FonaDyn v3.1.0 release notes

These release notes are longer than usual. They summarize the many *cumulative* changes that have been made since the year 2019, FonaDyn version 2.0.2, which till now was the version stored at the published repository at the journal *SoftwareX*. An updating article has been submitted. The latest public version of FonaDyn is always available at https://www.kth.se/profile/stern/ .

FonaDyn version 3.1 is a major new release with several significant new features. Most prominent among these are the addition of a second mode of clustering, of phonation types. Also, this release introduces smoothing/interpolation of voice maps, listening from maps, expanded multi-channel support, better layout control, optional 140 dB max SPL, a built-in level calibration tool, and 24-bit recording as default. Several usability aspects have been improved, with more checking for consistent settings, etc. Under the hood, much of the code has been further optimized and restructured. A terse description of all 'revision commits' can be found in the source code file VRPMain.sc.

IMPORTANT: SuperCollider version

Version 3.13.0 of SuperCollider, released in February 2023, contained some breaking changes and new features. The present FonaDyn version 3.1.0 relies on a few of these changes, and therefore requires SuperCollider 3.13.0 or later. If you need to update SC and/or FonaDyn, please follow the instructions in the file UPDATING.txt in the distribution ZIP file.

On Apple MacOS, when running FonaDyn for the first time, the computer must be connected to the Internet, or Apple's safety checks will not work.

New features

Multiple maps, map smoothing, and difference maps: For comparing voice maps, it is now possible to display up to four extra maps at the same time as the primary map (also called "Now"):

- "Twin" which shows and updates the same map as "Now", but of some other metric;
- "Before" which shows a previously saved map without updating it;
- "Diff" which computes and displays the differences between the maps "Now" and "Before";
- "Smooth", which is an interpolated and smoothed version of the first map.

Multiple maps can be stacked on top of each other, or can be tiled within the borders of the FonaDyn main window.

With FonaDyn 3.1, difference maps can be made also of the layers for clustered data.

SPL calibration tool: there is now a dedicated walk-through tool for SPL calibration. The texts and photos in the SPL Calibration Tool are intended to support a consistent calibration procedure. This tool also shows level meters for all inputs and outputs, a multi-trace oscilloscope, and a spectrum analyzer, so it is very useful for visualizing signal levels and signal quality.

EGG noise suppression: the conditioning of the EGG signal has been reworked. The 100 Hz HP filter is still in place, but the 10 kHz brickwall LP filter and the median filter have been removed. Instead, a spectral threshold is implemented, with 4:1 dynamics expanders on all spectral bins whose levels are below the threshold (Handbook section 3.7.1). This greatly increases the usefulness of the metrics Q_{Δ} and I_c . The threshold may need to be adjusted for the noise inherent in your particular system. To some extent, it can clean up noisy EGG recordings. Use it with moderation. If the threshold is set too high, the EGG signal will start to lose not only noise, but also its weakest (highest) harmonics.

Listen from maps: FonaDyn can now play back those sounds that contributed to a particular place in a voice map. This feature requires the creation of a _Log.aiff file that is consistent with the currently loaded signal file and voice map, in the same folder. If one exists, a horizontal strip called the “Signal window” opens to show the envelope of the voice signal. Left-click in a map to select the relevant parts of the signal; shift-left-click to also hear them. The segments highlighted in the signal can also be clicked for listening.

Clustering of phonation types: combinations of audio and EGG metrics can be clustered, using a mechanism similar to that for EGG wave-shape clustering. This can be used to map complex voice properties such as pressedness, breathiness etc. Centroids are displayed as radar plots. The metrics to be used for clustering, as well as their normalizations, are user-selectable, by pre-editing a centroids file *_cPhon.csv . Phonation-type centroids can also be picked manually from a voice map, guided by listening, and then used for classification.

