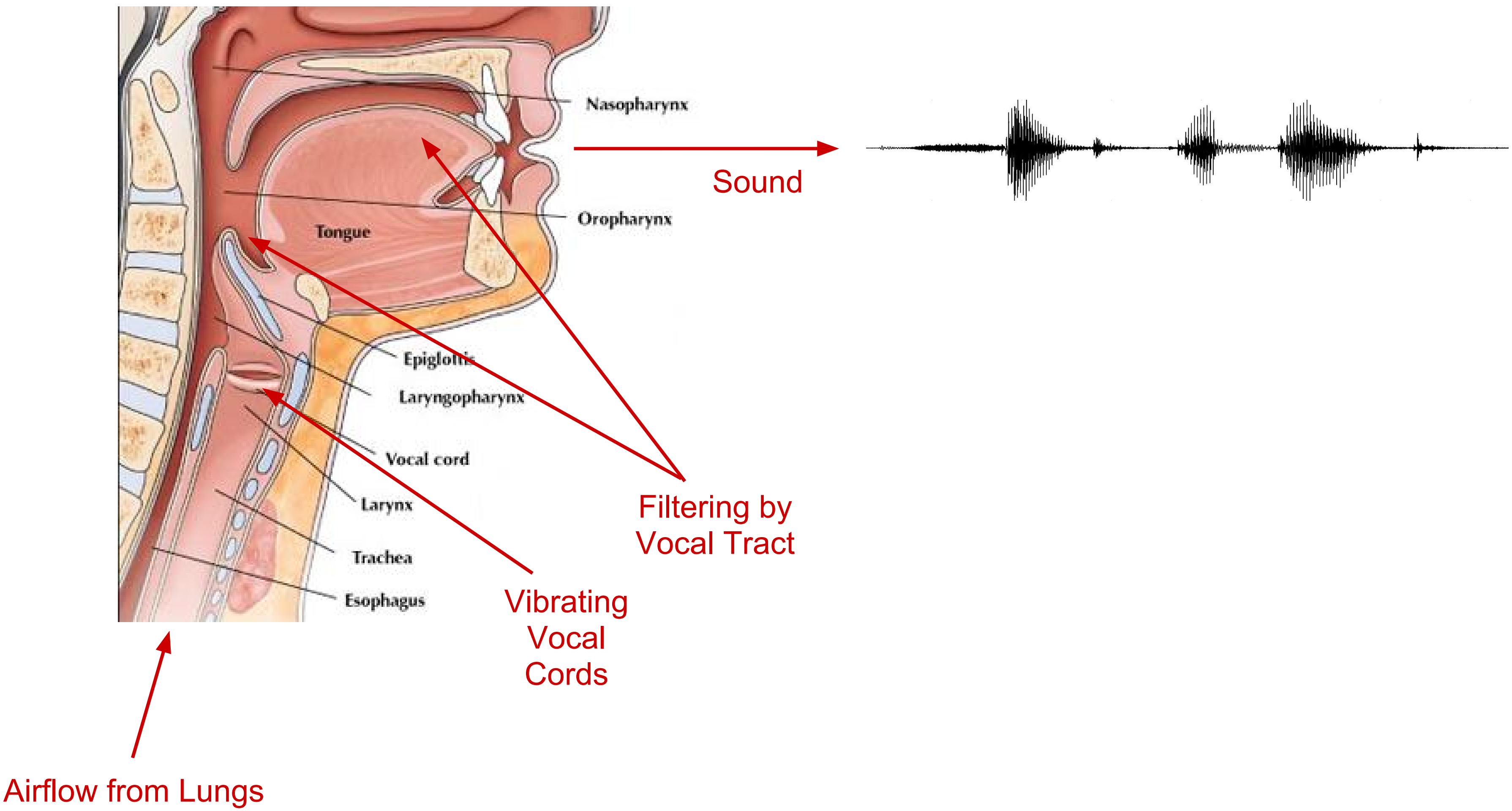
Control of Zebra Finch Vocalizations

Mike Schachter
NCASO Project Presentation
June 1, 2013

Human Vocalization

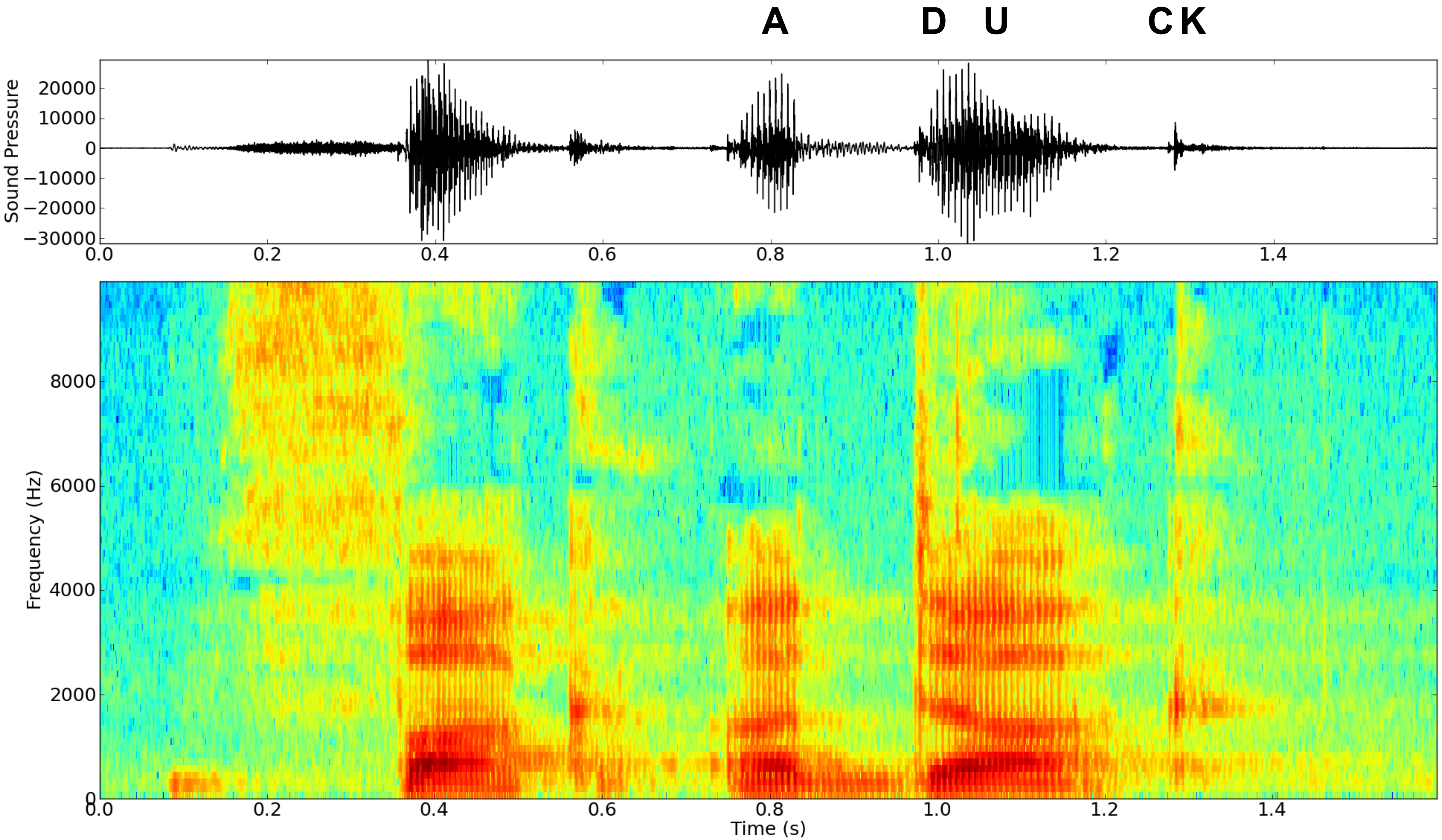


Air flow from the lungs makes the vocal cords vibrate, which produces a sound signal. That sound signal is then filtered through the rest of our vocal tract to produce speech.

