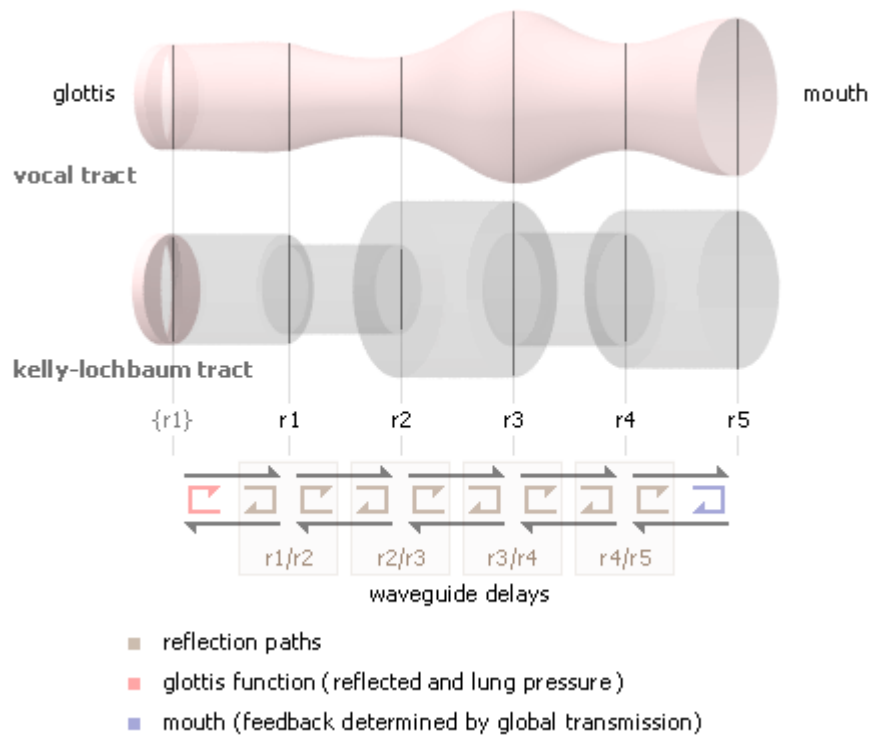


# FAUNA vSTi

<http://www.xoxos.net>

- I. Introduction
- II. Parameter Description
- III. Programming Tips
- IV. Parameter Reference
- V. VST use for Non-Musicians
- VI. License
- VII. Acknowledgements



## I. Introduction

Fauna is a simple waveguide model of the vocal tract intended to provide a flexible platform for the synthesis of abstract and animal voices.

Waveguide synthesis is a physical modeling technique that makes use of delay lines to model the transmission of acoustic vibrations. A sample rate of 44.1kHz translates to a distance of less than 1cm under normal atmospheric conditions, so modern computers can achieve a degree of accuracy in acoustic modeling.

The delay bore is divided into five sections. Reflection coefficients control the transmission between segments, which effectively emulate the area of the tract at each position. Kelly and Lochbaum developed this speech synthesis technique in 1962.

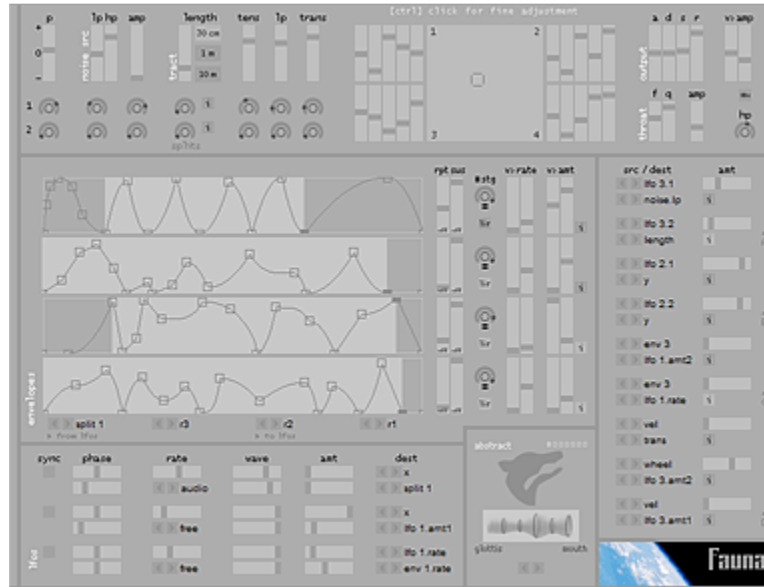
Unlike conventional speech synthesizers, which use the tract to add formants to an oscillator, the fundamental pitch is determined by the overall size and form of the model.

