



**HAL**  
open science

# VOCALAB: A USER-FRIENDLY SOFTWARE FOR VOICE ANALYSIS AND THERAPY

Etienne Sicard, Anne Menin-Sicard, Louise Ruffle

► **To cite this version:**

Etienne Sicard, Anne Menin-Sicard, Louise Ruffle. VOCALAB: A USER-FRIENDLY SOFTWARE FOR VOICE ANALYSIS AND THERAPY. Choice for Voice, Sep 2021, London (on line), United Kingdom. hal-03353490

**HAL Id: hal-03353490**

**<https://hal.science/hal-03353490v1>**

Submitted on 24 Sep 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Conference “Choice for Voice: Crossing boundaries in voice”

Virtual 3-day conference: 3-5th September 2021, The British Voice Association (BVA)

Extended paper

# VOCALAB: A USER-FRIENDLY SOFTWARE FOR VOICE ANALYSIS AND THERAPY

E. Sicard<sup>1</sup>, A. Menin-Sicard<sup>2</sup>, L. Ruffle<sup>3</sup>

(1) LURCO research director, INSA Toulouse, France, [etienne.sicard@insa-toulouse.fr](mailto:etienne.sicard@insa-toulouse.fr)

(2) Speech & Language Therapist and LURCO research affiliate, France, [anne.sicard2@orange.fr](mailto:anne.sicard2@orange.fr)

(3) Speech & Language Therapist in independent practice, France, [ruffle.orthophonie@gmail.com](mailto:ruffle.orthophonie@gmail.com)

## Abstract

This paper presents the Vocalab software program, a voice analysis tool designed for use by speech therapists in private practice. Vocalab has a user-friendly interface presented in the form of different modules: patient information, assessment, therapy, and the media library. The assessment module displays voice recordings as a spectrogram and produces a voice report with 5 main voice perturbation measures: pitch instability, amplitude instability, noise/signal ratio, harmonic poverty and attack perturbation. The therapy module offers exercises for working on specific aspects of voice production with real-time visual feedback, presented in an attractive way in order to appeal to both adults and children, and facilitate learning. Finally, the media library module offers access to a large database of normal and pathological voices, often with relevant case history information and pre- and post-therapy samples.

## Introduction

Existing software for voice analysis in the English-speaking world has a number of limitations and outside the Ear Nose and Throat (ENT) clinic, therapists generally have limited access to objective measures of voice quality. Perceptual measures of voice are subject to significant inter- and intra-rater reliability and alone are insufficient as outcome measures. Many voice therapists in the UK are reliant on software and/or equipment not specifically designed for clinical use (e.g. acoustic speech analysis programs), which require significant technical knowledge to interpret. More bespoke tools often require bulky equipment and are inaccessible to National Health Service (NHS) teams with restricted budgets.

Perceptual measures alone are insufficient to meet the demands of evidence based practice, and to demonstrate the effectiveness of therapy. Whilst vocal fold stroboscopy can be used to obtain objective measures of voice function, it is not always easily accessible. Regarding therapy, visual biofeedback can enhance patients' understanding and motivation by making the vocal target visible and easy to understand. The VOCALAB Media Library also represents a invaluable source of Continuing

Professional Development (CPD) with an array of learning materials, including tutorials and a large library of case studies to help familiarise practitioners with less common voice pathologies.

While PRAAT [Boersma 2021] is free and versatile thanks to a fully programmable environment, it has been designed primarily for speech analysis and phonetics. It requires significant technical knowledge, and does not provide real-time feedback, which significantly limits its use in voice therapy. VOCALAB offers a more versatile and much more user-friendly tool at low cost. It has been designed with the speech therapist in mind, with a large number of modules offering real-time feedback, each focused on one particular aspect of the voice: timbre, range, resonance, noise, voicing, etc.

VOCALAB originated from the need for an accessible and reliable tool to meet the need of French speech & language therapists (SLTs), the majority of whom are in independent practice. Development began in 2000 and a research team was created to validate the prototype within the French National Speech Therapy Research Lab LURCO [Lurco 2021]. The first edition was commercialised in 2005 by GERIP [Gerip 2021]. A major new version came out in 2014 (version 4.0), followed by a further update in 2019 (version 4.1).

