recording analyses.
- ④ Longitudinal: starts longitudinal analyses of multiple recordings; results can be visually inspected through time course plots.
- ⑤ Distribution: activated upon completion of a single recording analysis; it lets users browse through the resulting distribution plots: “Pause Duration”, “Utterance Duration”, “Loudness (Energy)”, “Vocal Pitch (Intonation)”, “F0-Amplitude”, and “55-440 Hz Power”.
- ⑥ Time Course: activated upon completion of the longitudinal analysis of multiple recordings; it lets users browse through the resulting time course plots.
- ⑦ Delete File: lets users delete one or more recordings.
- ⑧ Exit: exits the «voxApp» application.

Select Recording for Single Analyses

The previously saved results of a single recording can be opened for inspection. A recording is selected by highlighting the respective file name. Once a recording is highlighted, the “Next” button is activated. Single assessment data analysis yields directly interpretable quantitative results in terms of “Pause Duration”, “Utterance Duration”, “Energy”, “Dynamics”, “Vocal Pitch”, “Intonation”, “F0-Amplitude”, and “Power 55-440 Hz”. The within-subject variation of these quantities is highly informative in the sense of “biofeedback” regarding the speaker’s current affective state and stress response. The analysis is started by clicking the “Next” button. Typically, the analyses take some 10-15 seconds to complete.



7 **Next:** lets users inspect the previously saved results of a recording.

8 **Return:** goes back to the previous section.

Please note: File names are of the form “TS230227060728_2.vxx” with extension “vxx”. Files are stored in standard ASCII format and can be opened with any text editor such as “notepad”. Consecutive data values are separated by “;” from each other so that files can easily be imported into MS Excel.

Longitudinal Analysis: Inspecting Results

Of interest are within-subject fluctuations in speaking behavior and voice sound characteristics over time, along with the question of whether or not the fluctuations exceed “normal” values and, therefore, deserve special attention.

Longitudinal Analysis (1)

Date / Time	Pauses [msec]	Utters [msec]	Energy [erg]	Dynam [erg]
0 19-02-2017 19:48	358.2	953.1	24.0	8.5
1 19-02-2017 19:50	327.1	926.2	27.9	9.3
2 20-02-2017 09:33	348.3	1035.0	26.7	7.6
3 21-02-2017 20:20	388.9	980.0	30.0	10.6
4 22-11-2017 04:00	0.0	0.0	16.5	4.5

Records: 5 Date: 19-02-2017 < 22-11-2017

Save
Next

1
2

IFMA Preventive Health Management Inc.
New York, USA
<https://ifma-health.com>

Longitudinal Analysis (2)

Date / Time	Pitch [Hz]	Inton [QT]	F0-Amp [mV]	Power [erg]
0 19-02-2017 19:48	113.2	4.2	48.7	721.5
1 19-02-2017 19:50	110.0	3.7	49.7	699.5
2 20-02-2017 09:33	127.1	2.1	45.5	690.6
3 21-02-2017 20:20	127.1	2.6	40.5	725.3
4 22-11-2017 04:00	130.8	16.8	17.2	460.6

Records: 5 Date: 19-02-2017 < 22-11-2017

Save
Return

3
4

IFMA Preventive Health Management Inc.
New York, USA
<https://ifma-health.com>

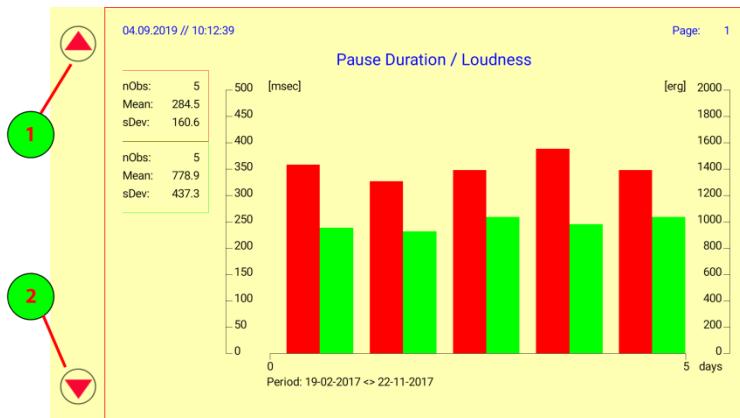
- ① **Save:** lets users save the results (part 1).
- ② **Next:** proceeds to the 2nd part of the results.
- ③ **Save:** lets users save the results (part 2).
- ④ **Return:** goes back to the previous section.

