Functioning Model of the Larynx and Voice Analysis System

Onyekachi “Kachi” Odoemene

ECE 158 Final Product Review

Professor T.J. Manuccia
Professor A.B. Hancock, PhD

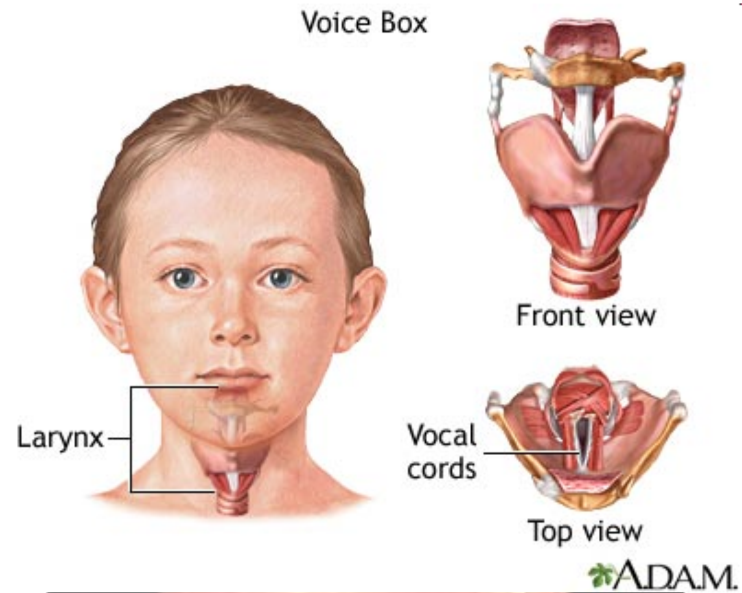
May 5 2009

Project Goals

- Design, Build, Test:
  - Biomechancial model of larynx with active vocal folds which vibrate to produce sound in response to airflow
    - Simulate common vocal disorders
  - Software system to analyze voice signals and extract parameters that will aid clinicians in vocal pathology diagnostics

Introduction: Larynx

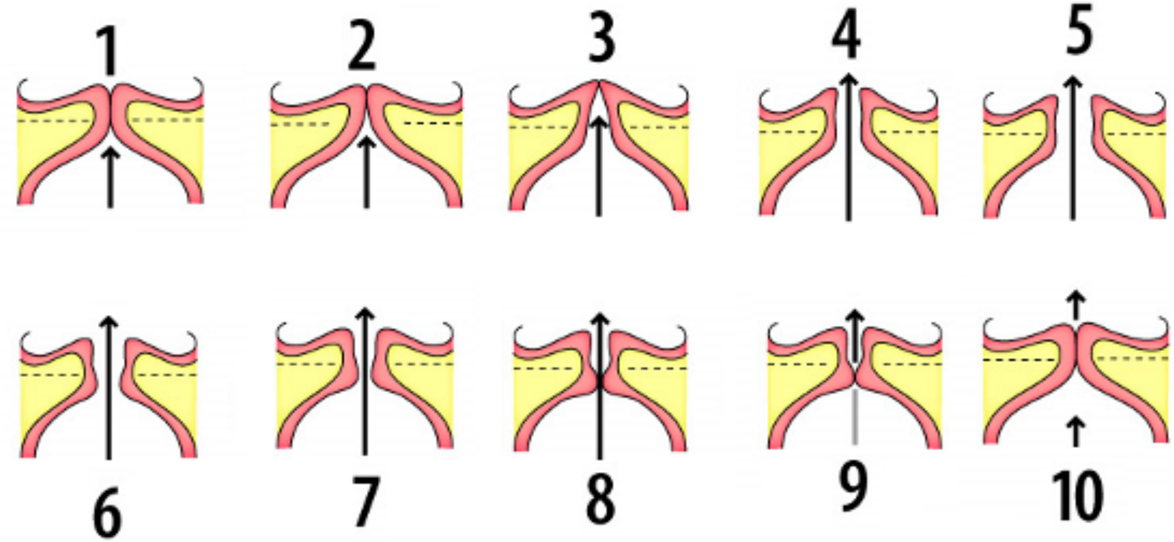
- Commonly known as 'voice box'
- Houses vocal folds (or vocal cords)
  - Mucous membrane
- Vocal fold vibration produces sound
  - Achieved via mucosal wave



Images from: http://apps.uwhealth.org/health/hie/2/19708.htm

http://www.voicemedicine.com/normal_voice_functioning.htm

Mucosal Wave

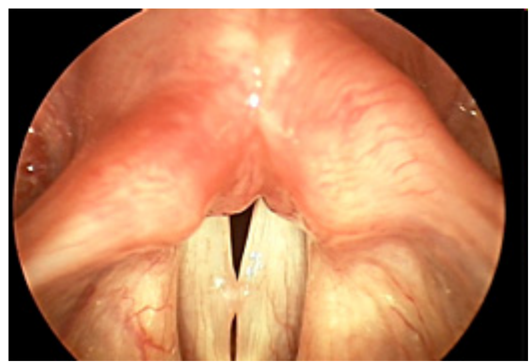


- Subglottal pressure (pressure from lungs) blow apart vocal folds from bottom to top
- Elastic nature of vocal folds attempts to return to adducted state
  - Bottom begins to close as top opens
- Bernoulli effect “slurps” the folds back to adducted state
- Repeated several times during voicing → Mucosal wave

Common Voice Disorders

[1]Normal

[1]Polyp



[1]Nodule

Cyst[2]

Images from: [1]http://www.voicemedicine.com/disorders.htm [2]http://www.yorku.ca/earmstro/journey/images/cyst2.jpg

Vocal Pathology Diagnostics

- Measurable acoustic parameters:
  - Average Fundamental Frequency (F0)
  - Jitter Percent→ measure of variation in pitch
  - Shimmer Percent→ measure of variation in volume
  - Noise to Harmonic Ratio→ voice quality measure of hoarseness and breathiness