## Research phases

The first phase of the research project attempted to identify the most clinically useful measures of voice, and a review was carried out of more than 50 articles discussing markers of pathological voice [Sicard 2013]. There was a reasonable consensus regarding fundamental frequency, vocal range, noise and maximum phonation time (MPT) [Kent 2000][Baken 2000]. However, the consensus was unclear for *Jitter* (frequency perturbation) and *Shimmer* (amplitude perturbation) with a number of critics raising doubts as to their relevance, specifically for altered and pathological voices [Maryn 2009, 2010].

In order to streamline the final voice report and make it accessible to clinicians and patients, as well as to facilitate comparison between pre- and post therapy measures, 5 key measures were selected: attack perturbation, pitch instability, amplitude instability, noise/signal ratio and harmonic poverty. The attack perturbation measure is unique to VOCALAB, and focuses specifically on the initial portion of a sustained /a:/ which contains valuable acoustic information and may help to identify abnormal voices. The threshold of normal voice was set at 1.0 for all measures, simplifying comparison and eliminating the need to refer to normal and abnormal values (for example: 3% for Jitter, -17 dB for NHR), trend and order of magnitude (5% Jitter is worse than 3%, while -20 dB is better for HNR). In VOCALAB, any value below 0.8 is considered as normal, values between 0.8 & 1.4 are considered threshold, and values above 1.5 are highly likely to be pathological.

The next phase of the research focused on the creation of a database of voices. 10 speech therapists contributed to an initial data set of 250 case samples including sustained /a:/, glides and conversational speech. Case history information was provided (age, gender, pathology, background.). Files were anonymised and coded according to the voice disorder classification from [Verdolini 2006]. Open access to this database is available on the LURCO website [Link to data base]. These cases were used to validate the chosen voice measures over a large number of samples. Further studies focused on the characteristics of children's voices [Sicard 2019] using a corpus of 50 normal cases, and the characteristics of ageing voice [Barbera & Frenay 2016] using 100 normal voices from patients aged between 65 and 90 years old.

## Introducing Vocalab

Vocalab is a Windows-based program requiring a microphone along with the computer. A virtual machine may emulate the Windows environment on Apple computers as well. Suggested microphone references include unidirectional devices such as Thomann MB85 BETA or Shure SM58.

The four main modules on home page are the patient module, the assessment module in green, the therapy module in orange and the media library in blue (Figure 1).



Figure 1: The four main modules on VOCALAB home page

## Patient information

The patient module is very easy to use and it displays all the information, measurements and images available for each patient. It allows the therapist to generate a report in Word format containing the spectrogram, all voice measures and associated thresholds, saving the therapist valuable time.

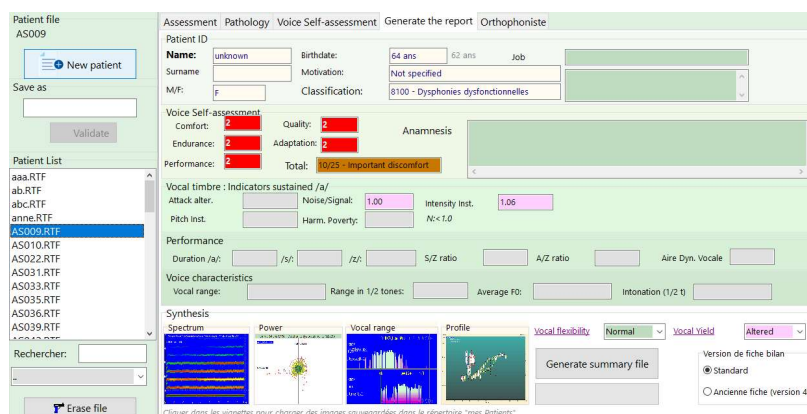
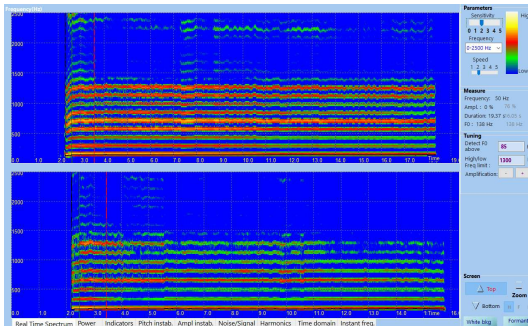