Vocal Cords Vibrate

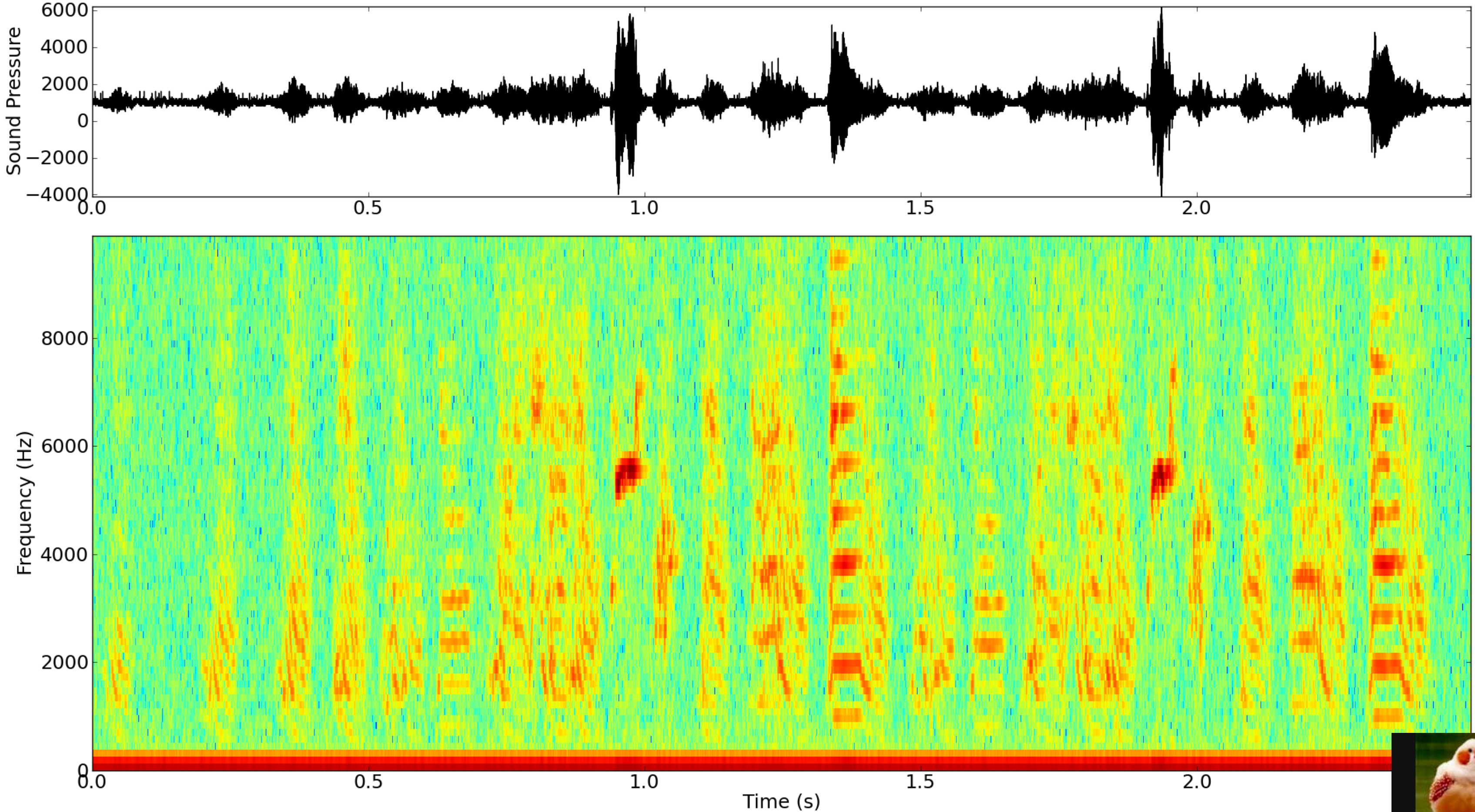
The vibration of the vocal cords produces an *oscillation* in the air flow which is then filtered in the vocal tract to produce speech.

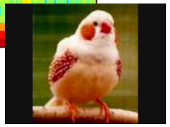
Analysis of Sound



Human speech is comprised of a continuous sequence of syllables.

