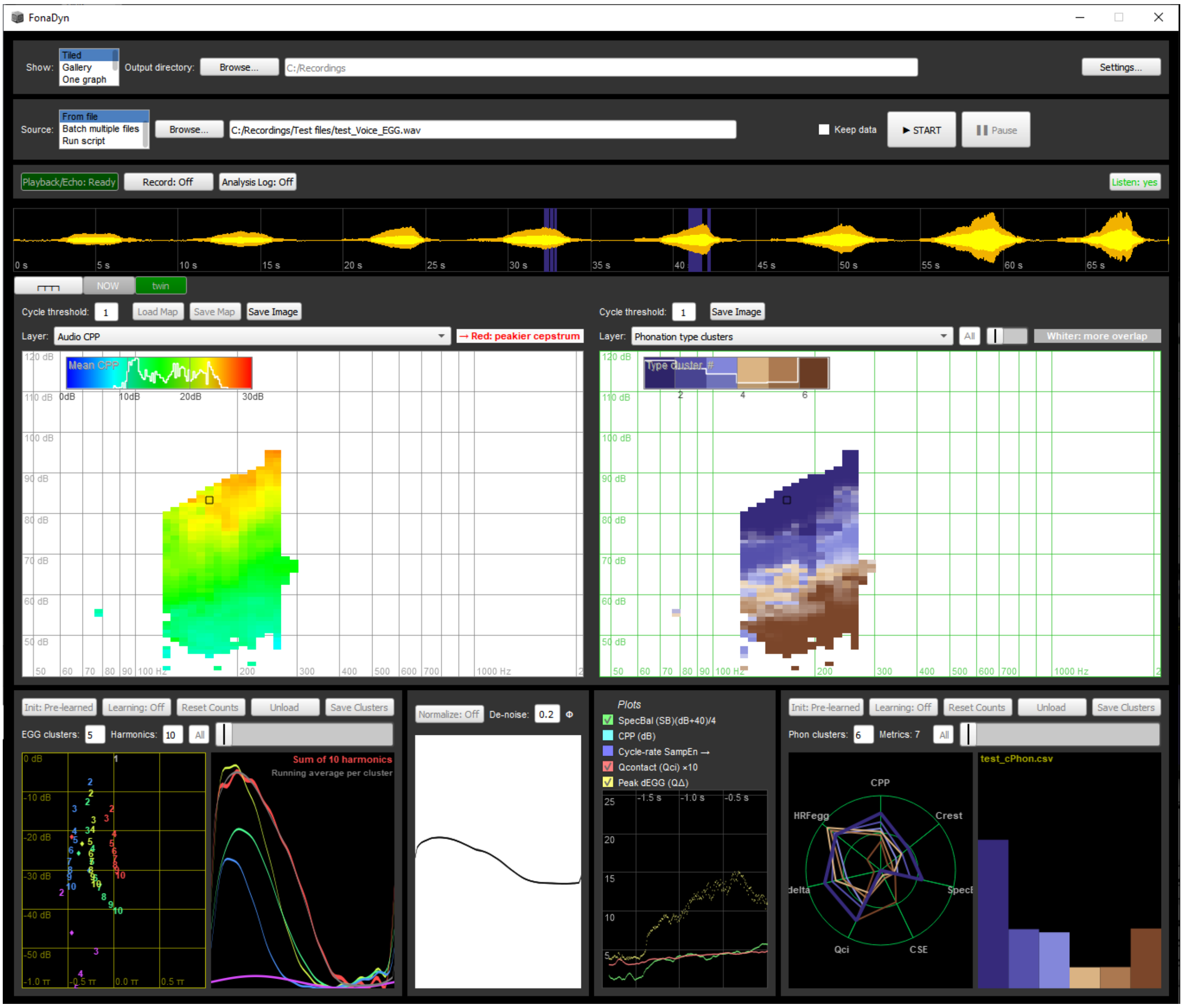
A quick look at FonaDyn

Sten Ternström, Division of Speech, Music & Hearing, KTH Royal Institute of Technology, Sweden, stern@kth.se



The 'maximal' FonaDyn user interface, with most options made visible. Sub-panels can be displayed or hidden as desired. In most cases, only one or two panels are needed at a time. There are is no window menu; most controls are visible all the time.

Abstract

The voice is notoriously variable, and conventional measurement paradigms are weak in terms of providing evidence for effects of treatment and/or training of voices. New methods are needed that can take into account the variability of scalar metrics across the voice range. The voice map is a generalization of the phonetogram, with the potential to be used in many ways, for teaching, training, therapy and research. FonaDyn is primarily a proof-of concept workbench for research on phonation and for exploring and validating the voice-mapping measurement paradigm.

Introduction

Since 2015, FonaDyn has evolved from an EGG analyzer into a more general workbench for voice mapping and measurement, with the ultimate goal of improving the evidential value of quantitative metrics of voice. For every phonated cycle, it analyses many signal attributes over all or part of the voice range, with real-time visual feedback that visualizes the great variability that exists within and between individual voices. Phonatory regimes such as modal/falsetto and other types of phonatory differences can be discriminated by clustering. FonaDyn can be used to pursue many kinds of research and teaching, on phonation, voice source dynamics, source-filter interaction and effects of therapy and/or training. It can be used interactively, or with scripted batches, or as a data-acquisition front-end. It has rich graphics and file export to Matlab® etc. Maps can be compared pre/post intervention. The only constraint is that the phonation must be reasonably periodic in at least some part of the voice range.

Design criteria

- Because the voice is so variable, it is important to collect as much data as possible, in as short a time as possible.
- Real-time visual feedback is essential for establishing our cognitive connection between the sound and the metrics.
- Accurate SPL calibration remains crucial, and must be facilitated.
- A single software system cannot do everything. Easy data exchange with other software is important.