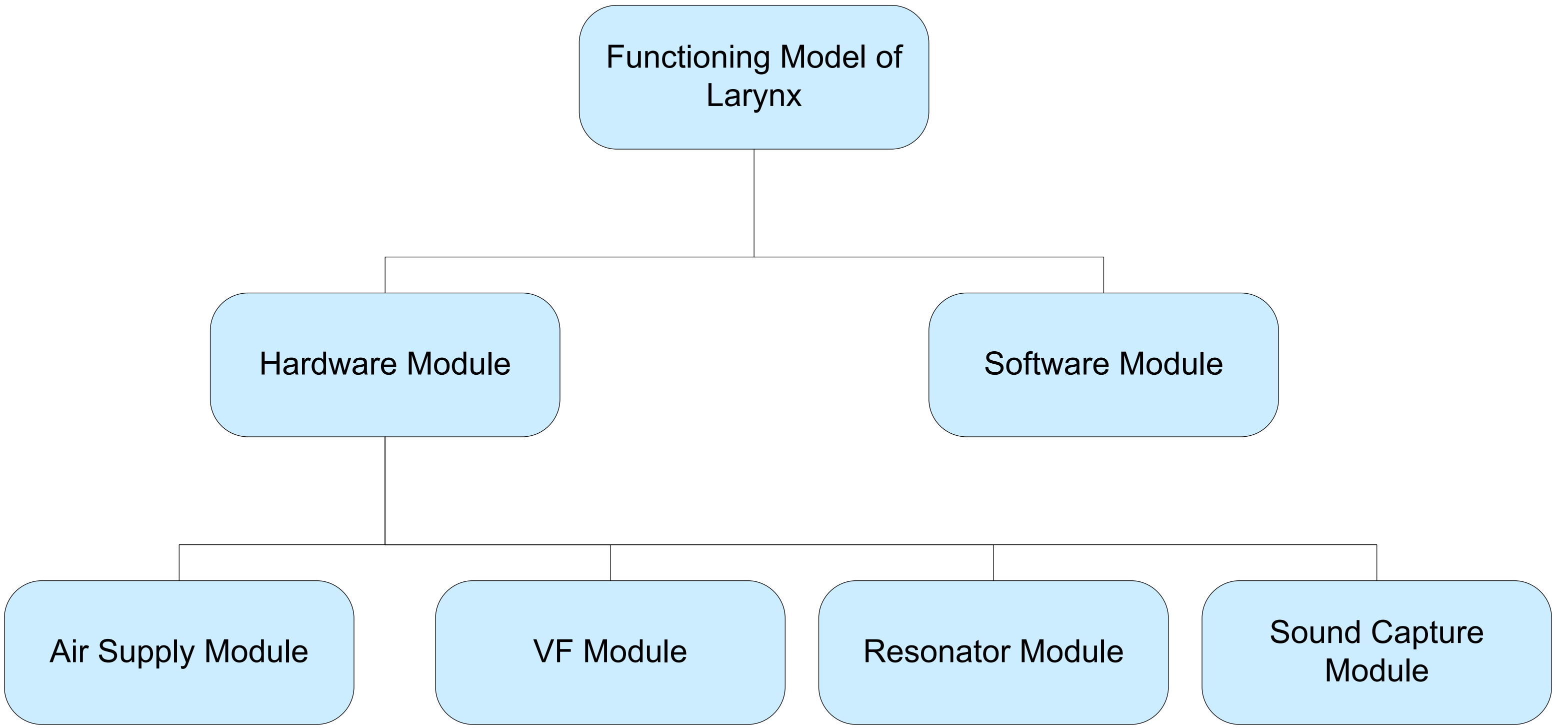
Motivation & Application

- Visual aid of voicing mechanism
  - Teaching and training tool
    - Simulation of voice disorders
  - Functioning model vs stationary model
    - Active vocal folds

State of the Art

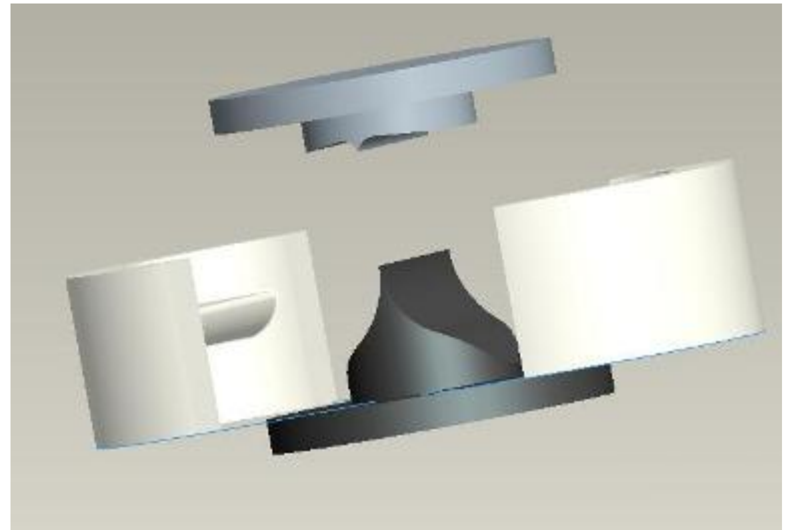
- Software
  - Multi-Dimensional Voice Program
    - Gold standard speech analysis software
    - Calculates 22 parameters on voice sample
- Model Vocal Folds
  - Anthropomorphic Talking Robot
    - Waseda University, Japan
    - Equipped with mechanical lungs, active vocal folds, articulators, tongue, teeth etc
    - Can really talk!