Please note: The longitudinal analysis includes automatically all available recordings that have been analyzed and explicitly saved (cf. p. 16) – sorted according to ascending date. Up to 30 recordings (1 month) can be evaluated as “time course”. The «voxApp» application takes the 30 most recent recordings as default.

A minimum number of 5 recordings is required for sufficiently reliable conclusions.

Longitudinal Analysis: Time Course Plots

Of interest are within-subject fluctuations in speaking behavior and voice sound characteristics over time, along with the question of whether or not these fluctuations exceed "normal" values and, therefore, deserve special attention.



- ➊ **Backward:** goes back to the previous time course plot.
- ➋ **Forward:** advances to the subsequent time course plot.

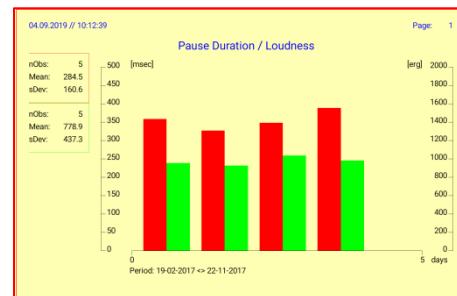
The buttons «▲» (backward) and «▼» (forward) let users browse through the time course plots: (1) Pause Duration / Loudness; (2) Energy / Dynamics; (3) Vocal Pitch / Intonation; and (4) F0-Amplitude / 55-440 Hz Power. The interpretation of these time course plots is given on the following page.

Interpretation of Time Course Plots

Of interest are within-subject fluctuations in speaking behavior and voice sound characteristics over time, along with the question of whether or not these fluctuations exceed "normal" values and, therefore, deserve special attention.

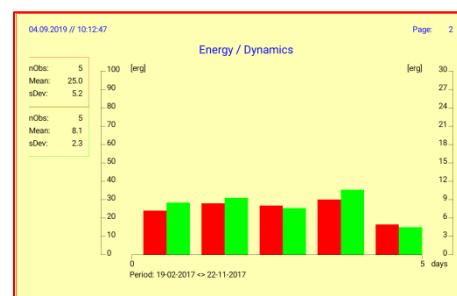
Pause Duration / Loudness

Despite their great stability over time, the parameters "Pause Duration" and "Loudness" often show a systematic trend toward shorter pauses and greater loudness when speakers get used to the test. Longitudinal analyses compensate for this effect.



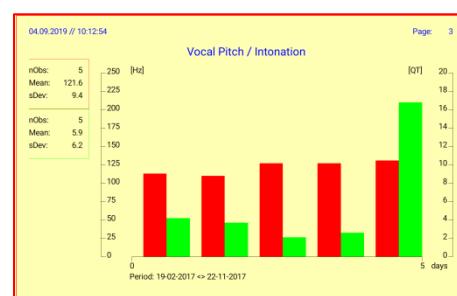
Energy / Dynamics

In contrast to healthy subjects, patients suffering from affective disorders speak in a low voice, slowly, hesitatingly, monotonously, sometimes stuttering or whispering. During recovery, however, patients regain their energy and dynamic expressiveness.



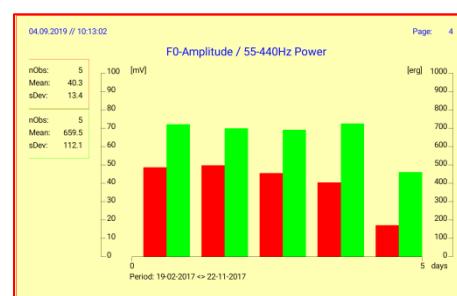
Vocal Pitch / Intonation

This plot monitors the time course of the test person's mean vocal pitch and intonation. It is worth noting that persons under chronic stress tend to speak a halftone above their "natural" mean vocal pitch. Vocal pitch shows "normal" values when coping skills improve.



F0-Amplitude / 55-440 Hz Power

This plot monitors the time course of the test person's voice sound characteristics. It is worth noting that persons under chronic stress tend to have a "sharp", sometimes "metallic" voice sound. Persons regain their bright and full timbre when coping skills improve.