Main features

FonaDyn records voice and EGG signals in parallel, and visualises in real time a large number of metrics in the voice field, with semitones or Hz on the horizontal axis and decibels on the vertical axis.

- Signals can be analysed live, or from signal files recorded with FonaDyn or other software.
- Voice maps display in real time, with multiple layers: density, clarity, crest factor, spectral balance, cepstral peak prominence (CPP), EGG contact quotient, EGG peak derivative, EGG index of contacting, EGG harmonic richness factor, and EGG cycle-rate sample entropy (at this writing).
- Features of the acoustic and EGG signals can be automatically clustered, to reveal regions of different phonatory regimes. Clusters are 'learned' in real time, then frozen for classification of similar signals. Or, cluster centroids can be computed elsewhere and imported into FonaDyn. Phonation types can also be specified interactively.
- The analysis is cycle-synchronous wherever possible. FonaDyn analyses every phonatory cycle, whenever the acoustic signal exceeds an adjustable periodicity threshold.
- There are real-time displays of the voice map, the EGG waveform, time contours of selected metrics, cluster centroids and EGG wave shapes.
- Multiple voice maps can be displayed simultaneously, including *difference maps* for assessment of changes across interventions.
- Voice maps and clusters data are saved to and loaded from ordinary CSV files, for simple exchange with Excel, Matlab, etc. for further analysis and customized visualizations. A set of introductory Matlab® m-files is provided.
- A multichannel log file with a time track and all metrics, for every phonatory cycle, can be created for detailed analysis.
- Click on a voice map and listen to the vocalizations that correspond to that place in the map, with the waveform displayed and selectable.
- Noise in EGG signals can be substantially reduced, on the fly, using thresholding in the frequency domain.
- Arbitrary extra signal channels can be recorded synchronously, at both audio and physiological rates.
- Batch operations are supported by a bespoke scripting mechanism. The analyses must still run in real time (not faster).
- For SPL calibration, an integrated 'wizard' supports four different calibration scenarios, depending on your equipment.
- Supports reading of most common signal file formats with at least two channels: voice and EGG. 44.1 kHz sampling rate only; 24-bit recording as default.
- The source code is provided, and editable, in the freeware SuperCollider development environment, which has to be installed first.
- Comes with a 90+ page FonaDyn Handbook, with sections for Setup, Theory and User Manual, plus other supporting materials.

Availability

FonaDyn is freeware, in the public domain, subject to the EUPL 1.2 licensing terms. The most up-to-date version can be freely downloaded from www.kth.se/profile/stern/, where you will also find links to several media presentations on voice mapping. The system is in constant development, so you are encouraged to read the Release Notes every time you update. FonaDyn runs on platforms supported by SuperCollider, including Apple MacOS and Microsoft Windows. Linux is possible with recompilation.

Acknowledgments

Voice mapping is a data acquisition and visualization paradigm that emerged initially from the phonetogram, most notably from the work of Peter Pabon on computerized voice profiling, since the 1980's. Ternström's doctoral students Anick Lamarche, Andreas Selamtzis, Peter Pabon and Huanchen Cai have all worked with voice mapping in different ways. The code foundation for FonaDyn was laid in 2015 by Dennis Johansson, for his M.Sc. thesis in Computer Science. Funding has been had from doctoral stipend organizations, from Swedish Research Council (VR) projects in 2010-2018, and KTH faculty funding.

Results?

In addition to the doctoral theses of the aforementioned students, which all explore voice mapping in different ways, the following publications give examples of research done at this writing. OA = Open Access.

Position paper: Ternström, S., and Pabon, P. (**2022**). "Voice Maps as a Tool for Understanding and Dealing with Variability in the Voice," *Appl. Sci.,* **12**, 11353. doi:10.3390/app122211353. OA.

Cai, H., and Ternström, S. (**2022**). "Mapping Phonation Types by Clustering of Multiple Metrics," *Appl. Sci.*, **12**, 12092. doi:10.3390/app122312092. OA.

Patel, R. R., and Ternström, S. (**2021**). "Quantitative and Qualitative Electroglottographic Wave Shape Differences in Children and Adults Using Voice Map–Based Analysis," *J. Speech, Lang. Hear. Res.*, **64**, 2977–2995. doi:10.1044/2021_jslhr-20-00717.

Lã, F. M. B., and Ternström, S. (**2020**). "Flow ball-assisted voice training: Immediate effects on vocal fold contacting," *Biomed. Signal Process. Control*, **62**, 102064. doi:10.1016/j.bspc.2020.102064.

Ternström, S. (**2019**). "Normalized time-domain parameters for electroglottographic waveforms," *J. Acoust. Soc. Am.*, **146**, EL65–EL70. doi:10.1121/1.5117174. OA.

Pabon, P., and Ternström, S. (**2018**). "Feature Maps of the Acoustic Spectrum of the Voice," *J. Voice*, **34**, 161.e1-161.e26. doi:10.1016/j.jvoice.2018.08.014. OA.

Ternström, S., Johansson, D., and Selamtzis, A. (**2018**). "FonaDyn — A system for real-time analysis of the electroglottogram, over the voice range," *SoftwareX*, **7**, 74–80. doi:10.1016/j.softx.2018.03.002

Ternström, S., D'Amario, S., and Selamtzis, A. (**2018**). "Effects of the Lung Volume on the Electroglottographic Waveform in Trained Female Singers," *J. Voice*, **34**, 485.e1-485.e21. OA. doi:10.1016/j.jvoice.2018.09.006