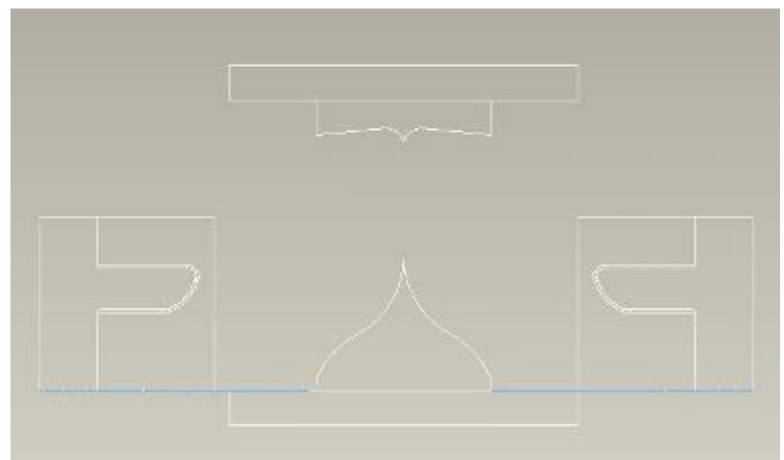
Modular Decomposition: Hardware



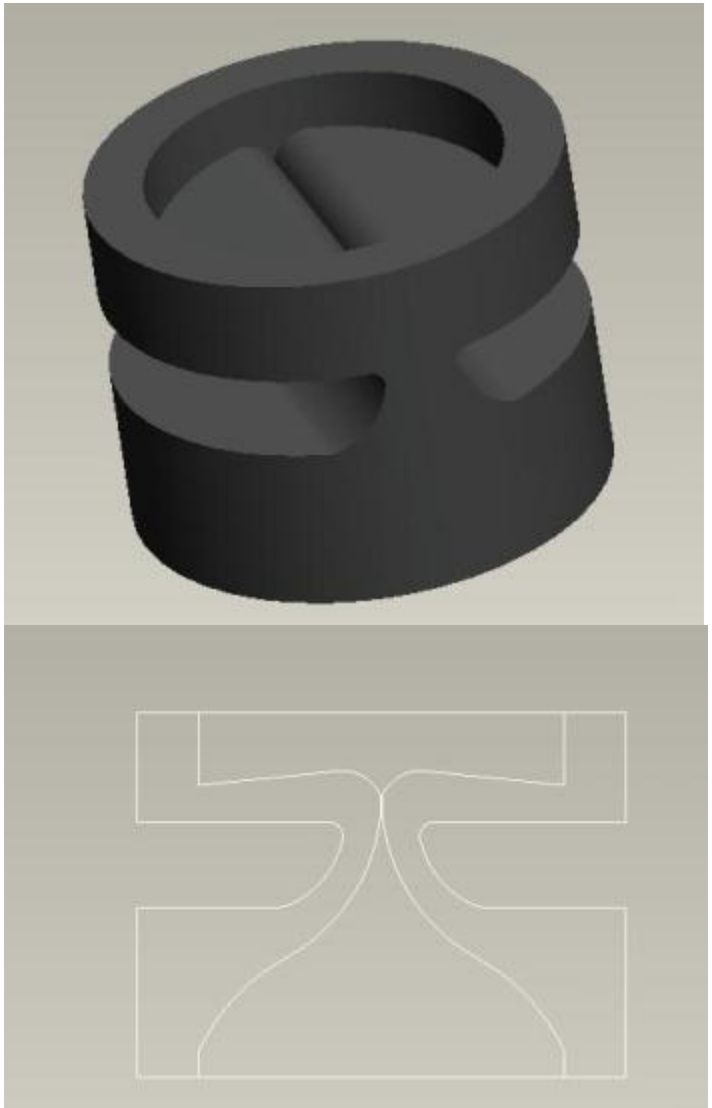
Vocal Fold Design

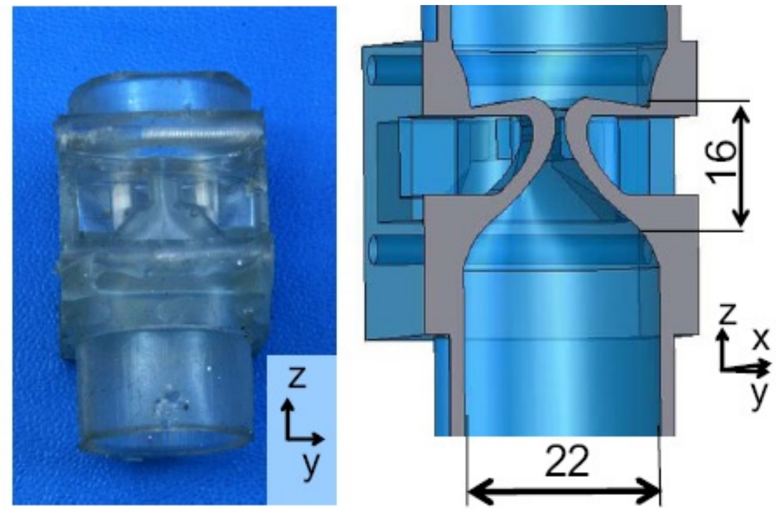
- Based off Japanese Talking Robot
- Use CAD toool (ProE) to design mold
- Rapid prototype mold
- Cast mold with silicone rubber





Vocal Fold Design





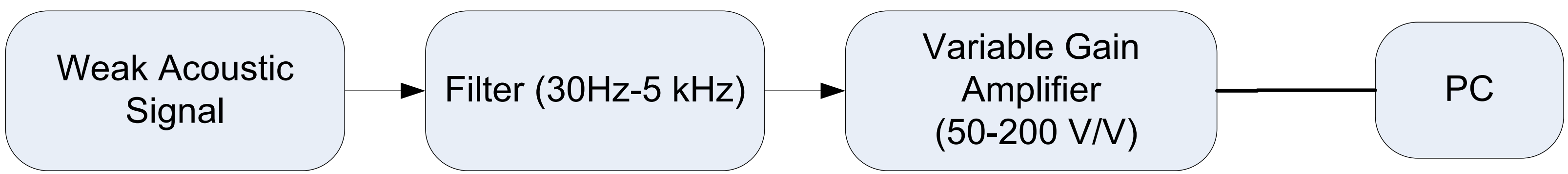
Vocal cords Shape Coronal

Japanese Robot Vocal Fold[1]