Script control: there is now a facility for initializing and running FonaDyn under the control of a text script. Since FonaDyn can only run in real time, processing large sets of many files had to be done manually, which takes time and patience, is prone to errors, and is hard to document. Now you can create a bespoke script text file containing instructions for those settings and actions that can be relevant for batch jobs. A script will process one _Voice_EGG.wav file at a time, and can accept new settings and save results before advancing to the next file. The progress of the script is logged in the post window, which can be copied and saved for documentation. The text script itself can be written manually, but long scripts are best generated by some other program, such as Matlab or MS Excel. Scripts can also be used to set up an initial state in FonaDyn, without starting an analysis. The script syntax is based on SClang and is specific to FonaDyn; it is not that of an OS shell script.

Auto Reset of cluster data: the Reset Counts button can now be pressed while FonaDyn is stopped. This ‘arms’ FonaDyn to initialize the clustering automatically, once a short period (≈200 ms) of continuously stable phonation has been detected. This feature improves consistency across repeated analyses, and reduces the probability of generating outlier clusters.

Drag-and-drop: Files with file names ending in ‘_Voice_EGG.wav’ can be dragged from outside FonaDyn and dropped on the field for the input file name(s). Files with file names ending in ‘_cEGG.csv’ or ‘_cPhon.csv’ can be dropped onto the corresponding ‘Load Clusters’ button. Files with file names ending in ‘_VRP.csv’ can be dropped onto the ‘Load Map’ button. This is often quicker than navigating through the File Open dialog boxes.

Metric distributions: the **colour bar** on voice maps now also displays a relative **distribution of the mapped metric** on its own axis, as counted by the number of cells in the map. The area under the white curve corresponds to the total number of occupied cells in the map. To reconstruct such a graph outside of FonaDyn, make a histogram of the corresponding column in the _VRP.csv file.

New audio metric: a layer for the **Spectrum Balance** (SB) of the audio signal has been added to the voice map. In FonaDyn, SB is computed as the level difference between a high- and a low-frequency band. Two filters divide the voice spectrum, one -24 dB/octave low-pass at 1500 Hz and one +24 dB/octave high-pass at 2000 Hz. The RMS amplitude of each filter output signal is computed and smoothed with a first-order low-pass filter at 50 Hz. Hence the SB metric is not cycle-synchronous. The resulting amplitudes are converted to levels in dB, and the high level minus the low level becomes the Spectrum Balance. Since there is usually less energy in the high band, SB is most often negative. This is by design. With increasing voice effort, SB typically increases (becomes less

negative). The SB average in each cell is the average of this level difference (in dB); it is not the ratio of the averaged powers in the two bands.

> **New in v3.0.5**: the internal synchronization between SB and EGG has been adjusted so as to improve the rejection of unvoiced fricatives in speech. It is still not perfect: *voiced* fricatives are *not* rejected, and will affect the SB metric.

New audio metric: a layer for the **Cepstral Peak Prominence** (CPP) of the audio signal has been added to the voice map. This metric has many analysis settings, so what is hoped to be a reasonable set of defaults has been chosen. These settings are described in the FonaDyn Handbook.

> **From v3.0.5, FonaDyn no longer smooths the CPP**, neither across time frames or across quefrency bins. Since many readings are averaged into each map cell, the maps appear quite even anyway, and the resulting CPP values are more consistent with the literature. A simple change of a file name can reinstate the CPP smoothing. Please contact us if you wish to know how to do this.
>
> **From v3.0.7, the CPP** gives a value that is about 1 dB higher than before, in strong voice. This is because of a small DSP optimization, involving a reduced dither noise. The new value is more correct.

New EGG metric: the **EGG Harmonic Richness Factor** (HRFegg). This metric is essentially an alternative to the Q_{Δ} as an indicator of vocal fold contacting. It is less sensitive to signal noise, but outside FonaDyn it would be more computationally intensive. HRFegg rises steeply when contacting occurs, and reaches its maximum for very short episodes of contacting, i.e. for small values of the contact quotient. Once minimum VF contacting is established, HRFegg *decreases* as the voice gets louder. HRFegg is not written to Log files, because the information is implicit in the harmonic levels, and can be recomputed e.g. in Matlab.

