



# Audio Recorder with Spectrum Analyzer

High-Quality wav-File Recorder for Smartphones and Tablets

## User Manual (Android)



# What is this App good for?

## Not just another wav-File Recorder

- 1) This is NOT just another wav-File Recorder for high-quality audio recordings at a sampling rate of 48 kHz, a 16-bit resolution, and no loss of information due to compression. Unlike other Apps in the field, the IFMA recorder possesses unique features that allow you (1) to monitor the quality of audio recordings by means of “Time Series Plots”; and (2) to explore the tonal composition of sounds regarding distribution and intensity of higher harmonics (“overtones”) by means of “Spectrum Plots” at a quartertone resolution over 7 octaves.
- 2) The App enables hassle-free audio recordings at high quality at virtually no cost. No other device is required that you have to carry with you and whose batteries have to be charged.
- 3) The App’s GUI does not show mainstream “material” styles. Rather, it looks quite different as our focus lies on the utmost functionality and a perfect readability such that all functions are at your fingertips even in low-light situations.
- 4) The App comes with integrated help files (that work offline) as well as a comprehensive manual with a series of detailed examples.

## Smart Companion for Rehearsals

- 5) A great tool to quantitatively assess the sound characteristics of your voice or musical instrument: find out how things improve over time by means of regular assessments.
- 6) Towards the perfect vibrato: improve your technique by using “Spectrum Plots” as immediate bio-feedback.
- 7) Towards the perfect speech: optimize a speech in terms of phrasing, intonation, accentuation, length of pauses and utterances by means of “Time Series Plots”.

## Getting to Know Yourself Better

- 8) Learn to modify the color (“timbre”) of your voice or musical instrument in a strictly controlled way using “Spectrum Plots” in the sense of an immediate bio-feedback.
- 9) Learn how regular voice recordings can improve your coping behavior under stress, so that you feel much more comfortable, for example, when speaking in public.

And, there is more to explore.

Copyright © 2017 by IFMA - Preventive Health Management Inc., all rights reserved.

# Table of Contents

Background	4
Chapter 1: Help Pages	5
Instructions for Use	
Carrying out Recordings	
Time Series Plots	
Spectrum Plots	
Chapter 2: Data Storage	6
List of Available Recordings	
System Initialization	
Volume Adjustments	
Imprint	
Chapter 3: Recordings	7
Carrying out a Recording Session	
Re-Initialize the Recording System	
Start New Recording Session	
External Microphone	
Chapter 4: Time Series Plots	8
Select Recording	
Browsing through the Plot Pages	
Screenshots	
Exit Time Series Plots	
Chapter 5: Spectrum Plots	9
Select Recording	
Browsing through the Plot Pages	
Screenshots	
Exit Spectrum Plots	
External Microphones	10
References	11

## Background

Speech production is the result of a joint effort of mind and body. It involves a cascade of steps from utterance planning to final sound production with hundreds of degrees of freedom. Rhythm, stress, and intonation (“prosody”) greatly influence the verbal and nonverbal content of the transmitted speech. Voice sound characteristics relate to the subjects’ mean vocal pitch along with the distribution and intensity of “overtones” that make up the subjects’ individual voice “timbres”. Overtones are the higher tones which faintly accompany a fundamental tone. These overtones are responsible for the tonal diversity of sounds and make voices and musical instruments distinguishable from each other.

Despite this complexity, speech characteristics can be roughly described by a few major features. Speaking behavior can be modeled in terms of “speech flow”, “loudness”, and “intonation”. Speech flow describes the speed at which utterances are produced, as well as the 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, and leads to tonal shifts in either direction of the speaker’s mean vocal pitch.

Experienced speakers, vocalists and instrumentalists can voluntarily modify the “timbre” of their voice or their musical instrument in a strictly controlled way, for example, from “lyric” to “dramatic” or from “lovesome” to “hateful”. This is accomplished by varying the amplitude envelope and spectral envelope of a particular sound or a sequence of sounds (“vibrato”). Vibrato is a technique that produces a shimmery variation in pitch and is used to “warm up” the sound of a voice or a musical instrument such as a flute or a violin. On the other hand, when attempting to cope with stressful situations, subjects tend to speak in a hectic, abrupt, and often monotonous way. Most notably, however, subjects under stress typically speak a half-tone above their “natural” mean vocal pitch – an effect that is highly unwanted as it can immediately be perceived by any attentive listener in the audience.

