

The voice-mapping system FonaDyn – overview and demo

Sten Ternström

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

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, a generalization of the phonetogram, offers a frame of reference that can be used in many ways, for research and in the clinic. FonaDyn is a proof-of-concept workbench that we are developing in order to explore and validate the mapping measurement paradigm. In this demo, you can try FonaDyn, to visualize and measure your own phonation faster and in greater detail than ever before.

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 analyzes 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. Phonation types such as modal/falsetto, or any 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, with scripted batches, or as a data-acquisition front-end with extensive preview facilities and file export to Matlab etc. Maps can be compared pre/post intervention. The only constraint is that the phonation must be reasonably periodic throughout.

Design criteria

- Because of the voice is so variable, it is important to collect as much data as possible. FonaDyn analyses every phonatory cycle, when the phonation exceeds an adjustable stability threshold.
- Real-time feedback is essential for establishing our cognitive connection between the sound and the metrics. FonaDyn's displays have negligible visual latency.
- Accurate SPL calibration is essential, and must be as painless and portable as possible. FonaDyn can guide the user to a uniform level calibration scaling in all signal files.
- A single software system cannot do everything. Easy data exchange with other software is essential.

Main features

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

- Signals can be analysed live or pre-recorded.
- Voice maps have the layers density, clarity, crest factor, spectral balance, cepstral peak prominence

smoothed (CPPS), EGG contact quotient, EGG peak derivative, EGG index of contacting, EGG cycle-rate sample entropy, and clustered EGG features.

- The analysis is cycle-synchronous wherever possible.
- EGG waveshapes are automatically clustered, to reveal regions of different phonatory regimes.
- Real-time displays of the voice map, the EGG waveform, time contours of selected metrics, and EGG cluster centroids and waveshapes.
- Multiple voice maps and *difference maps* can be displayed simultaneously, for assessment of changes across interventions.
- Voice maps and clustering data are saved to and loaded from ordinary CSV files, for simple exchange with Excel, Matlab, etc. for further analysis and customized visualizations.
- A log file in .aiff floating-point, with a time track and all metrics for every phonatory cycle, can be created for detailed analysis.
- In conjunction with a log file, FonaDyn lets you point in a voice map and listen to the sounds that correspond to that place in the map.
- Extra signal channels can be recorded in synchrony at both audio and physiological rates.
- Batch operations can be executed using a simple scripting mechanism.
- The source code is provided and editable in the SuperCollider development environment, which needs to be installed first.
- 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.)
- Comes with a 70+ page Handbook with sections for Setup, Theory and User Manual.

Availability

FonaDyn runs on both Windows and Mac. Linux is possible with recompilation. 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 at <https://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 it is important to read the Release Notes every time you update.

Acknowledgments

Voice mapping is a technique that has emerged initially from the phonetogram, most notably from the work of Peter Pabon since the 1980's on computerized voice profiling. S.T.'s doctoral students Anick Lamarche, Andreas Selamtzis, Peter Pabon and currently Huanchen Cai have

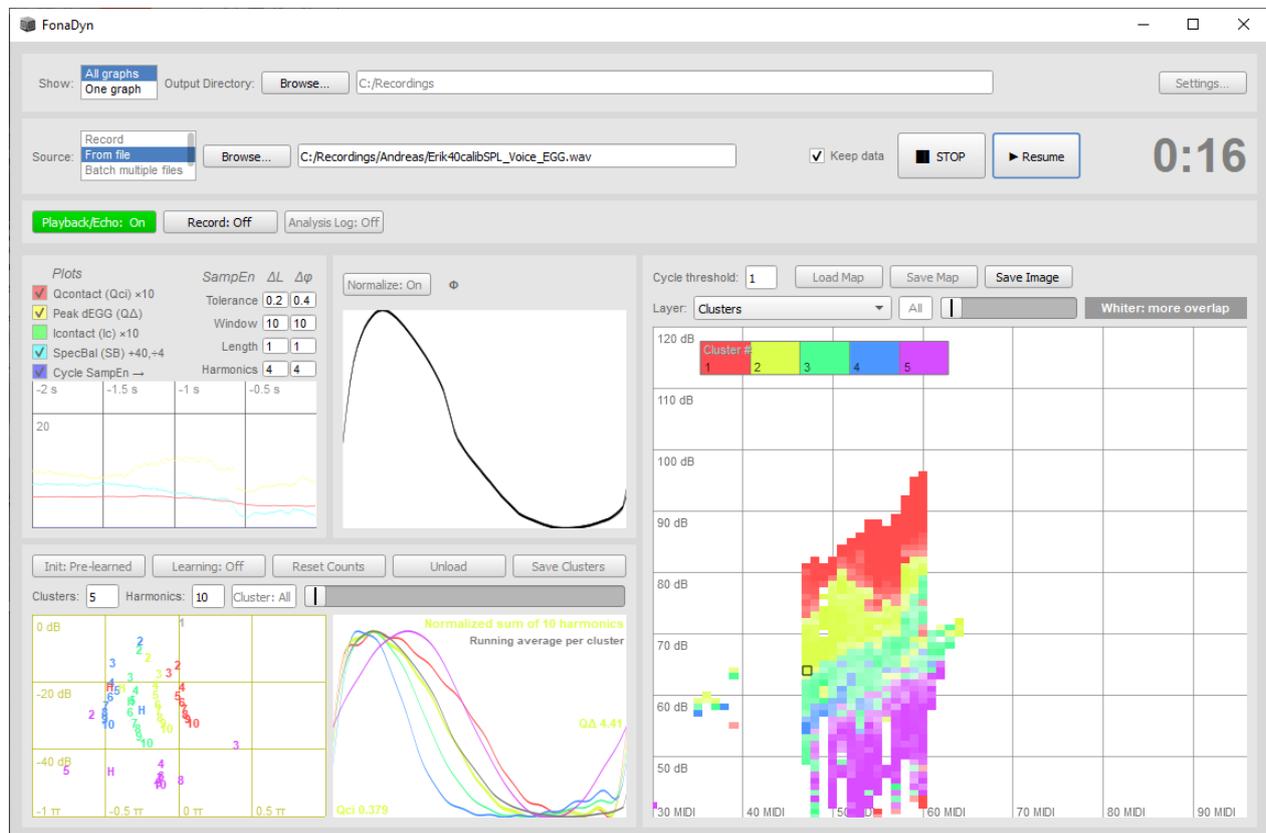
all worked with voice mapping in different ways. The foundation of FonaDyn was laid in 2015 by Dennis Johansson for his M.Sc. thesis. Funding has been had from their doctoral stipend organizations, from VR projects in 2010-2018, and KTH faculty funding.

References

In addition to the doctoral theses of the aforementioned students, which document the foundational work that has led to FonaDyn, the following publications give examples of research hitherto done using voice mapping.

- 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
- 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

- 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
- Ternström, S. (2019). “Normalized time-domain parameters for electroglottographic waveforms,” *J. Acoust. Soc. Am.*, **146**, EL65–EL70. doi:10.1121/1.51117174
- 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. doi:10.1016/j.jvoice.2018.09.006



The FonaDyn user interface. There are no menus, most controls are visible all the time. Subpanels can be displayed or hidden as desired.