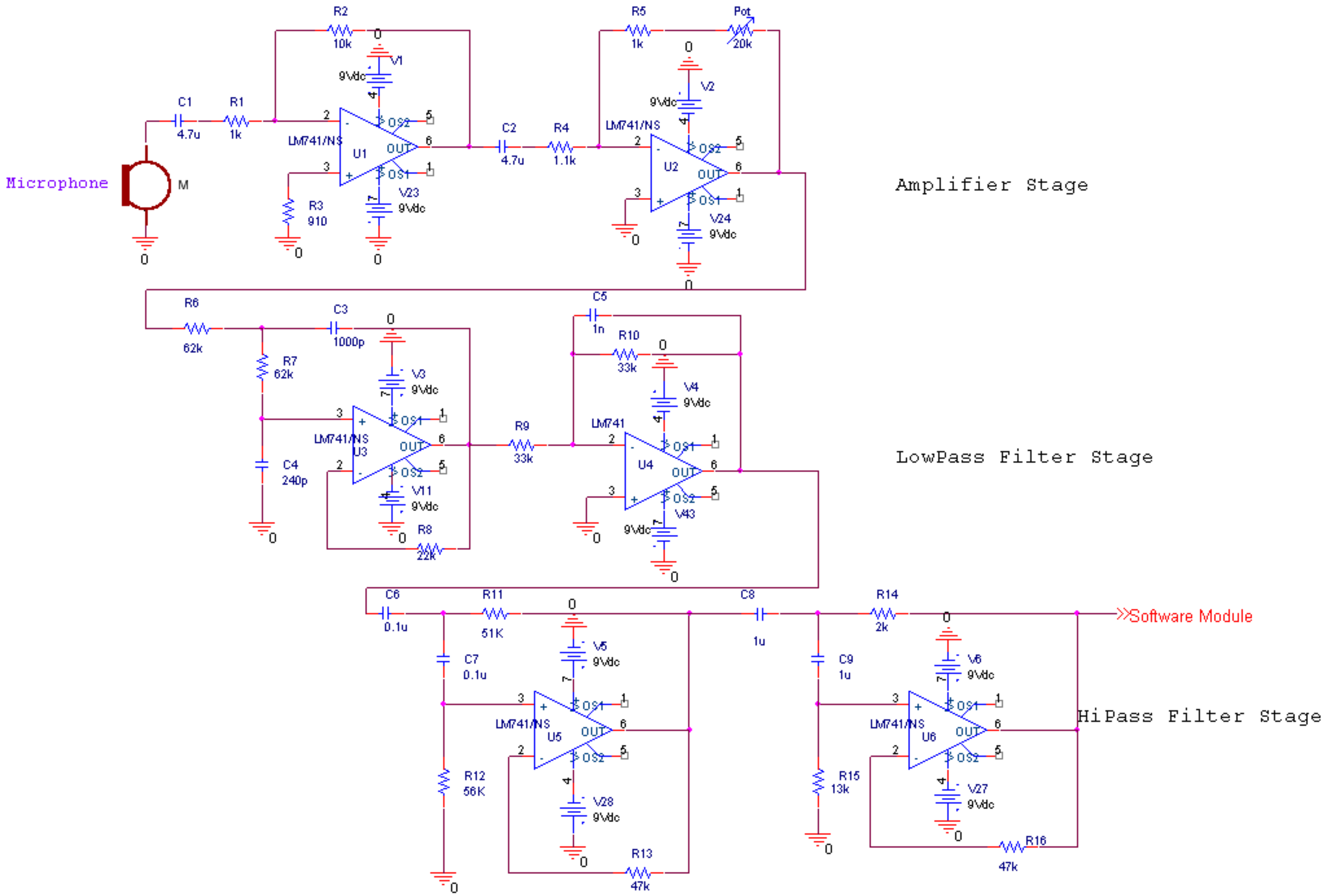
←My Vocal Fold Design

[1]Images from: Figure 2c. Fukui, Kotaro. "New Anthropomorphic Talking Robot having Sensory Feedback Mechanism and Vocal Cords based on Human Biomechanical Structure." IEEE International Conference on Robotics and Automation 2006 13th ser. 8 (2006): 1095-1100.