Input channels: FonaDyn can now analyze signal files with more than two channels. As before, the voice signal is expected in the first channel, and the EGG signal in the second channel. Any remaining channels are ignored (but can be useful when analyzing with other software). FonaDyn can also *record* extra audio-rate signals as additional channels into the `_Voice_EGG.wav` files. This makes it a lot more useful for experimental setups.

Startup file configuration: The new `FonaDyn.config(…)` statement can be given once or several times in the SuperCollider startup file. See the FonaDyn Handbook section 1.2.8 for the syntax for these options.

> **Record in 16 bits**: FonaDyn now by default records into 24-bit files rather than 16-bit files. FonaDyn still reads 16-bit files as it has always done. For backward compatibility with other audio applications, you can still configure explicitly for 16-bit recording.
>
> **Singer mode**: FonaDyn can now support a maximum SPL of 140 dB, for recording very loud voices.
>
> **Selectable audio inputs**: While the default inputs are still 0 and 1, these can be customized in the startup file, or in the **Settings…** dialog.
>
> **Constrain the voice map aspect ratio**: a 2:1 aspect ratio is often recommended as a standard, although it may waste space on the screen. To display voice maps with a fixed 2:1 aspect ratio, add `FonaDyn.config(fixedAspectRatio: true)` to the startup file.

Choose vertical tiling of maps: sometimes it is more convenient in the mode Show: Tiled to have the maps one above the other, instead of side by side.

Auto-initialize your FonaDyn sessions with a script: create a text script as described in the Handbook, and specify its pathname in double quotes, with forward slashes / :
`FonaDyn.config(runScript: "D:/full/path/to/scriptfile.txt")`

The **Settings…** dialog box:

- has a new **Record inputs** field for specifying optional audio rate channels.
- has a new popup menu for choosing the sample rate for Extra files.
- has a new option for playback of the EGG signal on the second (right channel) output. Listening can help in diagnosing EGG signal problems. The default is to play the Voice signal on both left and right outputs. When checked, the EGG signal too will be displayed in the Listening waveform window.
- has a new option: "Save Map also saves a context script" for restoring the context of settings in which the latest map was made.
- has a new option: "Suppress Gibbs' ringing in resynthesized EGG shapes". This affects only how the clustered wave shapes are displayed, not how they are saved.

Matlab integration: The **new Matlab file** "runDemos.m" plots a few pages of example plots, using the files supplied in the folder "Test files".

A new mechanism is provided for *patching* Matlab® m-files from within FonaDyn. This mechanism uses tagged template m-files into which the SClang compiler can embed FonaDyn's current settings. Typically this is used for matching the colours displayed by Matlab to those used in FonaDyn. This mechanism can be exploited by advanced users for closer SuperCollider-to-Matlab integration, and not only with FonaDyn. This new feature requires an additional folder "matlab" to be copied during installation from the ZIP file to the *userAppSupportDir* (not to *Extensions*). See readme-files therein, and the online documentation for the new class **MfSC**.

Changed behaviours and formats

Because there are now **two kinds of clusters**, the cluster data files have been renamed *_cEGG.csv and *_cPhon.csv. The earlier name for EGG clusters ('_clusters.csv') and the beta name for phonation type clusters ('_phonclusters.csv') are still accepted, but not recommended.

All CSV files (*_cEGG.csv, *_cPhon.csv and *_VRP.csv) can have either semicolon or comma as the column separator – but not mixed in the same file. FonaDyn will attempt first to parse columns using the semicolon, and if that gives unexpected results, it will try the comma instead. This can be helpful when editing and saving CSV files using spreadsheet apps. FonaDyn always saves CSV files with the semicolon as the column separator.