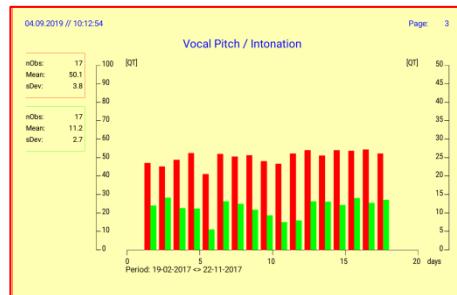


Typical Examples of Time Course Plots

Of interest are within-subject fluctuations in speaking behavior and voice sound characteristics over time, along with the question of whether or not these fluctuations exceed "normal" values and, therefore, deserve special attention.

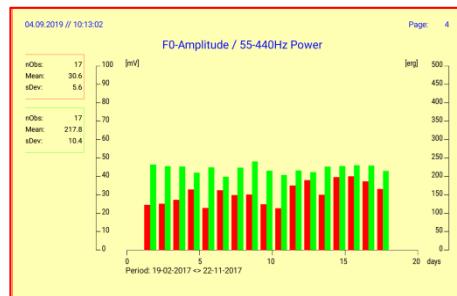
Vocal Pitch / Intonation

Mean vocal pitch and intonation typically display a picture of great stability over time. However, habituation effects along with routine may occasionally lead to better intonation. Short-term deviations from "normal" values in the range of 1-2 days are constituents of human life and do not require specific action.



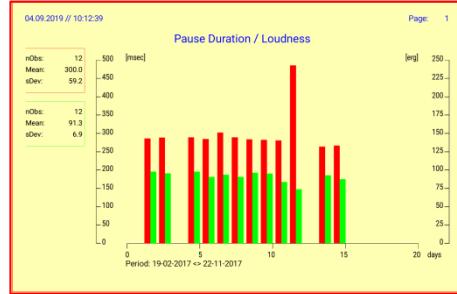
F0-Amplitude / 55-440 Hz Power

In contrast to healthy subjects, patients suffering from affective disorders speak in a low voice, slowly, hesitatingly, monotonously, sometimes stuttering or whispering. During recovery, however, patients regain their energy and dynamic expressiveness.



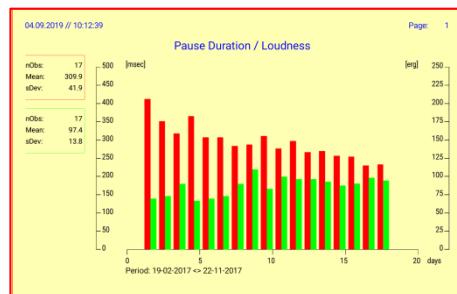
Pause Duration / Loudness

Pause duration (red bars) and loudness (green bars) over an observation period of 14 days: Speaking behavior is virtually unchanged over time except for day 12 with longer pauses and a lower voice (subject may have been tired). No recordings on days 4 and 13.



Pause Duration / Loudness

Pause duration (red bars) and loudness (green bars) over an observation period of 14 days: Speaking behavior shows a systematic trend towards shorter pauses and greater loudness as a function of time. This finding may indicate habituation effects.



Standard text

The thing which attracted her most, however, was the waving and roaring of the three old fir trees on these windy days. She would run away repeatedly from whatever she might be doing, to listen to them, for nothing seemed so strange and wonderful to her as the deep mysterious sound in the tops of the trees. She would stand underneath them and look up, unable to tear herself away, looking and listening while they bowed and swayed and roared as the mighty wind rushed through them.

There was no longer now the warm bright sun that had shone all through the summer, so Heidi went to the cupboard and got out her shoes and stockings and dress, for it was growing colder every day, and when Heidi stood under the fir trees the wind blew through her as if she was a thin little leaf, but still she felt she could not stay indoors when she heard the branches waving outside.

Then it grew very cold, and Peter would come up early in the morning blowing on his fingers to keep them warm. But he soon left off coming, for one night there was a heavy fall of snow and the next morning the whole mountain was covered with it, and not a single little green leaf was to be seen anywhere upon it.

There was no Peter that day, and Heidi stood at the little window looking out in wonderment, for the snow was beginning again, and the thick flakes kept falling till the snow was up to the window, and still they continued to fall, and the snow grew higher, so that at last the window could not be opened, and she and her grandfather were shut up fast within the hut.

Heidi thought this was great fun and ran from one window to the other to see what would happen next, and whether the snow was going to cover up the whole hut, so that they would have to light a lamp although it was broad daylight. But things did not get as bad as that, and the next day, the snow having ceased, the grandfather went out and shoveled away the snow round the house, and threw it into such great heaps that they looked like mountains standing at intervals on either side the hut.