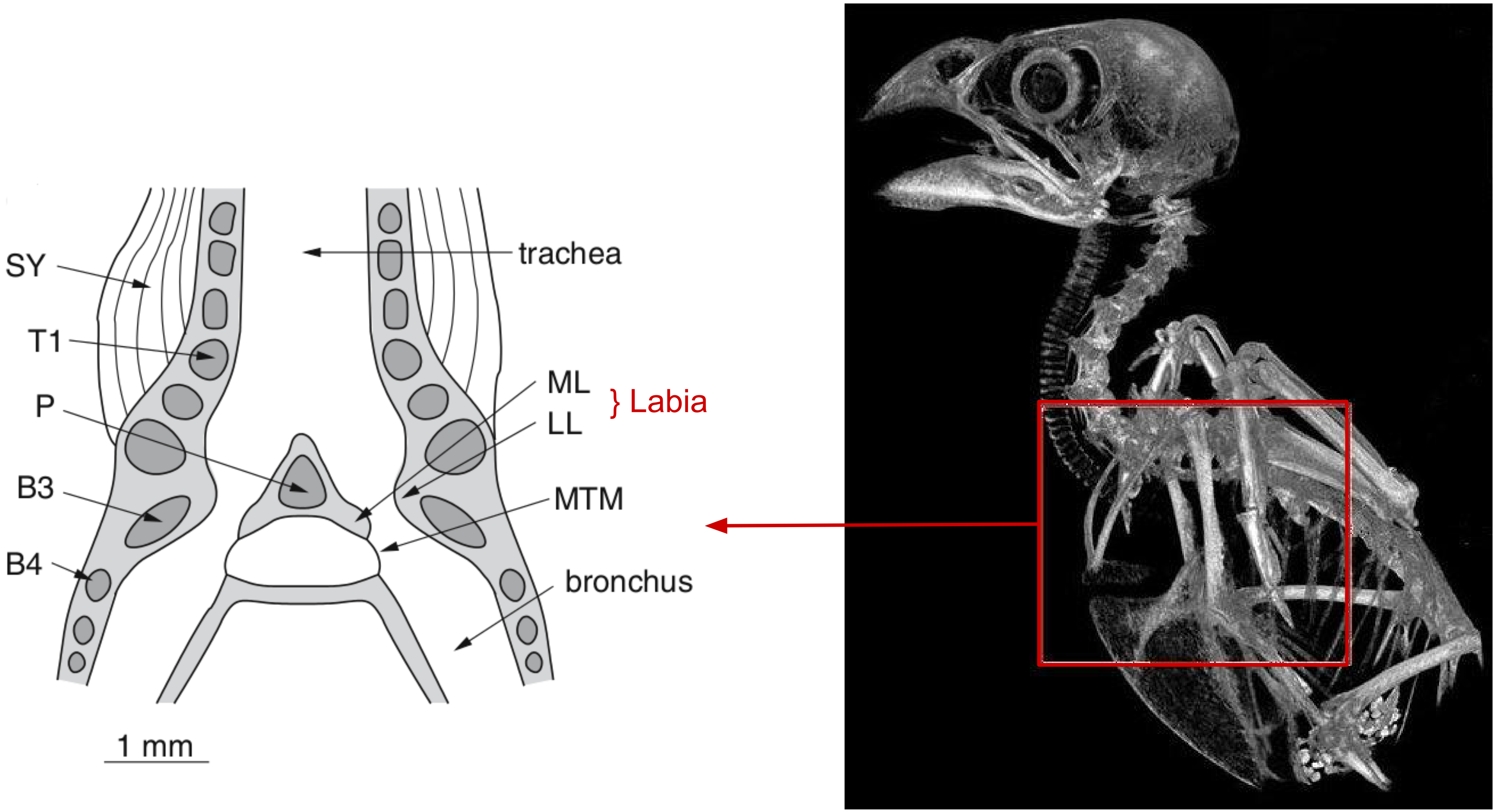
Zebra Finch Song





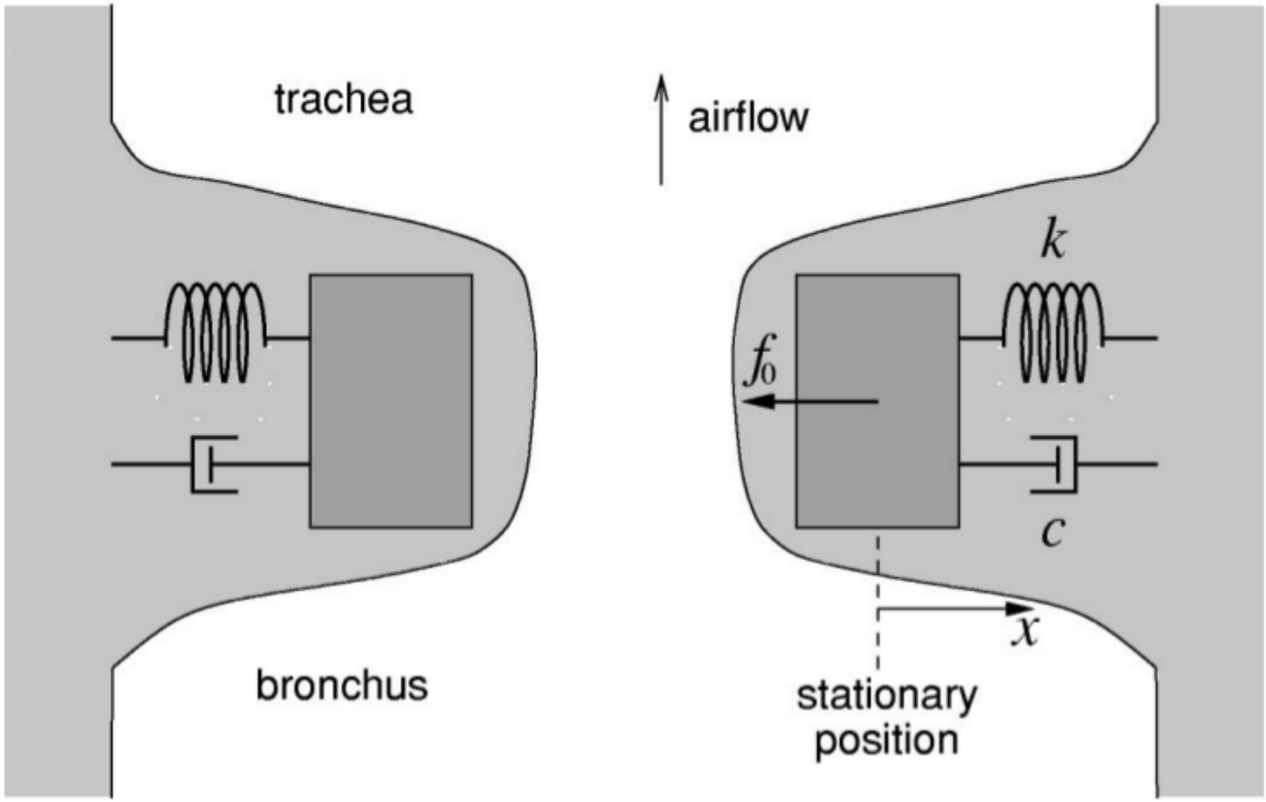
Zebra Finch song is comprised of a sequence of acoustically-complex syllables.

Zebra Finch Vocal Production



The Zebra Finch has two independently-controllable *synrinxes* that produce oscillations using the same mechanism that human vocal cords do.