EGG: the HF residual has been replaced by the level of the fundamental. In EGG analysis, FonaDyn now uses the average level of the fundamental to reconstruct clustered EGG wave shapes with their correct relative amplitudes. To see these non-normalized shapes in the EGG centroids display panel, choose Normalize: Off in the Moving EGG panel. The symbol 'H' in the EGG cluster centroid display, which stood for the residual power in higher un-tracked harmonics, has been replaced by a diamond '◊' marker. This marker displays the level of the EGG fundamental partial relative to full scale, and its phase relative to the cycle segmentation trigger (the latter is as before). These properties of the fundamental are internally scaled such that they do not influence the resulting clustering of the EGG wave shapes. The reasons for making this change emerged with experience: the HF residual is rarely used; and at low voice effort levels, it is sensitive to system noise. It was therefore not a very

suitable feature for clustering. Although the EGG waveshapes clustered by earlier versions will not change noticeably, the cluster extents on voice maps may change a little. All signals within a project should be reclustered with FonaDyn version 3.1; or, stay with your old version until the project is finished. The content of _cEGG.csv files has changed, in that the old HF residual column is now instead used for the level of the fundamental partial. FonaDyn will detect this, and will issue a warning when an old file is opened.

The notion of **voice map "layers"** is now made more explicit. In FonaDyn terminology, a map is a single entity made up of several layers. It is stored as a matrix in a _VRP.csv file, with one column per layer and one row per cell. This is as before.

For the new metrics, files of type `_VRP.csv` now contain **several additional columns**: "SpecBal" for the average SB, "CPP" for the average CPP, "HRFegg" for the average EGG harmonic richness factor, "maxCphon" for the number of the dominant phonation-type cluster in a cell, and "cPhon *n*" for the cycle counts for each phonation-type cluster. Old `_VRP.csv` files are still parsed and shown correctly; although the layers for the new metrics and clusters will display as empty. If you have written software that reads or creates such files, that software may need to be revised. The number of columns in `_VRP.csv` files is now 14 + *nClustersEGG* + *nClustersPhon*.

The map layers for individual clusters used to display cycle counts, like the Density layer does. Instead, they now display in each cell **the percentage of the total cycle count** in that cell which pertains to the chosen cluster. This is in line with the philosophy that the cycle count as such is seldom informative. However, the relative occurrence of a cluster in a given cell is often of interest. In difference maps, the change in the percentage is mapped. In _VRP.csv files, though, only the cycle counts are stored. When a file is loaded, the percentages are recomputed relative to the column 'Total'.

In _VRP.csv files, the columns representing the different layers of the map can now appear in any order. This can be useful if you want to insert extra columns of your own. FonaDyn will ignore such columns, on Load Map, and they will not be saved by Save Map. The clusters represented in the file must nevertheless form contiguous sets (1...*N*), for both EGG and PhonType clusters; for instance, if there are five EGG clusters, then columns "Cluster 1" through "Cluster 5" must all be present, if not in order. When using _VRP.csv files with the supplied Matlab® routines, the cluster columns must however be contiguous and in the order in which FonaDyn writes them.

_VRP.csv files are now saved with a reasonable number of decimals, rather than with the full 14-decimals precision. This saves a little space, and is more readable. Cluster centroid files are still saved with full precision. The decimal character is still always the period '.' .

WAV signal files containing **floating-point data** have changed calibration. FonaDyn used to assume ±1.0 to mean 120 dB peak. Now the value 1.0 is assumed to represent 1 Pa instantaneous pressure, since that is the SI-metric value, and also how it is often written by other software, such as *Sopran*. This calibration is relevant only for the first channel of signal files. A peak amplitude of 1 Pa corresponds to 91 dB RMS re 20 µPa, for a sine wave. (Note: floating-point WAVs are still unusual. The preceding does not apply to WAV files with fixed-point data, which are more common.)

Save image: press 'F' to invoke the Save File dialog box, or ESC to close the image window. Press 'C' to close all open image windows.

Log files: if Learning is **On** when a Log file is written, a warning is issued that the cluster numbers in the file will be inconsistent. Their centroids will drift over the course of the recording, as FonaDyn "learns". This will matter only if you are using the clustering functionality.

In the **Source** mode **"Live signals"** (formerly **"Record"**), that is used for recording, a **Calibrate…** button appears, as a short-cut to the SPL calibration tool.

The **Show** layout options "All graphs" and "One graph" have been extended to "All tiled", "All gallery" and "One graph". The "All gallery" layout makes better use of the screen space when multiple maps are displayed. There are also the options "Tiled" and "Gallery", which remember the chosen visibilities of the graphs. Pressing Enter on this list restores the visibility of all graphs.