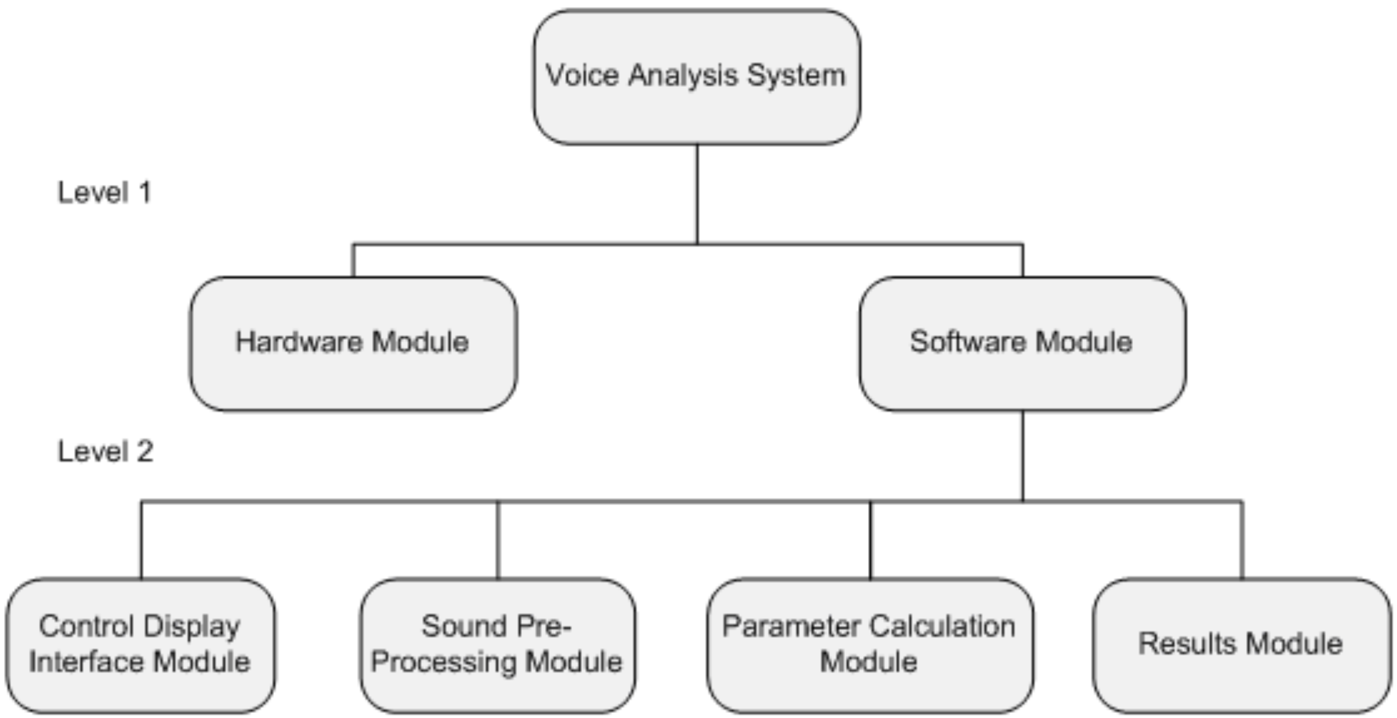
Sound Capture Design



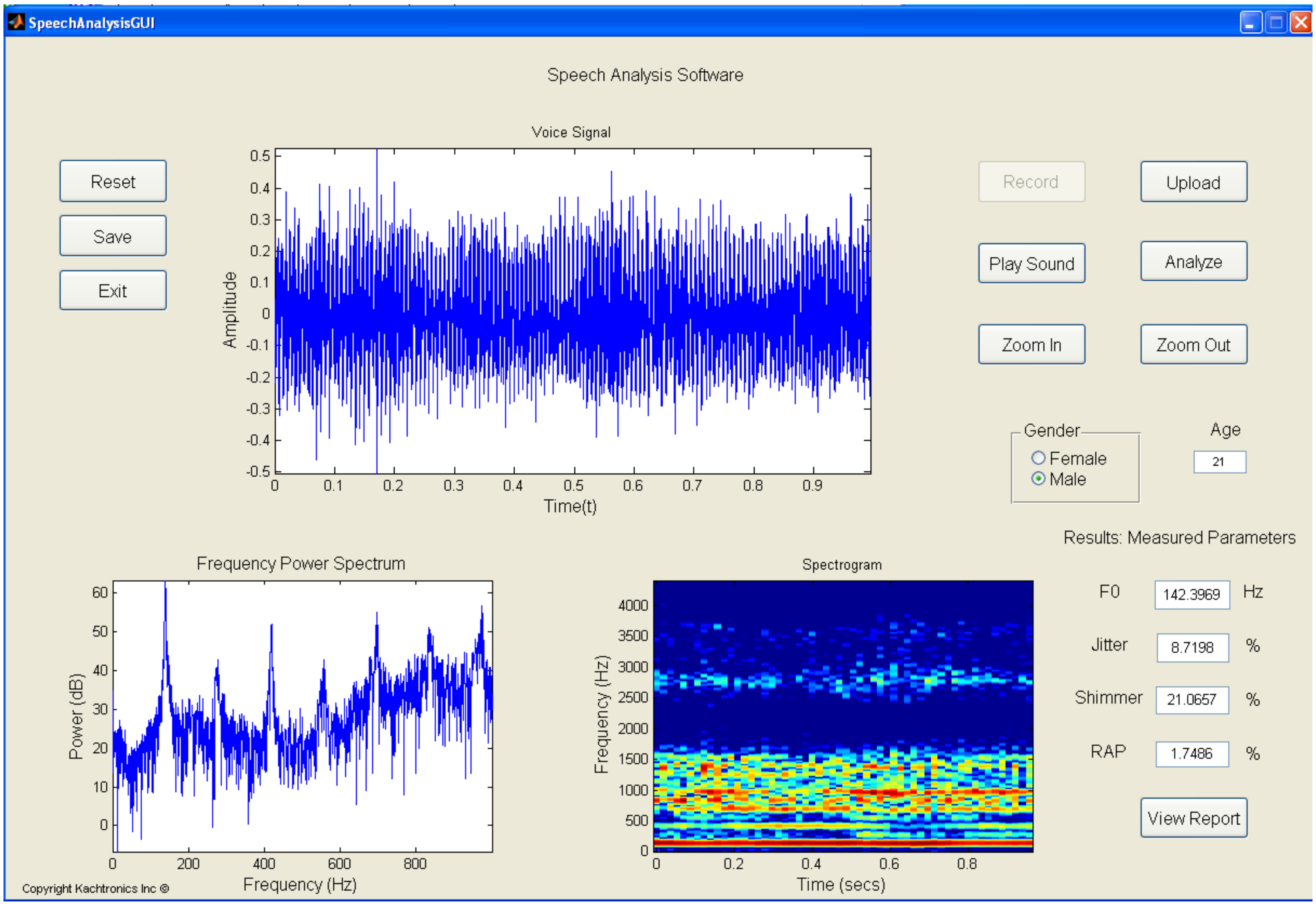
Sound Capture Design



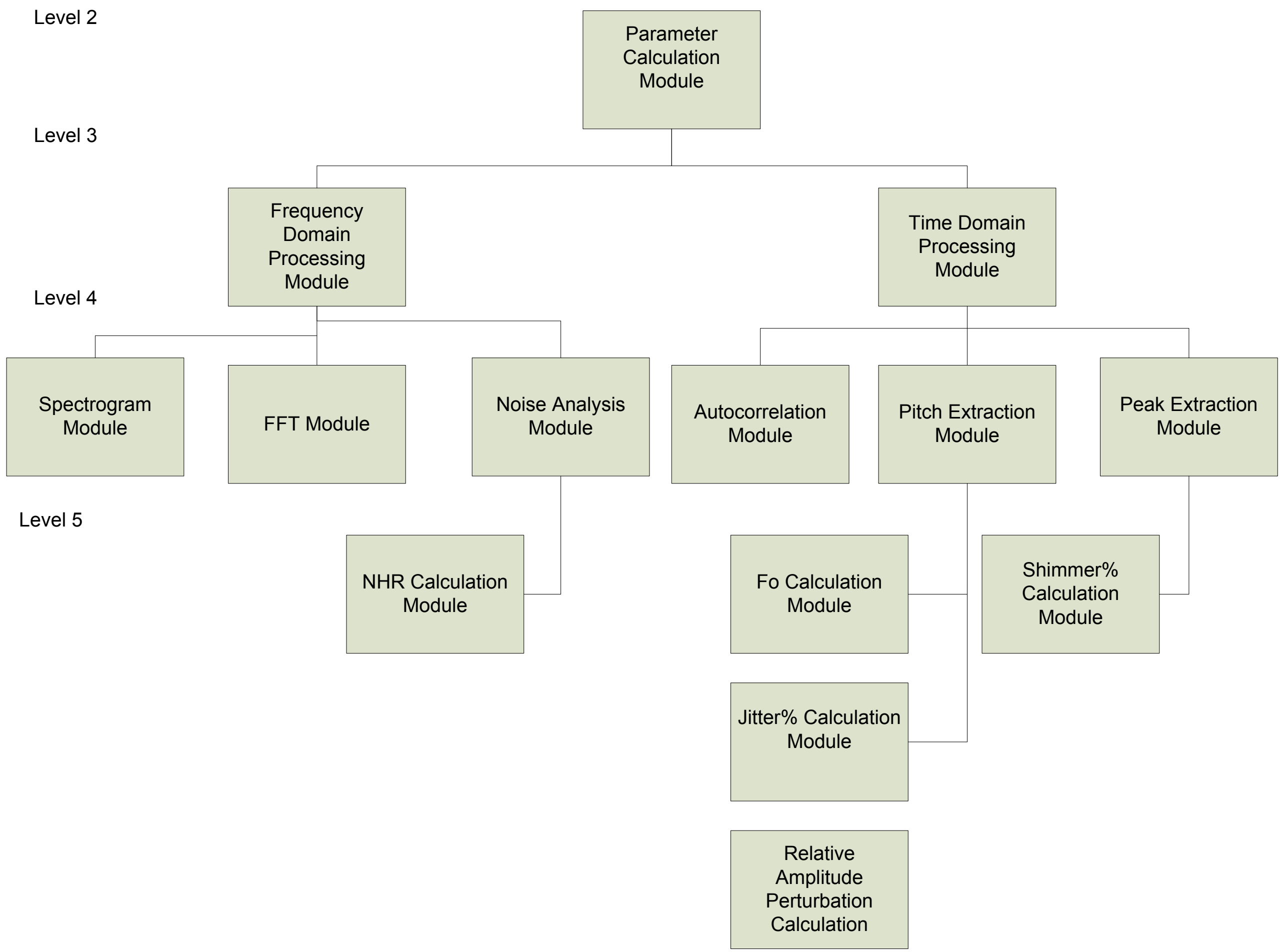
Modular Decomposition: Software



User Interface



Parameter Calculation Module



Results (Report) Module

Speech Analysis Report				
Developed By Kachtronics Inc				
All Rights Reserved				
Measured Parameter	Value	Female Norm	Male Norm	Unit
Average Fundamental Frequency (F0)	106.5421	243.973	145.223	Hz
Jitter %	1.531256	0.633	0.589	%
Shimmer %	6.659135	1.997	2.523	%
Relative Amplitude Perturbation (RAP)	0.32988	0.378	0.345	%

Testing Strategy & Results