Vocal Cord Model

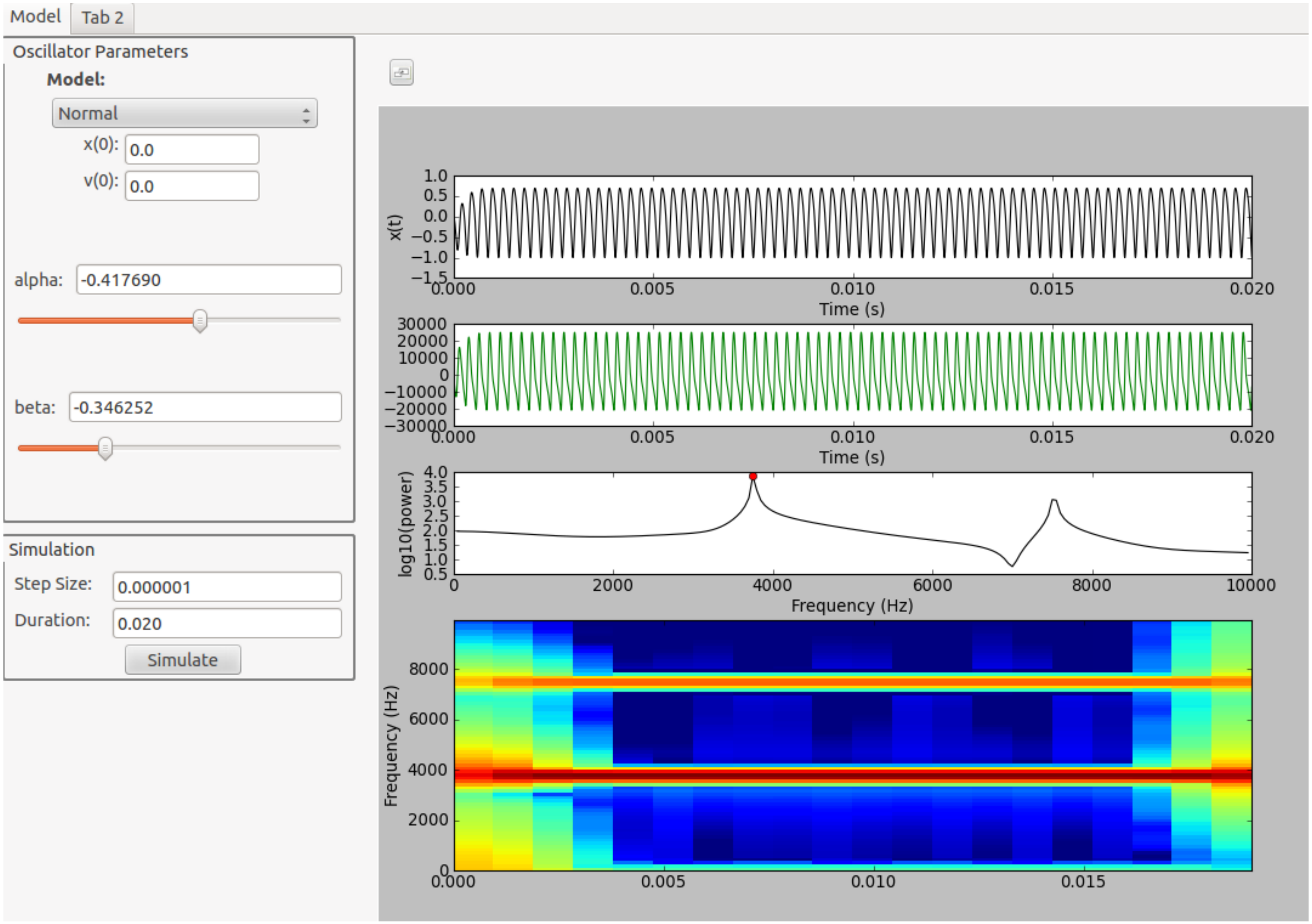


$$\dot{x} = v$$

$$\dot{v} = \gamma^2\alpha + \gamma^2\beta x - \gamma^2 x^3 - \gamma x^2 v + \gamma^2 x^2 - \gamma x v$$