You can now right-click to see **a popup-menu with the layout options** and their shortcut keys. This works anywhere in the main window, except in those panels where a right-click means something else.

The calculations of the **audio SPL** and the **audio crest factor** are now cycle-synchronous, which improves the resolution in timing, and practically eliminates such 'drooping', as earlier could be caused by level-smoothing filters. You may still see drooping at phonation offsets, but it is caused only by the voice itself. This means that maps remade with this version will be very similar but probably not identical to maps made with versions prior to 2.1.5.

The `_Log.aiff` file type now contains additional tracks for the spectrum balance (SB), the CPP (CPP), and the phonation-type cluster number. The remaining tracks have been moved up. If you have software that reads these files, it will need to be revised. The number of tracks in `_Log.aiff` files is now 12 + 2*(*nHarmonics*+1). A track list is given in the SuperCollider documentation for the class `VRPViewMainMenuOutput`.

The Entropy (CSE) layer of the voice map now contains in each cell the *mean* of the cycle-rate sample entropy. It used to display the maximum.

On Apple MacOS, both ARM and Intel hardwares are now supported by the FonaDyn plugins.

The post window now logs more verbosely what is going on.

Selected bug fixes

Closing FonaDyn while a playback or analysis was in progress would cause erratic behaviour. Of course, you should not close while something is running, but FonaDyn now copes with it.

The Sclang interpreter would sometimes crash after running FonaDyn. We believe that this has now been fixed.

In v3.0.4d, a bug was introduced that could cause audible glitches when listening from maps. This has been fixed.

Loading cluster data files _cEGG.csv and _cPhon.csv from a script would sometimes change the settings of the cluster panel buttons. This has been fixed.

Various small GUI bugs have been fixed, that were of no consequence to calculations.

The delay that would sometimes occur when "Stopping…" has become shorter.

Analyzing the test file containing tones at 114 dB would crash the server because of true silence between tones, which creates problems in the CPP calculation. A very low-level dither noise has been added to this file, solving the problem.

macOS only: In FonaDyn versions 2.2.0 - 2.4.7, signal files would be saved as empty upon recording. This has been fixed.

On some high-resolution displays, FonaDyn would sometimes display an awkward mix of font sizes, especially on MacOS computers. This problem has been reduced, but may still be annoying on some screens.

FonaDyn currently stores temporary files in a 'tmp' subdirectory in the userAppSupportDir. These are numerous, small, randomly named files that may accumulate if FonaDyn hangs. They serve no purpose between sessions, and can safely be deleted if they persist. The temp-file location may change in a future release.

On creating an `_Extra.wav` file, it would sometimes be marked with the sampling rate 44100 Hz instead of the actual (decimated) sample rate. This has been fixed.

Known issues

On Windows, the main window maXimize/Restore operations can be done only using Alt+space followed by X or R. Double-clicking the window caption does not work; we do not yet know why.

On Mac, the pop-up menu for controlling the layout does not show the keyboard shortcuts. Please find them in Table 5 on page 48 of the FonaDyn Handbook.

When your laptop is in power-save mode (slower CPU), there could be aberrant behaviour, such as incorrect updating of cluster colours. This is because the internal data refresh cycle may not have had time to complete in one frame (24 frames per second).

Discontinued features

The estimation of EGG residual energy in harmonics higher than *N* (plus noise) has been discontinued. It has been replaced by an estimation of the level of the fundamental partial in the EGG (see above).

The optional time plot of the Crest factor has been replaced by one of the Spectrum Balance. The Crest factor is still computed, displayable in a voice map layer, and saved to `_Log.aiff` files.

The Settings… dialog box option for playing a calibration tone during recording has been removed. Its purpose is now fulfilled by the greater functionality in `FonaDyn.calibrate`.

System and installation aspects

FonaDyn versions from 3.1.0 require SuperCollider **3.13.x** or higher. Please see the file UPDATING.txt for how to do the update.

On Apple MacOS, when running FonaDyn for the first time, the computer must be connected to the Internet, or Apple's safety checks will not work.