And now the windows and door could be opened, and it was well it was so, for as Heidi and her grandfather were sitting one afternoon on their three-legged stools before the fire there came a great thump at the door, followed by several others, and then the door opened. It was Peter, who had made all that noise knocking the snow off his shoes; he was still white all over with it, for he had had to fight his way through deep snowdrifts, and large lumps of snow that had frozen upon him still clung to his clothes. He had been determined, however, not to be beaten and to climb up to the hut, for it was a week now since he had seen Heidi.

"Good-evening", he said as he came in; then he went and placed himself as near the fire as he could without saying another word, but his whole face was beaming with pleasure at finding himself there. Heidi looked on in astonishment, for Peter was beginning to thaw all over with the warmth, so that he had the appearance of a trickling waterfall.

Texte standard

La chose qui l'attirait le plus, cependant, c'était l'agitation et le rugissement des trois vieux sapins en ces jours de grands vents. Elle s'enfuyait à plusieurs reprises peu importe ce qu'elle devait faire, pour les écouter, rien ne lui semblait plus étrange et plus merveilleux que le profond et mystérieux son de la cime des arbres. Elle se tenait debout, sous eux, à les regarder, incapable de s'en détacher, à les regarder et à les écouter pendant qu'ils s'inclinaient et se balançaient à cause du puissant vent qui se précipitait à travers eux.

Il n'y avait désormais plus le soleil chaud qui avait brillé tout au long de l'été, alors Heidi alla au placard et sortit ses souliers, ses bas et sa robe, puisqu'il faisait de plus en plus froid chaque jour. Quand Heidi se tenait sous les sapins, le vent soufflait à travers elle comme si elle était une fine et petite feuille, mais elle sentait qu'elle ne pouvait pas rester à l'intérieur quand elle entendait les branches s'agiter dehors. Puis il fit froid, et Peter arrivant tôt le matin soufflait sur ses doigts pour les réchauffer. Mais bientôt il arrêta de venir. Pendant une nuit, il y eut une grande tombée de neige, et le lendemain matin, la montagne entière était recouverte de neige, et plus une seule feuille verte n'était visible nulle part.

Il n'y avait pas Peter ce jour-là, et Heidi se tenait à la petite fenêtre, en regardant dehors avec émerveillement la neige qui tombait à nouveau. Les flocons épais qui ne cessaient de tomber arrivèrent jusqu'à la fenêtre, et continuèrent encore de tomber, et la neige de monter, de façon à ce qu'à la fin, la fenêtre ne puisse plus être ouverte, et qu'elle et son grand-père soient rapidement enfermés dans la hutte. Heidi, pensant que c'était très amusant, courra d'une fenêtre à l'autre pour voir ce qui allait arriver, et si la neige allait recouvrir toute la hutte, de sorte qu'ils auraient dû allumer une lampe bien qu'il fasse grand jour. Mais les choses n'étaient pas si mauvaises que ça, et le lendemain, la neige avait cessé. Le grand-père sortit et pelleta la neige autour de la maison, il la jeta dans des grands tas qui ressemblaient à des montagnes positionnées à des intervalles de chaque côté de la cabane.

Maintenant, les fenêtres et la porte pouvaient être ouvertes, et il en était bien ainsi, puisque pendant qu'Heidi et son grand-père étaient assis un après-midi sur leur tabouret à trois-pieds, il y eut un grand bruit sourd à la porte, suivi de plusieurs autres, puis la porte s'est ouverte. C'était Peter, qui avait fait tout ce bruit pour enlever la neige de ses chaussures; il était encore tout blanc, car il avait dû se frayer un chemin à travers des congères profondes et de larges morceaux de neige étaient gelés sur lui, accrochés à ses vêtements. Il avait été déterminé qu'il viendrait à ce moment car cela faisait maintenant une semaine qu'il n'avait pas vu Heidi. "Bonsoir", dit-il dès qu'il entra; puis il alla se placer aussi près qu'il pouvait du feu sans dire un mot, mais tout son visage était rayonnant de plaisir du fait qu'il se trouvait là. Heidi le regarda avec étonnement, car Peter commençait à fondre de partout avec la chaleur, ce qui lui donnait l'apparence d'une cascade.