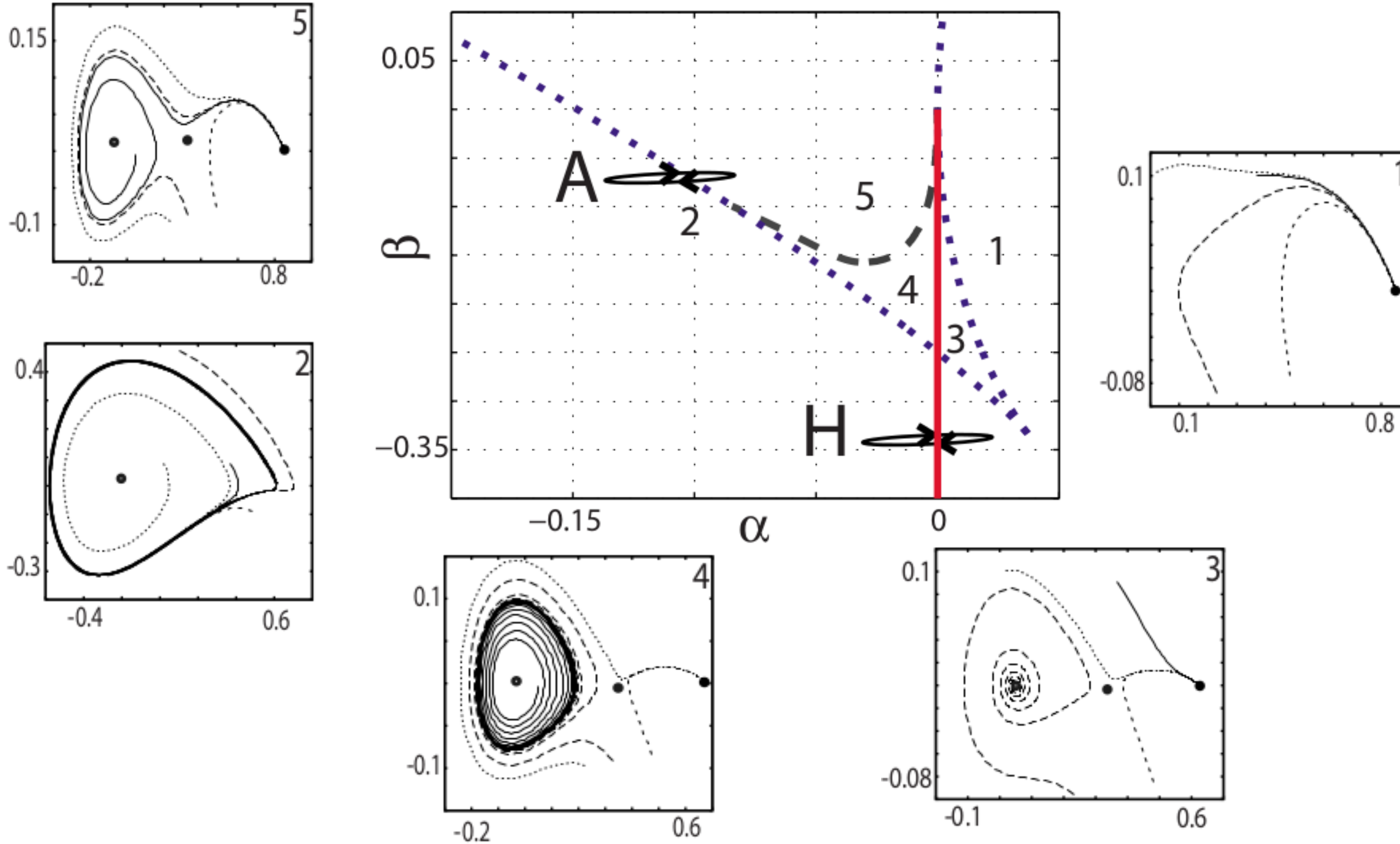
One half of the vocal cord is modeled as a mass attached to a nonlinear spring. Alpha is the air flow from the lungs, and beta is the stiffness of the vocal cords. Gamma is a constant.