Figure 2: Patient information used to create a report in word format

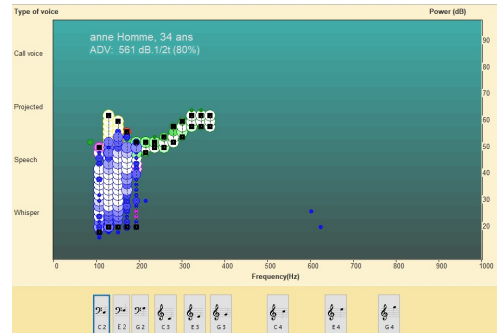
## Assessment Module

The assessment module consists of 4 sub-modules, including the spectrogram displayed in two separate windows, allowing the therapist to give a model and the patient to attempt to copy it. The Maximum Phonation Time module allows the clinician to record a sustained /a/, /s/ and /z/ and calculate the A/Z and S/Z ratio. The patient's score is displayed against the average for normal voice. The Vocal Range module displays the histogram of the patient's fundamental frequency in

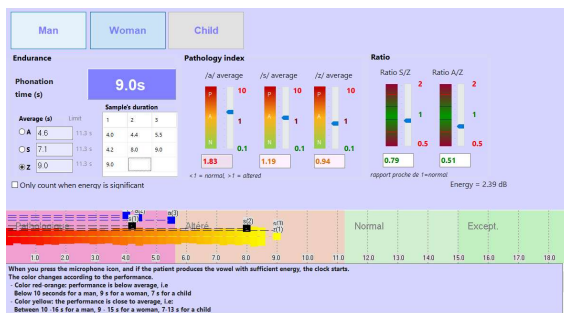
conversational speech or on a glide, underneath the usual ranges for men, women and children. The piano keyboard is also displayed, helping the patient to understand these abstract concepts. The speech range profile (SRP) is accessible through the Phonetogram module (Figure 4). The phonetogram displays the patient's frequency range across a range of volumes, and can provide professional voice users with a more in-depth analysis of their vocal profile.



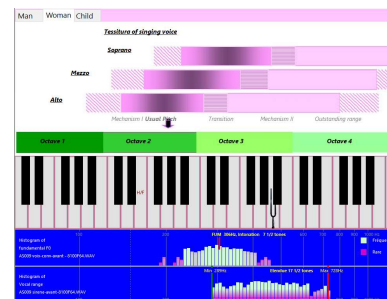
Assessment - Spectrogram



Assessment - Phonetogram (SRP)