Standardtext

Das schönste war für Heidi an solchen Windtagen das Wogen und Rauschen in den drei alten Tannen hinter der Hütte. Oft stand Heidi unten und lauschte hinauf. Jetzt schien die Sonne nicht mehr heiß wie im Sommer, und Heidi suchte Strümpfe und Schuhe hervor und auch ihr Röckchen, denn nun wurde es immer frischer.

Dann wurde es kalt, und auf einmal fiel über Nacht tiefer Schnee, und am Morgen war die ganze Alm weiß und kein einziges grünes Blättchen mehr zu sehen. Nun kam der Geißen Peter nicht mehr mit seiner Herde.

Heidi schaute durch das kleine Fenster und beobachtete verwundert, wie die dicken Flocken immerzu fielen, bis der Schnee an das Fenster hinaufreichte. Und dann lag er noch höher, so dass man das Fenster gar nicht mehr aufmachen konnte und in dem Häuschen ganz verpackt war.

Heidi fand das so lustig, dass sie von einem Fenster zum anderen rannte, um zu sehen, ob der Schnee noch die ganze Hütte zudecken würde. Es kam aber nicht so weit.

Am nächsten Tag schneite es nicht mehr. Der Großvater ging hinaus, schaufelte um das ganze Haus herum und warf große Schneehaufen aufeinander, hier einen Berg und dort einen Berg um die Hütte herum. Nun waren die Fenster wieder frei und auch die Tür.

Das war gut, denn als am Nachmittag Heidi und der Großvater am Fenster saßen, polterte auf einmal jemand gegen die Holzschwelle und machte endlich die Tür auf. Es war der Geißen Peter. Er hatte aber nicht aus Übermut so laut gepoltert, sondern um den Schnee von seinen Schuhen abzuschlagen, der in dicken Klumpen daran klebte. Der ganze Peter war von Schnee bedeckt, denn er hatte sich durch die hohen Schichten so durchkämpfen müssen, dass große Stücke an ihm hängen geblieben und in der scharfen Kälte an ihm festgefroren waren. Aber er hatte nicht nachgegeben, heute wollte er zu Heidi hinauf, denn er hatte sie acht Tage lang nicht gesehen.

„Guten Abend“, sagte er, stellte sich gleich so nahe wie möglich ans Feuer und sagte weiter nichts mehr. Aber sein ganzes Gesicht lachte vor Freude.

Heidi schaute ihn verwundert an, denn nun begann es überall an ihm zu tauen, so dass das Wasser an Peter herabließ wie ein Wasserfall.

„Nun, General, wie steht's?“ fragte der Großvater. „Bist du ohne Armee und musst am Griffel nagen?“

„Warum muss er am Griffel nagen, Großvater?“ fragte Heidi neugierig.

„Im Winter muss er in die Schule gehen“, erklärte der Großvater; „da lernt man lesen und schreiben. Das ist manchmal schwierig, da hilft's ein wenig nach, wenn man am Griffel nagt. Nicht wahr, General?“

„Ja, das ist wahr“, bestätigte Peter.

Testo standard

Ma quello che Heidi trovava più attraente era osservare l'ondeggiare e lo scrosciare dei rami tra i vecchi pini dietro la capanna, nei giorni ventosi. Heidi si metteva lì sotto ad ascoltare quel che avveniva in alto, ma non riusciva mai a vedere e a udire abbastanza di tutto quell'ondeggiare e strepitare possente che faceva il vento tra le cime. Ormai il sole non era più così caldo come d'estate, e Heidi tirò fuori le calze e le scarpe oltre all'abitino.

Poi venne il freddo. L'indomani tutta l'Alpe apparve candida, non si vedeva più una fogliolina verde all'intorno. Il pastorello non venne più col suo gregge. e Heidi guardava stupefatta fuori dalla finestrella: cominciava a nevicare di nuovo e i densi fiocchi cadevano sempre più fitti, finché la neve divenne così alta da arrivare alla finestra, che, alla fine, non si poté più aprire; ed eccoli ormai rinchiusi nella capanna.

Per Heidi era molto divertente correre da una finestra all'altra per vedere come sarebbe andata a finire e se la neve avrebbe seppellito l'intera capanna tanto da rendere necessario accendere il lume in pieno giorno. Ma le cose non andarono così.