Simulation



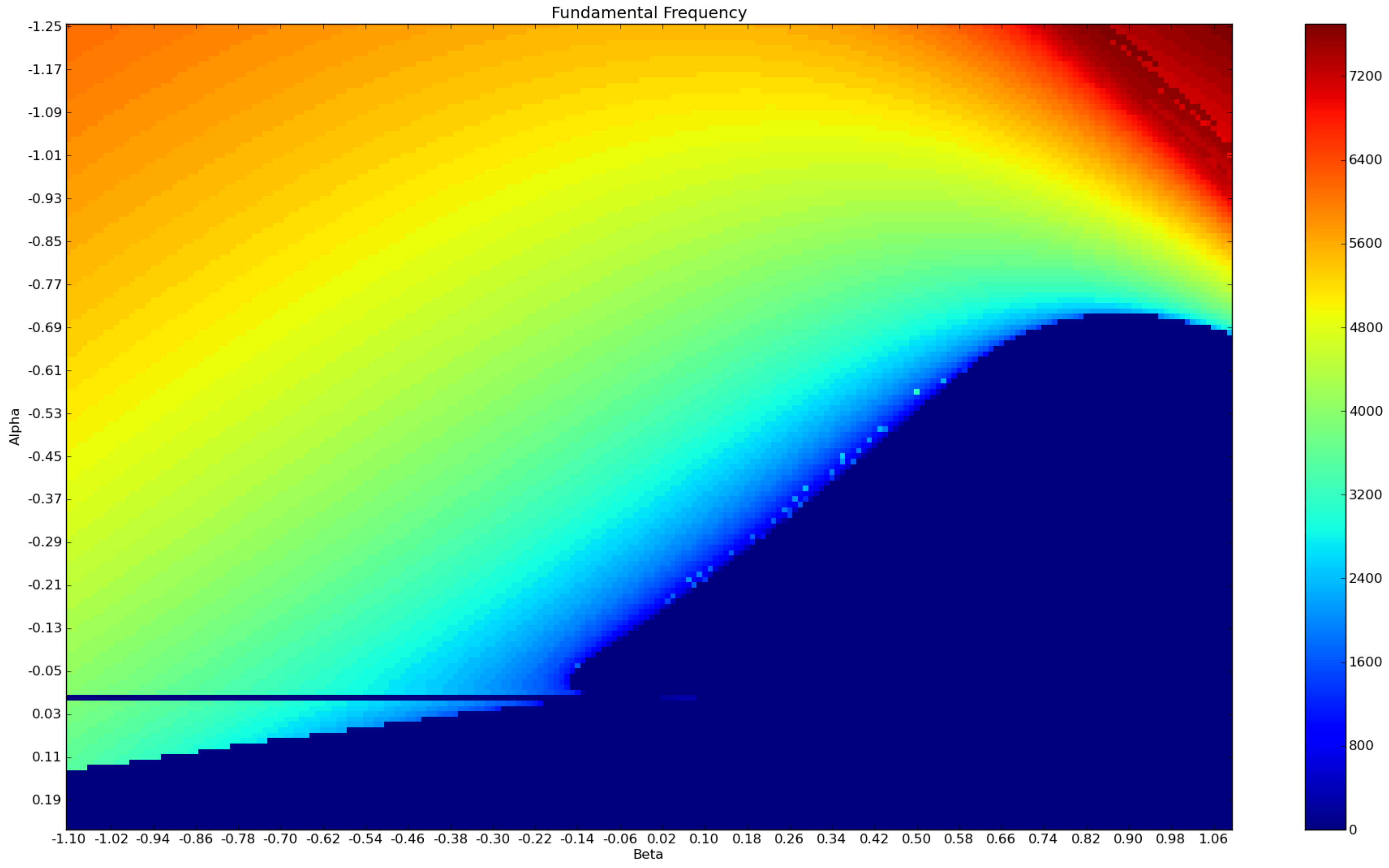
I wrote a program to simulate the model, and soon will add the capability to "design" synthetic syllables.

Bifurcations of Control Parameters



Oscillations are born in both Hopf (**H**) and Saddle-node-in-limit-cycle (SNILC, **A**) bifurcations.

Acoustics of Control Space



Fundamental frequency plotted as a function of control parameter.

Optimal Control of the Controls

Assume the control trajectory is a controlled linear dynamical system:

$$\boldsymbol{\phi}(t) = [\alpha(t)\ \beta(t)]^T$$

$$\dot{\boldsymbol{\phi}} = A\boldsymbol{\phi} + B\boldsymbol{u}$$

An optimal form for **u** can be obtained by minimizing an instantaneous cost function over time:

$$C\left(\boldsymbol{\phi}(t_{k-1}), F_f(t_k), \boldsymbol{u}\right) = \boldsymbol{\phi}(t_k)^T Q \boldsymbol{\phi}(t_k) + \boldsymbol{u}^T R \boldsymbol{u} + C_f(F_f(t_k), \boldsymbol{\phi}(t_k))$$

The cost of fitting the desired fundamental frequency at a given time point is specified as:

$$C_f(F_f, \boldsymbol{\phi}) \propto \frac{1}{p\left(F_f | \boldsymbol{\phi}\right)}$$