Assessment - Maximum Phonation Time



Assessment - Vocal Range

Figure 3: Assessment module with 4 distinct tools

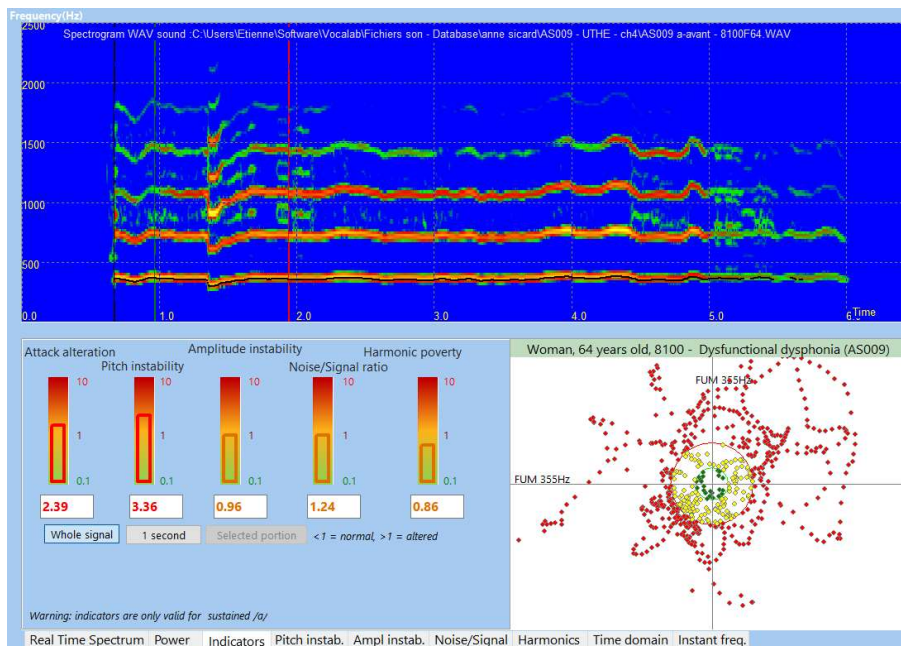


Figure 4: Voice perturbation measures calculated for a sustained /a:/ - patient AS009 before therapy.

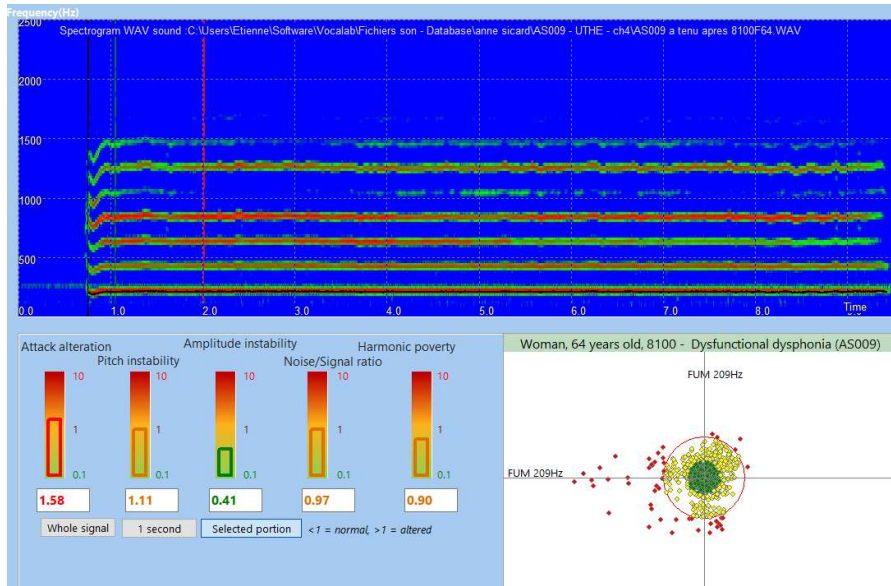


Figure 5: Voice perturbation measures calculated for a sustained /a:/ - patient AS009 after therapy.

Within the assessment tools, the spectrogram module calculates 5 main voice perturbation measures chosen to provide the clinician with the most pertinent information to inform future therapy (Figures 4 & 5). These are attack perturbation, pitch instability, amplitude instability, noise/signal ratio and harmonic poverty. Thresholds for 'normal' voice have all been standardised based on a sample of 1000 normal and pathological voices using expert ratings and patient self-assessment.

Values below 0.8 are considered as normal (green), while values above 1.5 may indicate pathology (red). In Figure 4, both the attack and the pitch are significantly altered, as well as the noise/harmonic ratio. The phonation time is very short. Significant improvements can be observed after therapy (Figure 5), with four measures at around or below 1.0, whilst the attack remains unstable.

## Therapy Module

The therapy module offers exercises for working on specific aspects of voice production with real-time visual feedback, through four distinct tools: breath control, pitch, articulation and timbre. These aspects are presented in a visually attractive way in order to appeal to both adults and children, and facilitate learning.