The glottis acts like a reed, opening with air pressure from the lungs (determined by a pressure coefficient) and closing with air reflected back from the mouth. The position of the glottis is a result of these and other forces, such as tension and springiness of the muscle tissue.

Oscillation is produced by the balance of pressure waves in the tract and their effect on the glottis; as the size of the aperture changes, so do its properties of transmission and reflection.

Fauna VST uses four sets of five reflection coefficients to describe the shape of the vocal tract. Four 16-stage graphic envelopes and three dual-contour LFOS are applied to the tract coefficients. A bank of nine sends and two 'splits' route modulators to parameters as groups to assist in the emulation of organic forms.

The structure is not complex enough for speech, keeping in mind that the five coefficients describe the contour of the throat from the glottis and not from the back of the mouth. I was surprised by how many sounds it could create given the low waveguide count.

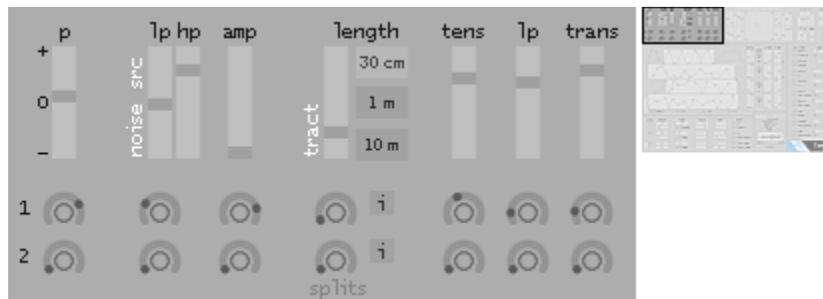


## II. Parameter Description

The top left panel includes the coefficients that apply to the tract, and a noise source that is summed with the pressure coefficient.

At the right, four sets of reflection coefficients are crossfaded with an x-y pad, and signal output parameters are situated to the far right. Signal flow in the model roughly corresponds with position from the left to the right side of the panel.

In addition to the envelope, LFO and send panels, a mini reference panel is situated toward the lower right hand corner. This contains routing heirarchy, examples of summed LFO shapes and other information that may be useful during patching, as well as the control parameter page.



The leftmost parameter is labeled 'p' for pressure. In most patches, lung pressure is modulated by another source, so this slider generally serves as a center for modulation.

The next three sliders amplify and filter a noise source, which can add realism, presence and roughness. The human glottis rarely closes all the way, and so produces more broadband signals than an instrumental reed, such as a clarinet.

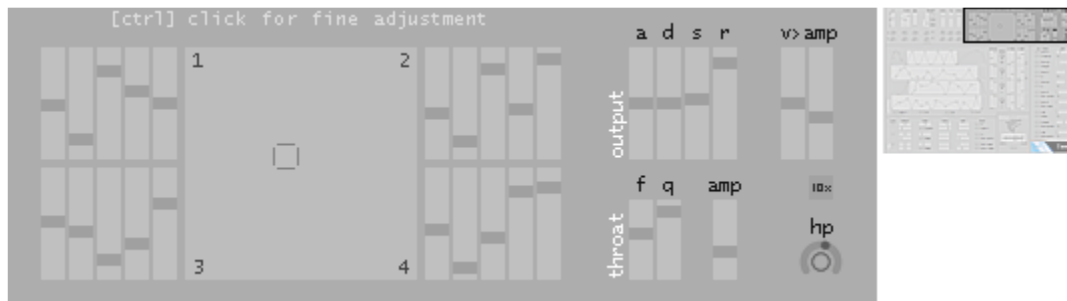
The length of the vocal tract has three ranges. The smallest length corresponds to "five samples," or between four and five centimeters at 44.1kHz.

Tension (tens) applies to how responsive the glottis is to ambient pressure waves. Lower settings create gentler timbres.

A low pass filter inside the vocal tract can be used to shape the damping properties of the organic tissue inside the throat. This 6db filter should also be set to indicate the size of the speaker.

The last coefficient, transmission, equates to feedback in a standard delay.

To assist in modeling organic movement, two 'splits' route modulators to these parameters as groups. This can create more realistic timbres that suggest a cohesive acoustic system.



The x-y pad crossfades between four sets of reflection coefficients. These are labeled 'r' in the modulation assignment lists. The sliders can be seen as a graphic representation of a cross-section of the vocal tract from glottis to mouth, arranged from left to right.