Il giorno seguente il nonno uscì, perché aveva smesso di nevicare, e spalò la neve tutt'attorno alla baita: gettò grossi mucchi di neve uno sopra all'altro, e fu come se si erigesse un monte qui, un altro là. Così le finestre e la porta non furono più ostruite dalla neve, e fu un bene. Infatti, nel pomeriggio, mentre Heidi e il nonno sedevano davanti al fuoco, ognuno sul suo sgabello, si udì a un tratto un gran rumore e colpi sempre più fitti contro la porta di legno. Alla fine, la porta si aprì: era il pastorello. Non aveva fatto tutto quel chiasso per mera maleducazione, ma per togliere la neve dalle scarpe. A dire il vero, Peter era tutto coperto di neve: aveva dovuto lottare talmente contro gli strati di neve, che gli erano rimasti appesi dei grossi pezzi che si erano gelati addosso a lui per il gran freddo. Ma non si era arreso, voleva andare su da Heidi: non la vedeva da otto giorni.

"Buona sera!" disse entrando. Si avvicinò il più possibile al fuoco e non aggiunse altro, ma il suo viso era radioso per il piacere d'essere lì. Heidi lo guardava con stupore, poiché, accanto al fuoco com'era, cominciò a sgocciolare da tutte le parti, così che sembrava trasformato in una cascatella.

Texto estándar

Por encima de todo, a Heidi le gustaba, en aquellos días en que soplaba el viento otoñal, el misterioso runrún de los tres abetos que había detrás de la cabaña. Por lo general Heidi estaba debajo de los abetos y oía el murmullo de los árboles. El sol ya no era tan cálido como en verano y Heidi sacó del armario sus calcetines y sus zapatos y también un vestido de lana, porque hacía cada vez más fresco. Y llegó el frío. Y una mañana todo amaneció teñido de blanco; durante la noche había caído la primera nevada y ya no se veía ni una sola mancha verde. Pedro el cabrero dejó de subir al monte con sus cabras.

Heidi, sentada junto a la ventana, contemplaba cómo caía la nieve en grandes copos, sin interrupción. Tan grande fue la cantidad de nieve caída, que al fin alcanzó el borde inferior de la ventana, y aún seguía subiendo de tal manera que ya no se podía abrir la ventana. Dentro se estaba bien calentito. A Heidi eso le pareció tan divertido que no paraba de correr de una ventana a otra para ver en qué iba a parar todo aquello. Se preguntaba si por fin la nieve cubriría toda la cabaña, y si sería preciso encender las luces en pleno día. Pero las cosas no llegaron a tanto.

Al día siguiente dejó de nevar y el abuelo salió fuera y se puso a quitar la nieve. Con una pala fue amontonando la nieve en varios sitios hasta que las ventanas y las puertas quedaron despejadas. Por suerte el abuelo lo había hecho en el momento oportuno, porque cuando él y Heidi se hallaban por la tarde sentados junto al fuego del hogar, oyeron de pronto recios golpes y patadas delante de la puerta, y al momento entró Pedro el cabrero, que hacía aquel ruido cuando se quitaba la nieve de los zapatos. De hecho, estaba cubierto de nieve porque tuvo que abrirse camino a través de una capa tan densa que grandes trozos quedaron pegados a su ropa por el frío. Pero ni la nieve ni el frío le hicieron renunciar a su empeño: hacía ocho días que no veía a Heidi y la echaba de menos.

“Buenas tardes” dijo al entrar. Después se acercó al fuego y no dijo nada más, pero su rostro expresaba franca alegría por estar allí. Heidi le miraba asombrada ya que se hallaba tan cerca del calor del hogar que la nieve empezó a derretirse y caía de su ropa en forma de lluvia.

“Bien, general, ¿cómo te van las cosas?” preguntó el abuelo. “Ahora te has quedado sin ejército y tienes que morder el lápiz”. “¿Por qué ha de morder el lápiz, abuelo” preguntó Heidi, muy curiosa.

“Durante el invierno, Pedro tiene que ir al colegio” explicó el anciano; “allí se aprende a leer y a escribir y eso, a veces, resulta muy difícil y morder el lápiz ayuda, ¿verdad, general?”

“Sí, es verdad” confirmó Pedro.