Spectrum analysis reveals the sound characteristics of voices and musical instruments by displaying a fundamental tone along with the distribution and intensity of its higher harmonics (“overtones”) in the form of intensity versus frequency plots. Thus, users can visually assess the tonal composition of a sound as well as its characteristic variation over time.

Time series plots allow users to verify the overall quality of a recording by looking for clipped amplitudes in intervals with audio signal and by examining intervals without audio signal for unwanted background noise. All this can help when testing the environmental conditions that are most appropriate for the envisaged recording sessions, or when selecting the optimal location for an external microphone (“Lavalier”) which has to be attached with its small clip to a suitable location of the speaker’s or vocalist’s clothing.

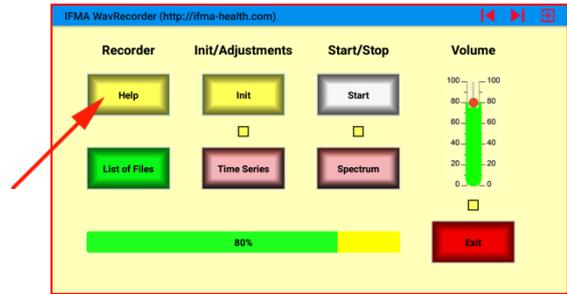
# Chapter 1

## Help Pages

The help pages are intended to give immediate support should questions arise when using the app. Therefore, the pages provide a short overview about the app's central features. A more detailed description is given in this manual which can be downloaded from website "http://ifma-health.com".

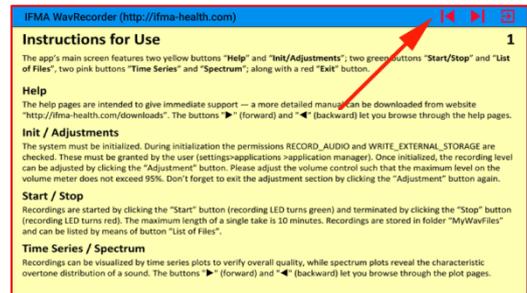
## Instructions for Use

The app's main menu features 7 buttons whose functions are described herein for first-time users. You can browse through the help pages by means of the buttons "▶" (forward) and "◀" (backward). The right-most button "Exit" brings you back to the app's main menu.



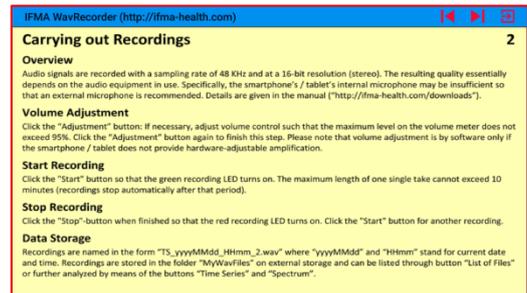
## Carrying out Recordings

Please find a quiet place for your recordings, adjust volume level adequately, and avoid amplitude clipping. Click the "Start" button so that the green recording LED turns on. The maximum length of one single session must not exceed 10 minutes. Click the "Stop"-button when finished so that the red recording LED turns on.



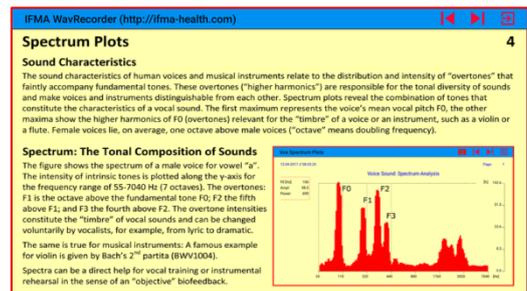
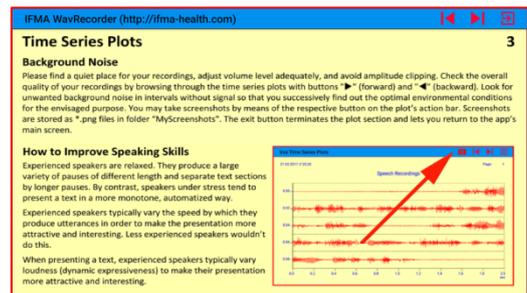
## Time Series Plots