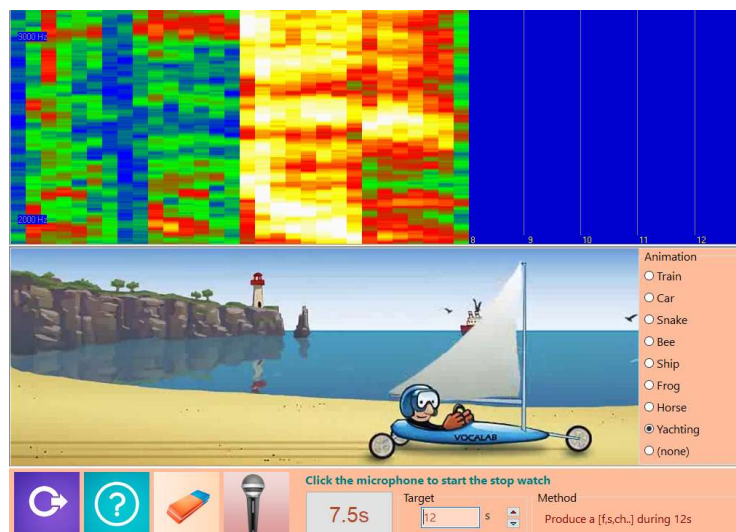


Figure 6: Breath control with real-time focus on high-frequency energy (top screen).

The breath control module offers a choice of different animations to represent breath flow which are particularly motivating for younger patients (Figure 6). Different colours show the amount of high-frequency energy. The therapist can choose the target duration (shown here as 12s sustained /f/) and the patient attempts to reach the desired target with significant constriction during the whole recording. The articulation module displays the sound waveform in real time, with the therapist model in red and the patient's production in blue. The therapist triggers the animation if progress is observed in the pronunciation of the desired target or if the patient achieves the target.

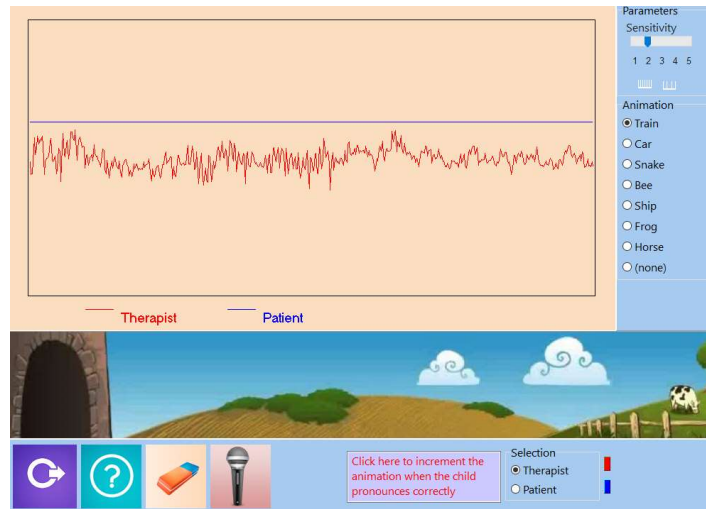


Figure 7: Comparing the therapist's model and the patient's reproduction.