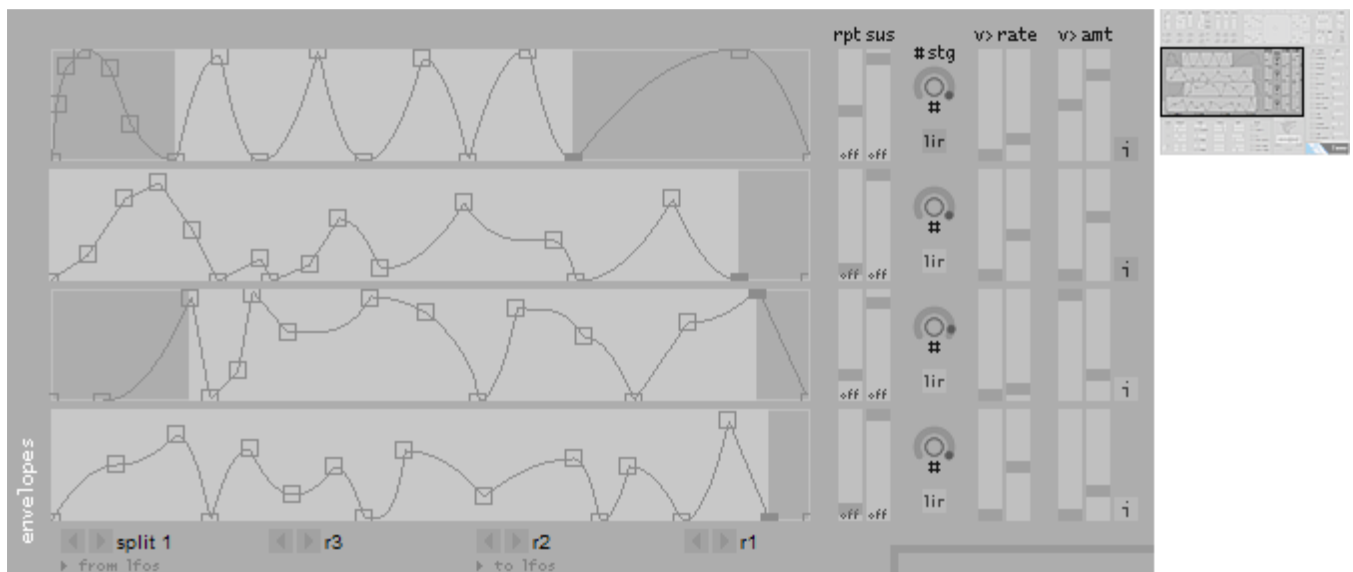
The rightmost slider corresponds directly to the mouth aperture, although the leftmost slider indicates the area one fifth of the distance from the glottis.

Remember that the acoustic length of the vocal tract runs from the glottis to the mouth, and not from the back of the throat, which for humans, is positioned roughly at the second slider.

Given the digital mileu, the ratios of slider values can produce a variety of subtle effects.

Below the ADSR envelope is a resonant lowpass filter that taps the tract at the glottis for modeling the radiation of sound from the throat. This, and the final highpass filter positioned below the amp slider, can add a sense of the size of the creature.

The 10x button under the amp slider amplifies the signal if your patch requires it.



The 16-stage graphic envelopes (coded by Chris Kerry) have a variety of contours which are selected by left-clicking. A list of the envelope contours is documented on the reference panel.

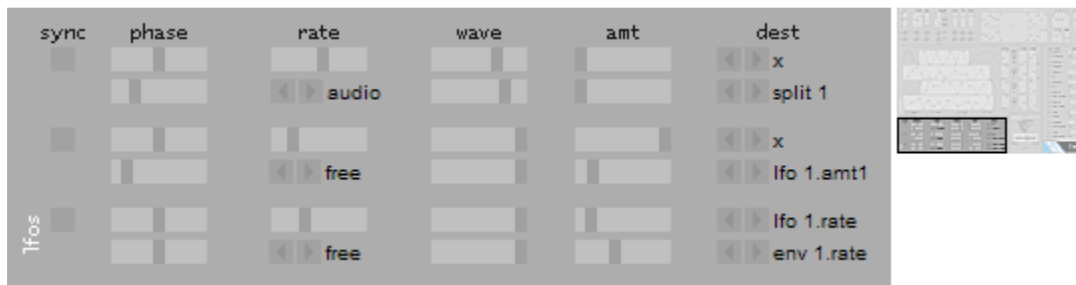
It is important to note that the repeat and sustain values should never be set at values higher than the current number of active stages. You will notice high CPU during the release if you do so.