http://giantrobots.thoughtbot.com/2008/10/17/non-test-code-is-trivial

Vocal Fold Module Testing

- Fundamental frequency measured in three ways
  - Stroboscope
  - Oscilloscope
  - Developed Software
    - Pitch Determination Code
    - Spectrogram
    - Frequency Power spectrum

Model Vocal Fold Testing

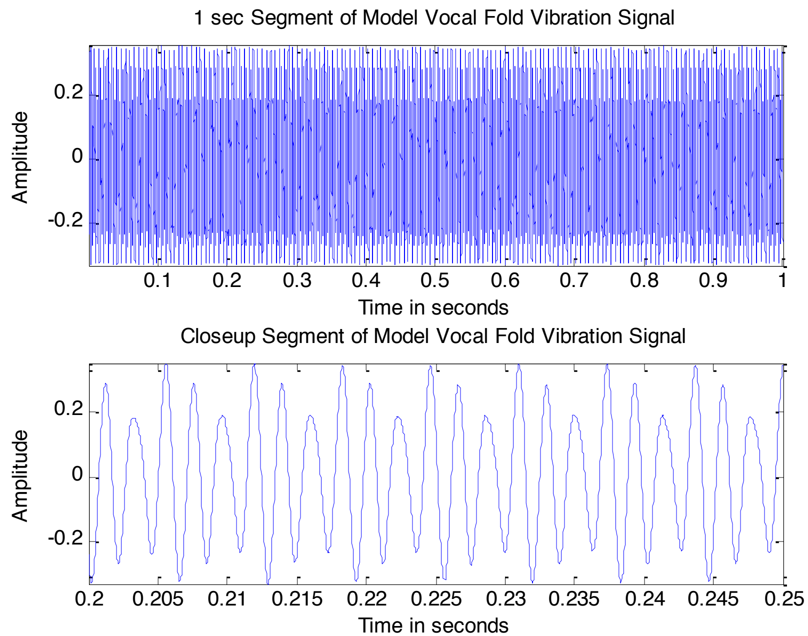
VF Model Results

- Stoboscopic Data: F0=152 Hz
- Oscilloscope : 769 Hz
  - Resonance effect, 5th resonance mode of F0
  - Tube length 17 inches

$$f_r = n * \frac{v}{2 * L}$$

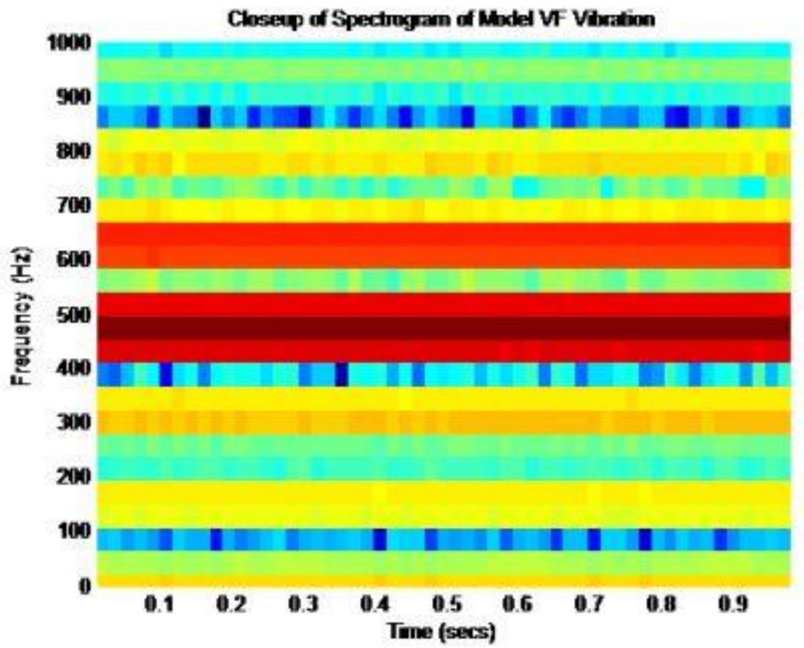
- Pitch algorithm: 154-157 Hz
  - Strong harmonic component :467 Hz
    - 3rd harmonic
- Power spectrum 154-157 Hz