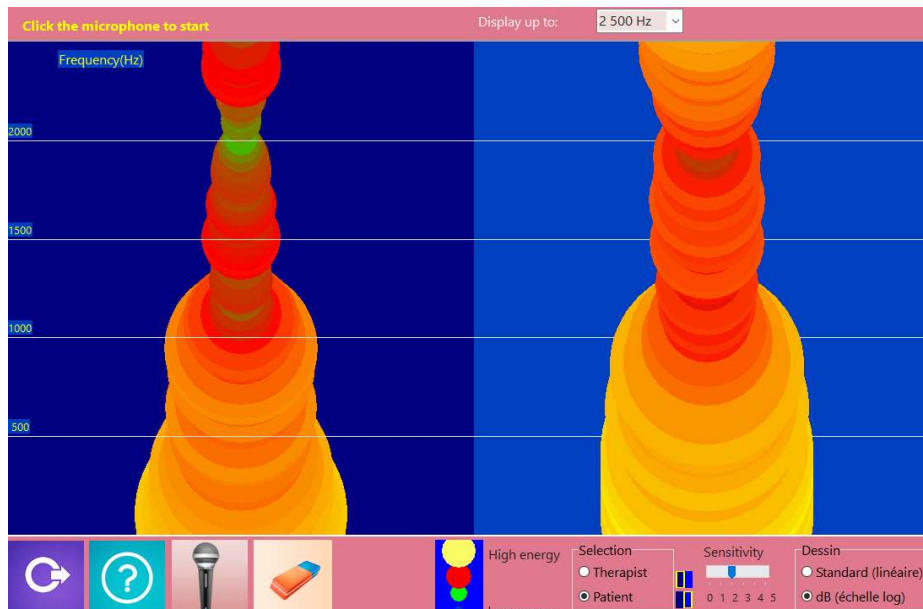


Figure 8: Real-time practice of the resonance on a sustained /a:/

The timbre module uses different coloured circles to represent the different formants and can be used to work on singing as well as the spoken voice (Figure 8). This module offers a wide range of possibilities for working on professional voice. When a sound is produced, circles appear for each harmonic. The

size and colour of the circles is linked to the harmonic energy. Two separate screens make it possible to contrast the therapist's model with the patient's production. The patient can visualise his ability to produce harmonics at high energy in real time instead of concentrating the energy at low frequency. This is highly effective exercise for working on harmonics in the voice, restoring resonance and improving speech efficiency. On the left-hand side of Figure 8, we can see the harmonics produced by a sustained /a:/ where most of the energy is concentrated at low frequencies, which limits the power of the voice and its ability to project. In contrast, the right-hand side shows the effects of efforts to transfer energy to the higher frequencies, which will have an immediate positive impact on the efficiency of the voice.

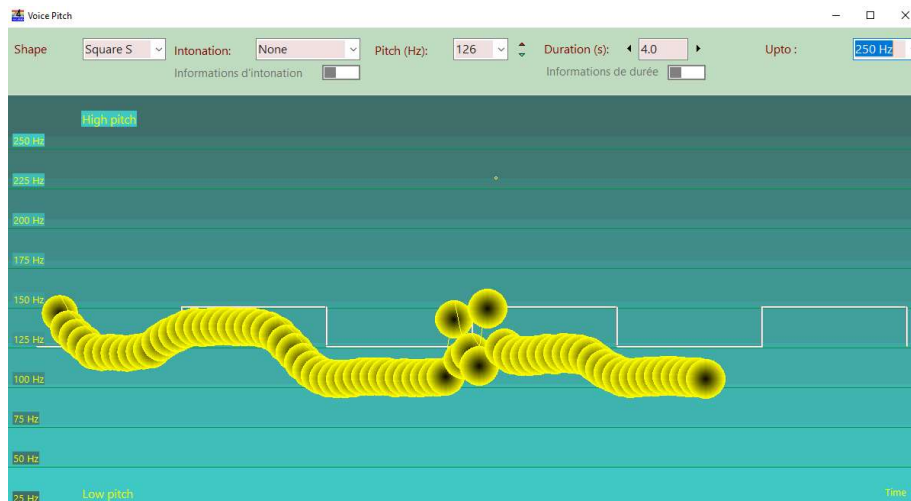


Figure 9: Real-time practice of pitch following a predefined profile

The pitch module offers an exercise where the patient attempts to reproduce a pitch profile chosen by the therapist or a predefined shape which can be adjusted in amplitude, duration and frequency according to the patient's ability. An example of a pitch profile with a square shape is shown in Figure 9. Note that the frequency has been adjusted to the male vocal range (here around 125 Hz). The patient is experimenting with the change of pitch by moving up and down in the voice and trying to follow the predefined pattern. There are a number of ready-made profiles with different shapes that can be used with the patient.

### Assessing progress

VOCALAB is also able to generate a pre and post-therapy report, based on 13 voice measures including sustained /a:/, glides, maximum phonation time for /a:/, /s/ and /z/, and the speech range profile. In Figure 10 on the right, the green circle denotes normal voice, and blue lines show the changes before and after therapy. [Sicard 2021] discusses 12 case studies with pre- and post-therapy reports assessing progress as a result of therapy.