LIR stands for 'loop in release.' When inactive (dark), releasing the key continues the envelope past the sustain point.

Note that while the envelopes have a 'rate' slider, the graphical time axis of the envelope display is relative to the adjustments that have been made to it, and it may be necessary to edit the envelope points to achieve the desired time scale.

Envelopes (and LFOs) can be assigned to modulate those with a lower number.

Envelopes 3 and 4 can be assigned to the LFOs, while envelopes 1 and 2 can be modulated by the LFOs.



The LFOs have a selection of five symmetrical waveshapes and a variable, or noise setting. The contours are illustrated on the reference panel.

LFOs are arranged in groups of two; they share a common rate setting, and have separate contours, trimmers and assignments.

By using phase to offset the contours, asymmetrical and varying waveshapes can be produced, some of which are also illustrated on the reference panel. These include square and saw as well as less common shapes. Modulating the phase of one waveform can effect pulsewidth, fade saw to triangle to ramp and other effects.

The LFOs have a variety of rate scalings:

FREE	conventional LFO range
AUDIO	audio rate oscillation for FM effects
KEY -2/0/+2	tracks the keyboard at audio rate
BPM /2	1/64m, 1/32m, 1/16m, 1/8m, 1/4m, 1/2m, 1m, 2m, 4m, 8m, 16m
BPM /3	1/48m, 1/24m, 1/12m, 1/6m, 1/3m, 2/3m, 1m, 4/3m, 3/2m, 8/3m, 3m, 16/3m, 6m

LFOs (and envelopes) can be assigned to modulate those with a lower number.

LFOs can be assigned to envelopes 1 and 2, and modulated by envelopes 3 and 4.

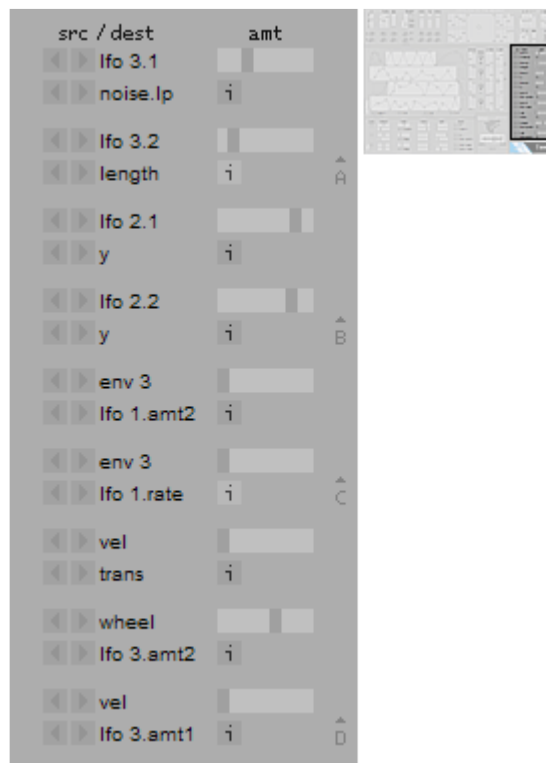
Sends assign LFOs, envelopes and controllers to tract parameters, other modulators, and lower numbered sends.