The overall quality of your recordings can be verified by browsing through the time series plots. Look for clipped amplitudes in intervals **with** signal as well as unwanted background noise in intervals **without** signal so that you can successively improve recordings by optimizing each session's environmental conditions.



## Spectrum Plots

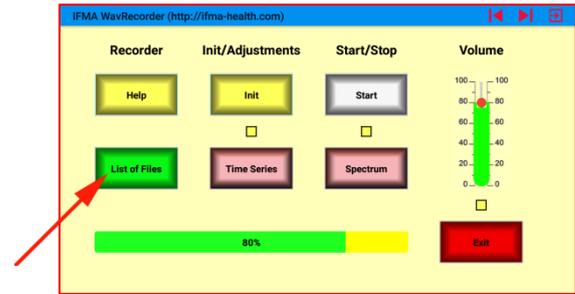
Spectrum plots reveal the combination of tones that constitute the characteristics of a sound. The first maximum shows the sound's fundamental tone F0, the other maxima relate to the higher harmonics of F0 relevant for the "timbre" of a voice or an instrument, such as a violin or a flute.



# Chapter 2

## Data Storage

Recordings are stored in the folder “**MyWavFiles**” on external storage and can be listed through button “**List of Files**” or further analyzed by means of the buttons “**Time Series**” and “**Spectrum**”. Recordings are named in the form “**TS\_yyyyMMdd\_HHmm\_2.wav**” where “**yyyyMMdd**” and “**HHmm**” stand for current date and time.



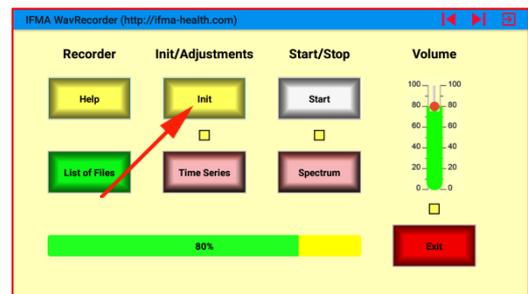
## List of Available Recordings

Provides a list of recordings available for further analysis, for example (1) “**Time Series Plots**” for testing the overall quality of a recording regarding clipped values and background noise, and (2) “**Spectrum Plots**” for analyzing combination and intensity of the tones that underlie a sound. The list is scrollable.



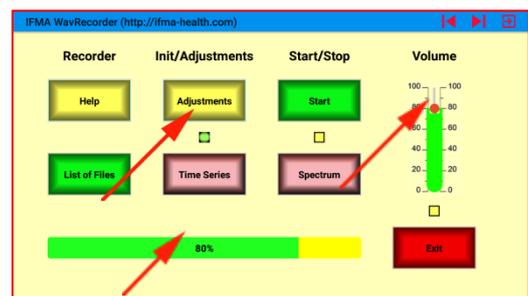
## System Initialization

The system has to be initialized prior to working with it. During initialization the permissions “**Record\_Audio**” and “**Write\_External\_Storage**” are checked. These must be granted explicitly by the user: **settings** ▶ **applications** ▶ **application manager**.



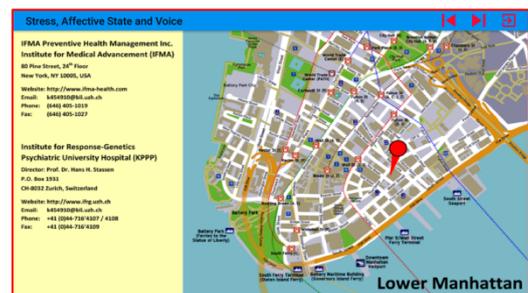
## Volume Adjustments

Once initialized, the recording level can be adjusted by clicking the “**Adjustment**” button. Provide a test sound that is representative of the envisaged recording and adjust the volume control such that the maximum level on the volume meter does not exceed 95%. You exit the adjustment section by clicking the “**Adjustment**” button again.