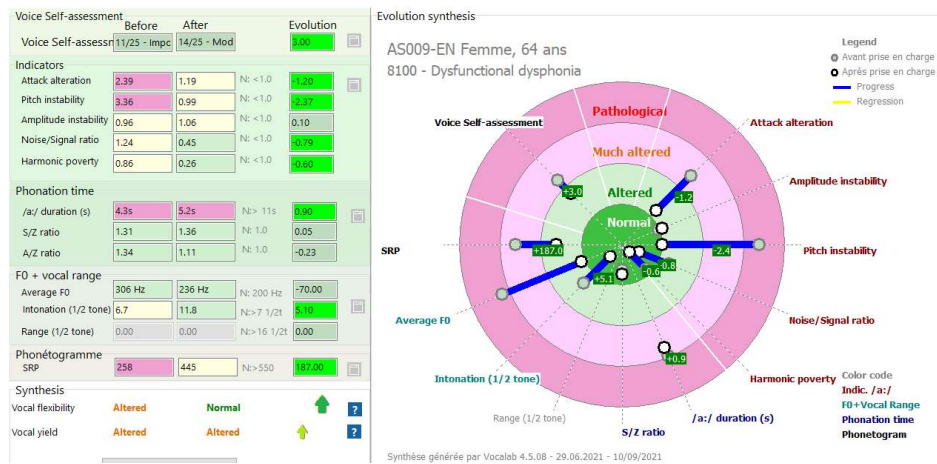


Figure 10: Real-time practice of pitch following a predefined profile

## Media Library

Finally, the media library module offers access to a large database of 350 normal and pathological voices, often with relevant case history information and pre- and post-therapy samples. These are an invaluable source for users wanting to deepen their knowledge of voice analysis as they can be opened within the spectrogram module, meaning that the spectral features of each type of voice can be clearly visualised and compared with the user's own recordings. There are also around 40 videos of normal and pathological voice with diagnostic information and suggestions for areas to target in therapy.

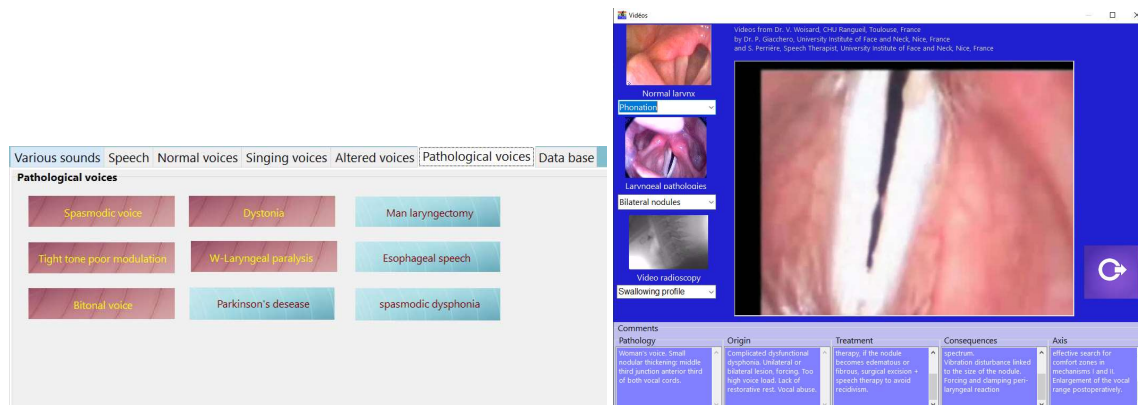


Figure 11: The media library includes 350 voice samples (sustained /a:/, glides). Around 40 videostroboscopy clips are also available

## Conclusion

In this paper, we described the VOCALAB software which can help patients understand their voice, and provide easy-to-understand voice measures at assessment and during speech therapy. The tool offers a wide variety of personalised therapy exercises to make voice work more tangible for the patient. We described the main modules and illustrated the key voice measures obtained from sustained vowels, glides and conversational speech. We also demonstrated the use of real-time biofeedback and therapist/patient comparisons to maximise client motivation throughout therapy. To date, the tool has been used by more than 3000 speech therapists in France and internationally.