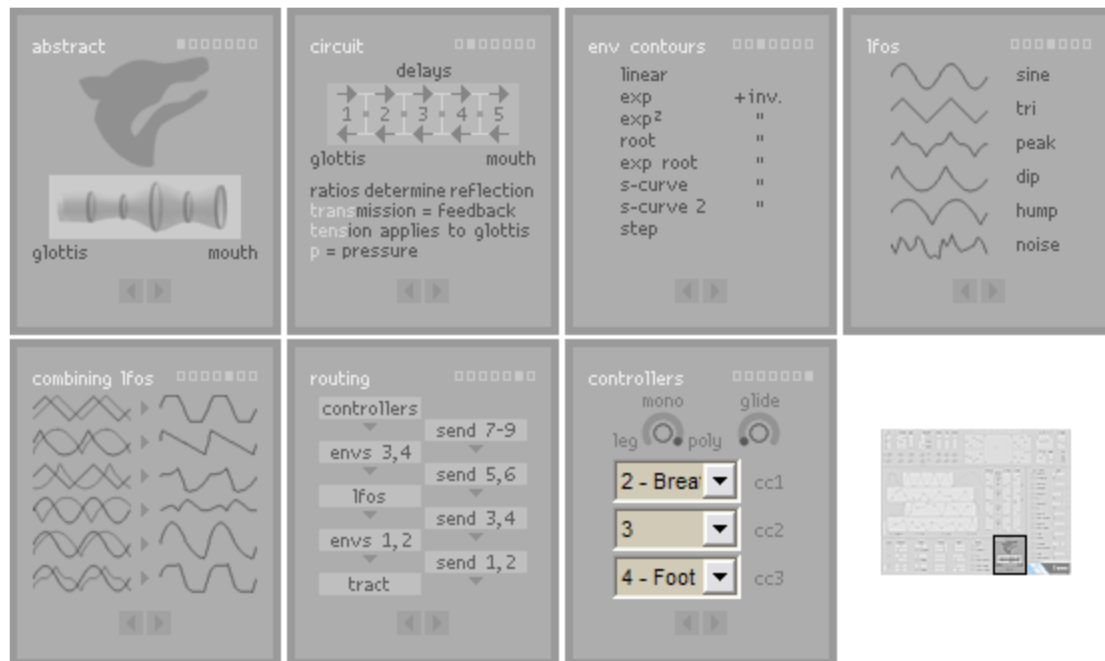
Not only does this allow modulators to control groups of parameters, it also allows them to control how much they do so..

The reference panel includes a chart of the modulation heirarchy. The sends are arranged in groups labeled A through D to indicate which modulators they control and are controlled by

Groups A and B (sends 1-4) are uniquely assignable to throat filter parameters. This routing is used in some of the laughing presets.

If you wish to assign a MIDI controller, such as key tracking or velocity, use group D to save the other sends for internal modulators.





The reference panel is designed to quickly orient experienced synthesists with Fauna's features. The first and second pages illustrate concept and form. The third, fourth and fifth display modulation contours, and the sixth page illustrates the control heirarchy. The heirarchy is linear - sends are positioned to one side for clarity.

The seventh panel contains the polyphony and glide parameters, and cc#s assigned to the three controllers used in the sends in case your host requires this method.

### III. Programming Tips

Although Fauna is significantly less complex than a real vocal tract, by modeling the fundamental dynamics of the target system, it can usually be expected to behave similarly. For instance, it takes significantly more pressure to drive a long vocal tract to oscillation than it does a small one. If you attempt to patch a tract longer then two meters, you will probably have to use at least one of the sends to duplicate the modulator assigned to pressure.

Use a split to drive pressure instead of assigning a modulator directly to pressure.

If you are attempting to produce a breathy tone, start with the noise hp at a very high setting. Adding lower noise frequencies creates more broadband frequencies and hysteresis than breathiness, and is better used for snarling or gravelly vocalisations.

In most situations, noise amplification is best performed with a modulator, since a constant amplification will sound unnatural when pressure crosses from inhalation to exhalation.

The human glottis rarely closes all the way. I have read that women have breathier voices because the folds make less contact.

Be careful about modulation ranges - many parameters, such as transmission, have a ceiling and floor. Excessive modulation may sound unnatural when clipped.

If you intend to use Fauna extensively, limit modulation to one set of reflection coefficients to familiarise yourself with the effect of each position: a higher r1 setting will produce a deeper tone, and smaller settings will produce a falsetto.

Use the x-y pad when recording to explore regions that produce the most natural timbres.

Fauna is a more complex system than many VST. Sometimes motion between two settings will produce effects that static settings in the same range do not.

The throat filter is usually set to accentuate the resonance of the throat. For some patches, such as screams, setting it at the resonance of the oral cavity improves realism.

In many of the presets, noise LFOs are used to provide continuous variation. The rate of modulation will be slower than the other waveshapes. Group noise modulators together when possible to reserve LFOs for other rate settings, and give each noise waveform a different phase setting to avoid identical output.

Stacking noise LFOs, eg. modulating the rate of one with another, produces more dynamic variation.