## Imprint

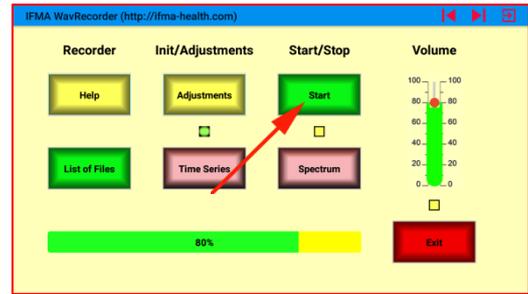
IFMA Preventive Health Management Inc., New York, USA;  
Institute for Response-Genetics, University of Zurich,  
Switzerland



# Chapter 3

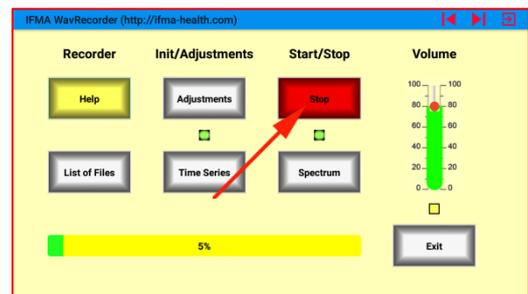
## Recordings

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 smartphone's internal microphone may be insufficient so that external microphones may be a better choice (they are less sensitive to background noise and pick up sounds more purely).



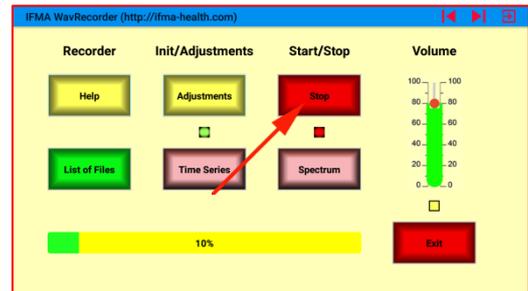
### Carrying out a Recording Session

Click the “Start” button so that the green recording light turns on and the “Stop” button shows up. Carry out the recording and click the corresponding “Stop”-button when finished. The maximum length of a single recording session is 10 minutes after which the recording stops automatically.



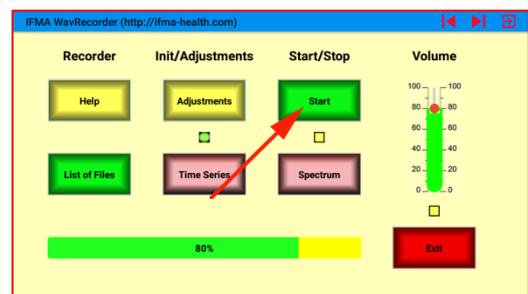
### Re-Initialize the Recording System

To re-initialize the recording system click the “Stop” button again so that the green “Start” button shows up. Then you have the opportunity to adjust the volume level again by clicking the “Adjustment” button. Do not forget to exit the adjustment section by clicking the “Adjustment” button again (the adjustment LED turns red).



### Start New Recording Session

Click the “Start” button so that the green recording light turns on and the “Stop” button shows up. Carry out the recording and click the corresponding “Stop”-button when finished. The maximum length of a single recording session is 10 minutes after which the recording stops automatically.



### External Microphone

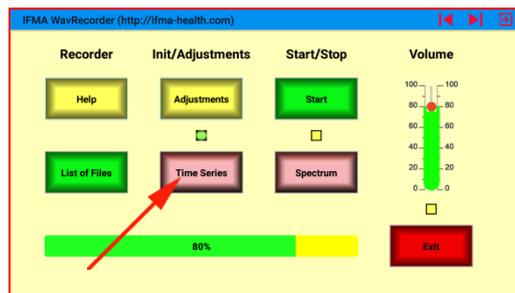
You can either use the internal microphone of your smartphone, or an external microphone attached with its small clip to a suitable location of your clothing. It may be a good idea to test various positions. Subsequently, use the optimal location for best results in repeated recordings.