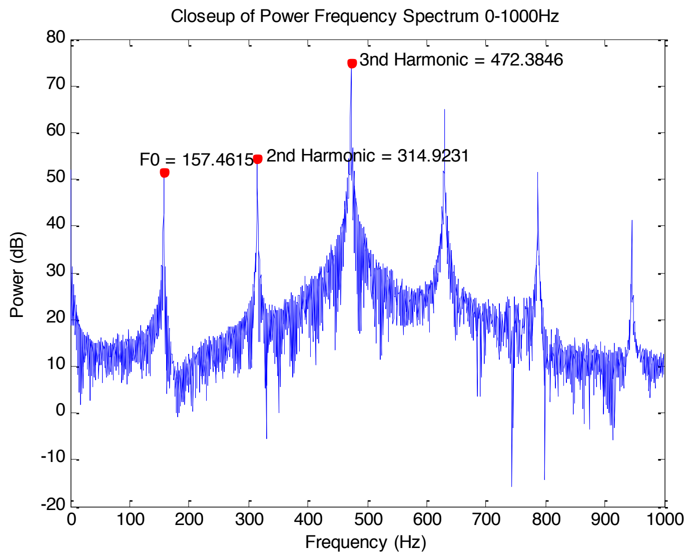
VF Results Cont’d: Time-based Measurements



Loudness	Max Amplitude	F0 (Hz)	Highest F0 (Hz)	Lowest F0 (Hz)	Jitter %	Shimmer %	RAP %	File Name
Medium	0.36	157.50	158.06	156.94	0.09	1.44	0.02	20090318 220439 Medium.wav

VF Results Cont’d: Frequency-based Measurements





Software Testing

- Three types of data
  - Randomly generated noise data using excel
  - Real human voice data
    - Kachi Voice
    - Dr. Hancock Voice
  - Vocal fold model data

Software Testing Strategy I

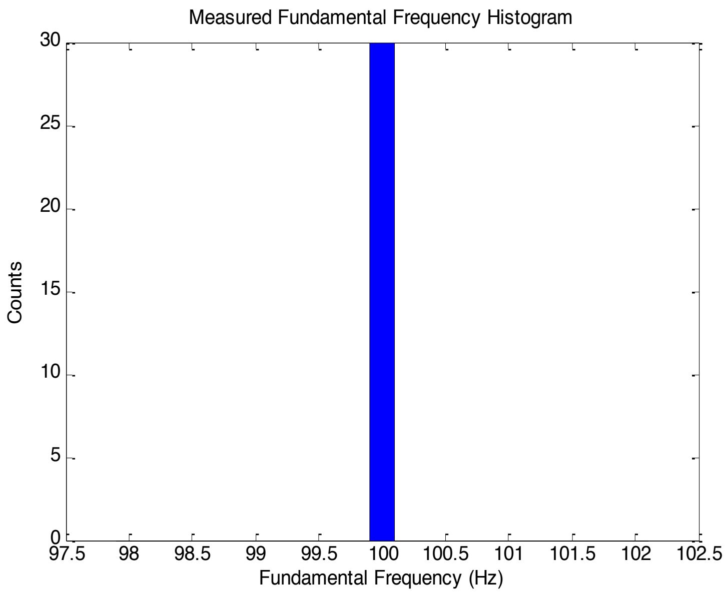
- Simulated Input Data
  - Normally distributed (Gaussian) Random Noise
    - More realistic in nature
  - Amplitude and phase noise
    - Change values of α and β
    - Simulate shimmer and jitter respectively

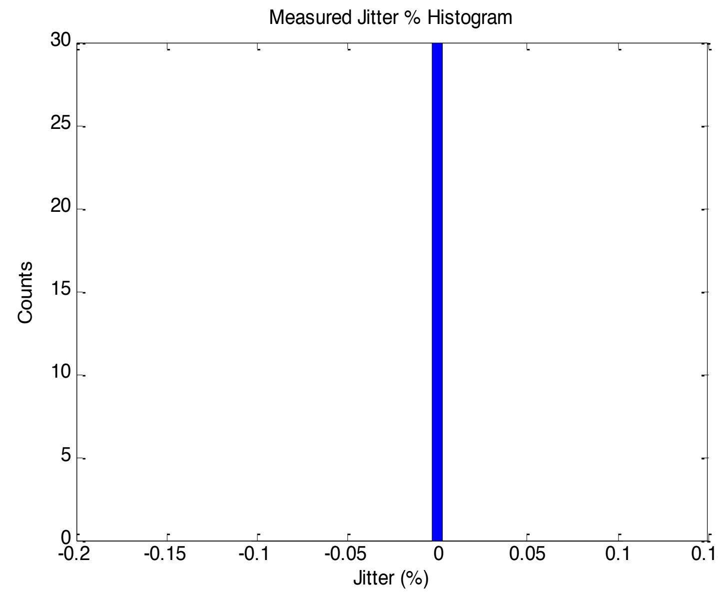
Amplitude and phase noise

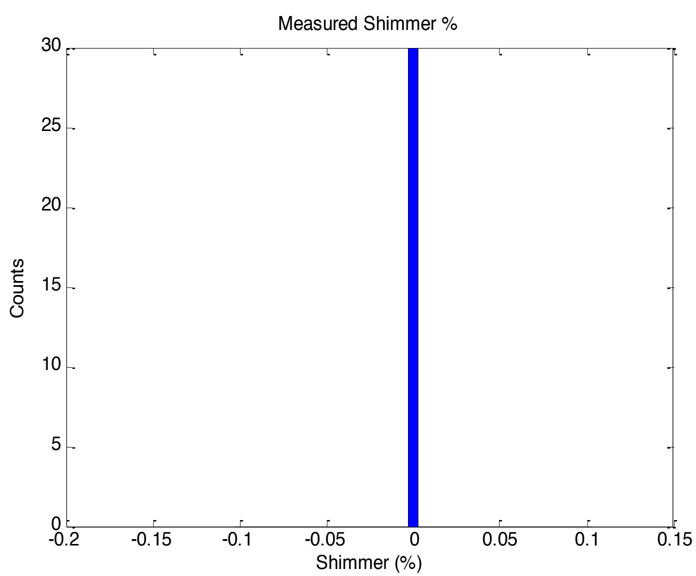
$$y = \left(1 + \alpha \cdot n_1(t)\right) \cdot \left(\sin\left(2 \cdot \pi\left(F_0 \cdot t + \beta \cdot n_2(t)\right)\right)\right)$$