## About the authors



Professor **Etienne Sicard** is currently a professor at INSA, an engineering school part of the University of Toulouse, France. He received a B.S degree and a PhD in Electrical Engineering from the University of Toulouse, France, in 1984

and 1987 respectively. Professor Sicard has authored or co-authored over 15 books, 10 commercial software packages and more than 200 technical papers in many areas of electrical engineering, including integrated circuit design automation and digital signal processing for speech therapy. He has developed best-seller software VOCALAB & DIADOLAB for voice & speech analysis applied to speech therapists. He is currently research director of a research team focused on pathological speech at LURCO Laboratories.

Site : [www.etienne-sicard.fr](http://www.etienne-sicard.fr)

Email : [etienne.sicard@insa-toulouse.fr](mailto:etienne.sicard@insa-toulouse.fr)



**Louise RUFFLE** is a speech and language therapist with specialist experience in voice disorders. She trained in the UK and spent a number of years working in London, including ENT departments at

Whipps Cross, Barts and the London and Whittington Hospitals. She has been based in the Paris region since 2015 and currently works in private practice in the Cergy-Pontoise area.

Louise is also a keen singer and has sung semi-professionally in a number of vocal ensembles, including Clare College Choir.

Email : [ruffle.orthophonie@gmail.com](mailto:ruffle.orthophonie@gmail.com)



**Anne MENIN-SICARD** is a speech therapist and received a master's degree in Language Sciences from the University of Grenoble, France. She was a speech & language therapist in Toulouse from 1994 to 2015.

Anne MENIN-SICARD is co-author of VOCALAB and DIADOLAB software. As part of AMS Formation, she organizes and conducts training for speech therapists on the assessment and management of pathological voice and speech. She is proposing personalized online training for speech therapists. As a researcher in ERU 46 team at LURCO laboratories, she is developing tools & methodologies for objective measurement and speech therapy. She is the author of two books on the objective assessment of the voice & speech as well as numerous publications in the field of speech therapy.

Site : <https://www.formationsvoixparole.fr>

Email : [anne.sicard2@orange.fr](mailto:anne.sicard2@orange.fr)

## References

[Baken 2000] Baken, R.J., Orlikoff, R.F. (2000). Clinical measurement of speech and voice (2nd ed.). San Diego, CA: Singular.

[Barbera & Frenay 2016] Barbera, O. & Frenay, C. (2016). Etude de l'évolution de la voix au cours du vieillissement. Master in Speech Therapy, University of Lyon. (in French)

[Boersma 2021] Boersma, Paul & Weenink, David (2021). Praat: doing phonetics by computer [Computer program]. Version 6.1.53, retrieved 8 September 2021 from <http://www.praat.org/>

[GERIP 2021] <https://boutiquesante.gerip.com/en/home/72-pack-vocalab-3.html>

[Kent 2000] Kent, R.D., Ball, M.J. (2000). Voice Quality Measurement. San Diego, CA: Singular Publishing Group.

[Lurco 2021] <https://www.unadreo.org/lurco/presentation-du-lurco/> (In French)

[Maryn 2009] Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality: A meta-analysis. *The Journal of the Acoustical Society of America*, 126(5), 2619-2634.

[Maryn 2010] 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(5), 540-555.

[Sicard 2013] Sicard, E., Perriere, S., & Menin-Sicard, A. (2013). Développement et validation d'outils de mesures de la qualité de la voix dans le logiciel VOCALAB. *Glossa*, (113), pp-63. (In French)

[Sicard 2019] Sicard, E., Meyrieux, L., Moreau, M., & Remacle, A. (2019). L'analyse acoustique des voix d'enfants de 5 ans: Proposition de valeurs de référence pour les logiciels PRAAT et VOCALAB.

[Sicard 2021] Menin-Sicard, A., & Sicard, E. (2021). Etude de cas de voix dans le cadre de la prise en charge orthophonique. Research Report LURCO, INSA Toulouse. (In French)

[Verdolini 2006] Verdolini, K., Rosen, C.A., Branski, R.C. (2006). *Classification manual for voice disorders—I*. Mahwah, NJ: Lawrence Erlbaum.