# Chapter 4

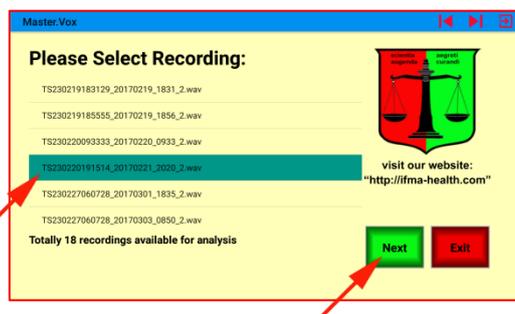
## Time Series Plots

The overall quality of your recordings can be verified by browsing through the time series plots. Look for clipped amplitudes in intervals **with** signal and unwanted background noise in intervals **without** signal so that you can successively improve recordings by optimizing each session's environmental conditions.



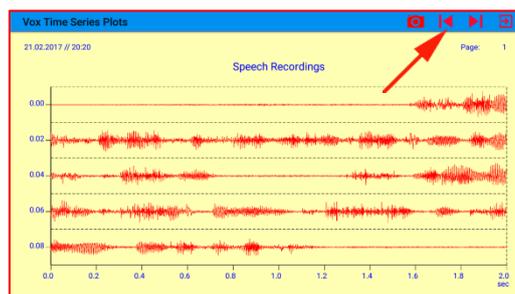
### Select Recording

Select the recording that you want to display by clicking the respective file name (here: recording date February 21, 2017 at 8:20 pm). Recordings that have already been analyzed are marked by “.” yet can be re-analyzed at any time. Click “Next” to display the time series plots or “Exit” to return to the main menu.



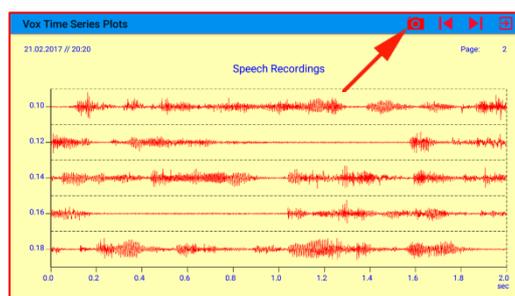
### Browsing through the Plot Pages

The buttons “▶” (forward) and “◀” (backward) let you browse through the time series plots. Look for clipped amplitudes in intervals **with** signal and unwanted background noise in intervals **without** signal. The plot also provides valuable insights in the length and distribution of pauses, or in the recording's organization regarding phrases and sections.



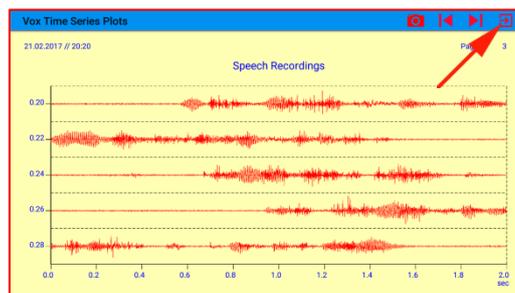
### Screenshots

You may take screenshots by means of the respective button on the plot's action bar. Screenshots are stored as “\*.png” files in folder “MyScreenshots”. The successful completion of a screenshot is shown by a short message.



### Exit Time Series Plots

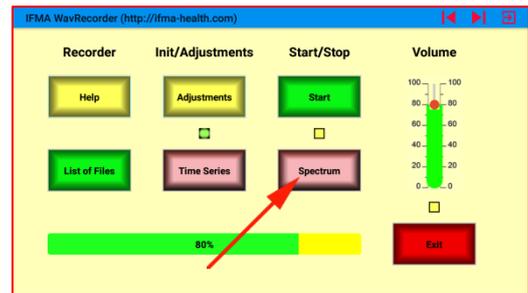
The plot's “Exit” button terminates the plot section and lets you return to the app's main menu.