## IV. Parameter Reference

### VOCAL TRACT

p	Pressure from lungs
noise lp	Frequency of 6db low pass filter applied to noise source
noise hp	Frequency of 6db high pass filter applied to noise source
noise amp	Volume of noise source (summed with pressure coefficient)
tract length	Ranges from 0 to 1 times the length range value (30cm, 1m, 10m)
tens	Tension of the glottis
lp	Damping applied within the waveguide circuit
trans	Transmission (feedback) 0 to unity
r1-5	Reflection coefficients
adsr	Amplification envelope
v>	Responsivity of amplification to velocity
throat f	Frequency of throat resonance
throat q	Resonance of throat lowpass filter
throat amp	Volume of filtered throat tap
hp	Frequency of 6db high pass filter applied to output

### ENVELOPES

rpt	Beginning of envelope loop. Keep in off position when not in use.
sus	Sustain stage or end of envelope loop. Keep lower than # of active stages.
#stg	Number of active stages (1-16)
lir	Loop in release
v>	Velocity modulation of parameter to right
rate	Envelope rate (also controlled by graphic settings)
amt	Height of envelope
i	Invert envelope output

### LFOS

sync	Reset LFO to phase position on MIDI gate
phase	Position or offset of LFO waveshape
rate	LFO rate - see above for selections
wave	Sine, tri, peak, dip (waves), hump (hills), noise
amt	Amplification of LFO

### SENDS

send src	Modulator or controller assigned to send
dest	Assignment of send src

### CONTROLLERS

legato	Mono, overlapping notes change MIDI note and do not retrigger gate
mono	Overlapping MIDI notes retrigger gate
poly	Overlapping MIDI notes trigger up to 4 voices of polyphony
glide	Portamento (MIDI note value generally used as modulator)
cc 1-3	Internal assignment of up to 3 controllers for use in some hosts

## V. VST Use for Non-Musicians

This section is intended for those who aren't familiar with music software and want to record sound effects for animation et c.

VST (Virtual Studio Technology) is an open-source plugin format developed by Steinberg. To use a VST plugin, a VST host is required. Today, there are a multitude of free hosts available suited for a wide range of applications, many of which are very simple to use. Note that Fauna only runs on windows operating systems.

Perhaps the simplest solution is Tobybear's MiniHost, available at:

[http://www.tobybear.de/p\\_minihost.html](http://www.tobybear.de/p_minihost.html)

To create and record sounds with MiniHost, download the zip file and extract the files to a folder of your choice. There is no installation executable, so MiniHost will not adversely affect your system.

Place the fauna.dll file inside the same folder as MiniHost.exe.

The first time you run MiniHost, a donation splash window will appear for a few seconds. This window does not appear on subsequent uses.

Once the GUI appears, select VST > Load VST Plugin ... from the menu at the top of the GUI. A standard windows file browser will appear, where you can select fauna.dll as the plugin to load. After a few moments, the Fauna GUI will appear.

At the top of the GUI is a black box reading 000: Blank, indicating the currently selected patch. Use the arrows inside this box to select patches.

You can play Fauna at any time by clicking on the 'keyboard' at the bottom of the screen.

To record the sounds, select Main > Show MID/WAV Window ... and a new dialogue box will appear. At the bottom of this window, simply click the [record] button, which will produce a dialogue box to enter the name the .wav file.

If you need a free .wav editor to trim the recordings, I recommend Audacity:

<http://audacity.sourceforge.net>

## VI. License

Licensing for the use and distribution of Fauna VST is as follows:

You may NOT redistribute Fauna or other xoxos software by any means without my permission.

Use of Fauna in any commercial or otherwise paid venture requires purchase of a license.

YOU decide how much you wish to pay. It is suggested that your payment is in the amount you generally pay for software synthesizers.

There is no expression of guarantee due to the use of third party resources.



## VII. Acknowledgements

John L. Kelly and Carol. C. Lochbaum for developing their vocal tract model. More information -  
[http://ccrma.stanford.edu/~jos/pasp/Singing\\_Kelly\\_Lochbaum\\_Vocal\\_Tract.html](http://ccrma.stanford.edu/~jos/pasp/Singing_Kelly_Lochbaum_Vocal_Tract.html)

Julius Orion Smith III for an abundance of contributions to synthesis, and especially in regards for his commitment to open source documentation.  
<http://ccrma.stanford.edu/~jos>

Rob Martino for his paper on the first animal voice synthesizer, Synthasaurus -  
<http://robmartino.com>

In memory of Mike W. Macon for Flinger -  
<http://www.cslu.ogi.edu/tts/flinger>

Jeff McClintock for SynthEdit -  
<http://www.synthedit.com>

Chris Kerry for his assistance and development of ASM-coded resources for SynthEdit -  
<http://www.chriskerry.f9.co.uk>

A general acknowledgement is given for the people who provide and maintain educational, artistic and communication resources.

And..

To those who provide voices for animals!