References

- Moragrega I, Bridler R, Mohr C, Possenti M, Rochat D, Sanchez Parramon J, Stassen HH: Monitoring the Effects of Therapeutic Interventions in Depression through Self-Assessments. *Res Psychother.* 2021; 24(3): 250-262
- Zhang M, Bridler R, Mohr C, Moragrega I, Sun N, Xu Z, Yang Z, Possenti M, Stassen HH: Early Detection of the Risk of Developing Psychiatric Disorders: A Study of 461 Chinese University Students under Chronic Stress. *Psychopathology* 2019; 52(6): 367-377
- Braun S, Annovazzi C, Botella C, Bridler B, Camussi E, Delfino JP, Mohr C, Moragrega I, Papagno C, Pisoni A, Soler C, Seifritz E, Stassen HH: Assessing Chronic Stress, Coping Skills and Mood Disorders through Speech Analysis. A Self- Assessment "Voice App" for Laptops, Tablets, and Smartphones. *Psychopathology* 2016; 49(6): 406-419
- Delfino JP, Barragán E, Botella C, Braun S, Camussi E, Chafrat V, Mohr C, Bridler R, Lott P, Moragrega I, Papagno C, Sanchez S, Soler C, Seifritz E, Stassen HH: Quantifying Insufficient Coping Behavior under Chronic Stress. A cross-cultural study of 1,303 students from Italy, Spain, and Argentina. *Psychopathology* 2015; 48: 230-239
- Braun S, Botella C, Bridler R, Chmetz F, Delfino JP, Herzig D, Kluckner VJ, Mohr C, Moragrega I, Schrag Y, Seifritz E, Soler C, Stassen HH: Affective State and Voice: Cross-Cultural Assessment of Speaking Behavior and Voice Sound Characteristics. A Normative Multi-Center Study of 577+36 Healthy Subjects. *Psychopathology* 2014; 47(5): 327-340
- Mohr C, Braun S, Bridler R, Chmetz F, Delfino JP, Kluckner VJ, Lott P, Schrag Y, Seifritz E, Stassen HH: Insufficient Coping Behavior under Chronic Stress and Vulnerability to Psychiatric Disorders. *Psychopathology* 2014; 47: 235-243
- Stassen HH, Delfino JP, Kluckner VJ, Lott P, Mohr C: Vulnerabilität und psychische Erkrankung. *Swiss Archives of Neurology and Psychiatry* 2014; 165(5): 152-157
- Lott PR, Guggenbühl S, Schneeberger A, Pulver AE, Stassen HH: Linguistic analysis of the speech output of schizophrenic, bipolar, and depressive patients. *Psychopathology* 2002; 35(4): 220-227
- Stassen HH, Kuny S, Hell D: The speech analysis approach to determining onset of improvement under antidepressants. *Eur Neuropsychopharmacology* 1998; 8,4: 303-310

External Microphones

Lavalier Microphones

External microphones are more versatile and better suited for general-purpose use than the smartphones' or tablets' internal devices. In particular, dynamic "Lavalier" microphones offer good response characteristics over the frequency range of 50 Hz - 18 kHz, are small in size, fit neatly near the neck of a speaker, and clip freely to almost any piece of clothing. Thus, lavalier microphones provide a hassle-free technique of recording an audio source at high quality.

Selecting the Optimal Microphone

Audio signals are recorded with a sampling rate of 48 kHz and at a 16-bit resolution (stereo). The resulting quality depends on the audio equipment in use. In particular, the smartphones' and tablets' internal microphones are sometimes insufficient so that dynamic "Lavalier" microphones can be a much better choice as such devices are less sensitive to background noise and pick up sounds more purely.

Sony ECMCS3

This is a budget-friendly, omnidirectional electret condenser microphone with a frequency response of 50 Hz - 18 kHz (stereo). A TRRS microphone adapter is required to connect the 3.5 mm stereo jack plug to the smartphones' combo audio port.

Audio-Technica ATR-3350iS

This is a high-quality omnidirectional condenser microphone that comes with a foam windscreens, a LR44 battery, and a TRRS microphone adapter. It features a frequency response of 50 Hz - 18 kHz.

Rode smartLav+

This is one of the best Lavalier microphones on the market (condenser capsule). It comes with a foam windscreens and features a frequency response of 60 Hz - 18 kHz. It plugs directly into your smartphone's combo audio port.

Microphone Adapter

Lavalier microphones with a 3.5 mm stereo jack plug are connected to the smartphones' and tablets' combo audio port through a TRRS adapter (\$ 8).



Sony ECMCS3 (\$ 19)



Audio-Technica ATR-3350 (\$ 29)



RODE smartLav+ (\$79)