# Chapter 5

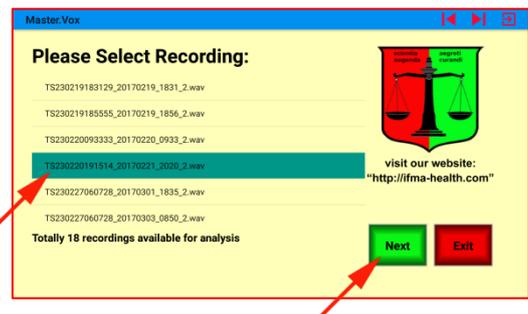
## Spectrum Plots

The sound characteristics of human voices and musical instruments relate to the distribution and intensity of “overtones” that faintly accompany fundamental tones. Overtones (“higher harmonics”) are responsible for the tonal diversity of sounds and make voices and instruments distinguishable from each other. Spectrum plots reveal these sound characteristics.



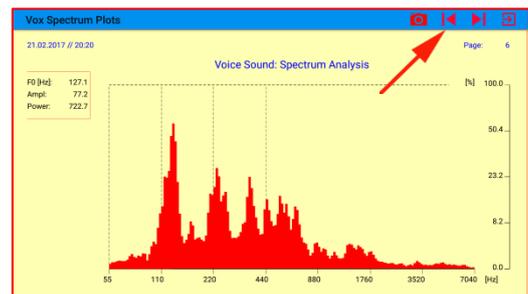
### Select Recording

Select the recording that you want to analyze by clicking the respective file name (here: recording date February 21, 2017 at 8:20 pm). Recordings that have already been analyzed are marked by “\*”, yet can be re-analyzed at any time. Click “Next” to display the time series plots or “Exit” to return to the main menu.



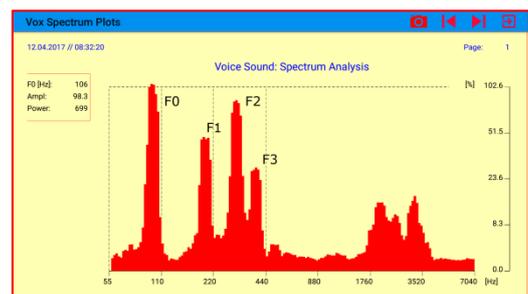
### Browsing through the Plot Pages

The buttons “▶” (forward) and “◀” (backward) let you browse through the spectrum plots. The intensity of tones inherent in a sound is plotted along the vertical axis over the frequency range of 55-7,040 Hz (7 octaves at a quartertone resolution). The first maximum represents the fundamental tone F0, the other maxima show the higher harmonics of F0.



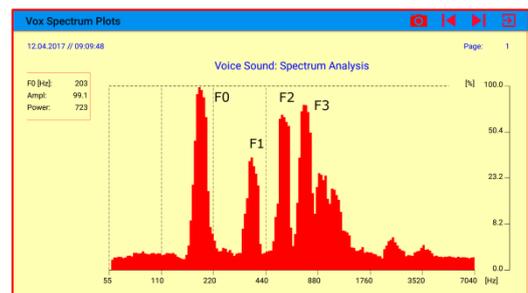
### Male Voice

The figure shows the spectrum of the sound of a typical male voice (vowel “a”). The intensity of intrinsic tones is plotted along the vertical axis over the frequency range of 55-7,040 Hz (7 octaves at a quartertone resolution). As to overtones: F1 is the octave above the fundamental tone F0 (100 Hz); F2 the fifth above F1; and F3 the fourth above F2.



### Female Voice

Spectra of typical female voices display a mean vocal pitch of 220 Hz which is one octave above that of the average male voice (110 Hz). The figure shows a female’s spectrum with the following overtones: F1 is the octave above F0 at 400 Hz; F2 the fifth above F1 at 600 Hz; and F3 the fourth above F2 at 800 Hz (vowel “a”).



# External Microphones

## Lavalier Microphones

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.

## 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 windscreen, 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 windscreen 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).



# References

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

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

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

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, Kuny S, Hell D: The speech analysis approach to determining onset of improvement under antidepressants. *Eur Neuropsychopharmacology* 1998; 8,4: 303-310