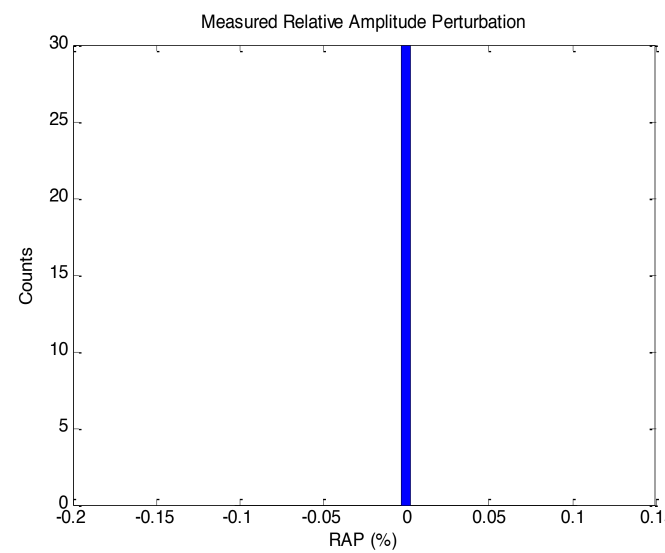
Software Testing Strategy I Results

No additive noise



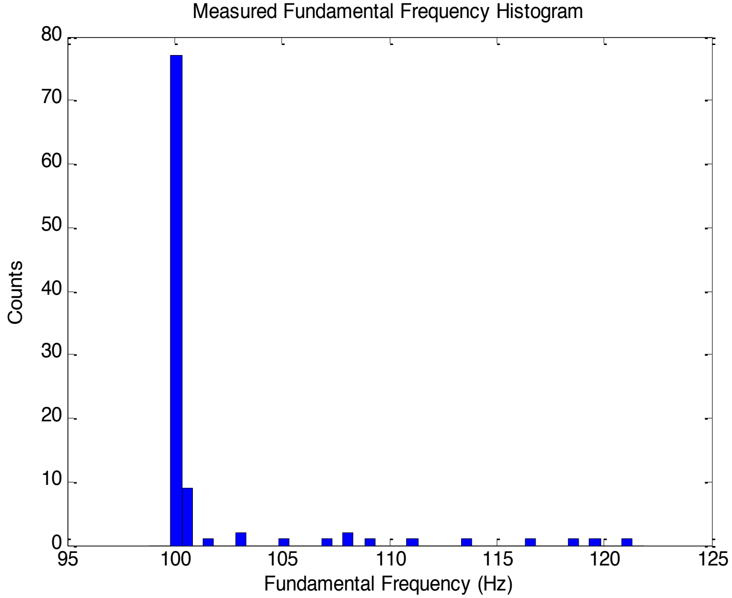


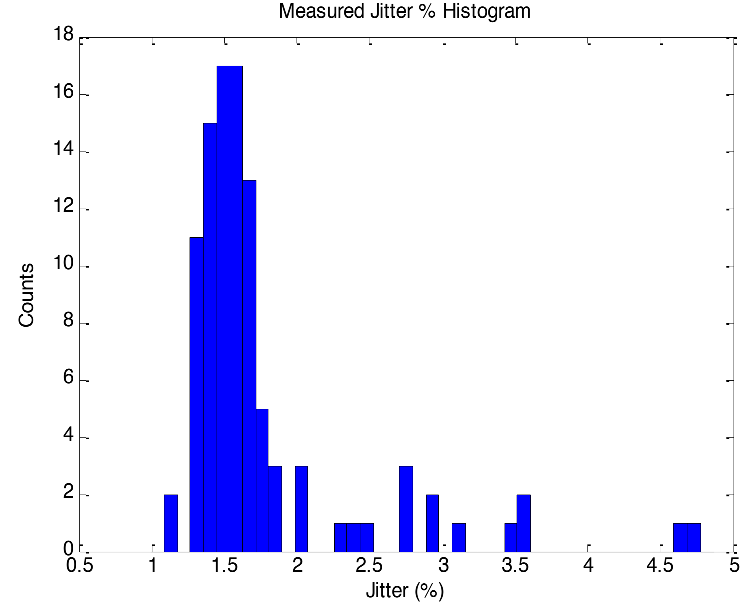


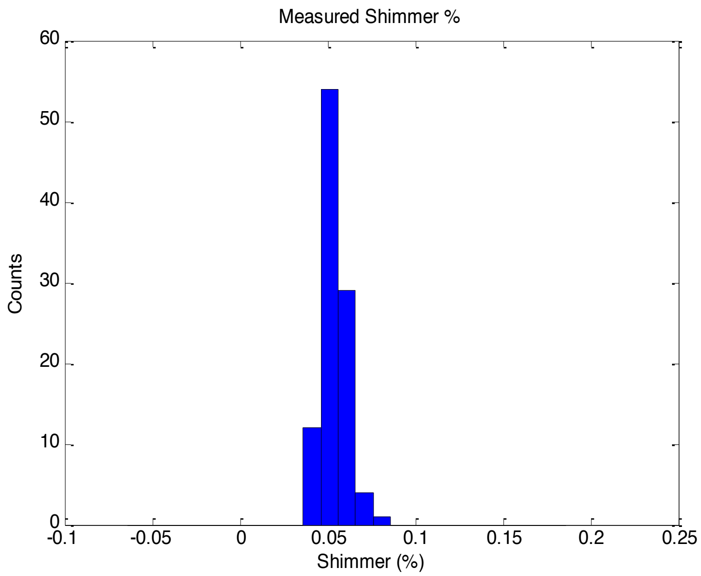


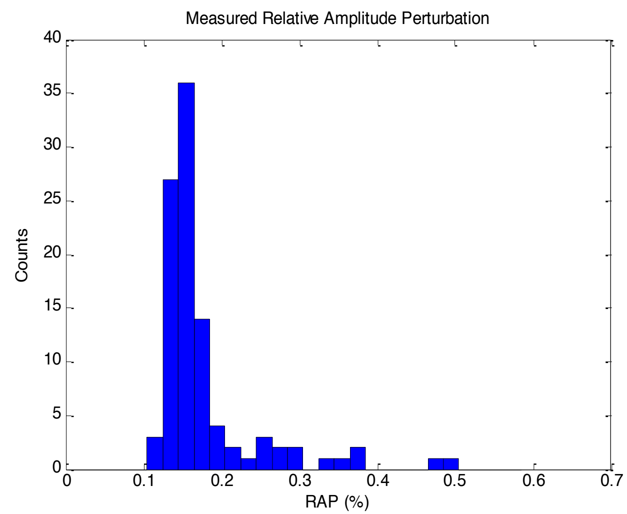
Software Testing Strategy I Results

Additive Phase and amplitude noise (α =0.005 and β=0.5)







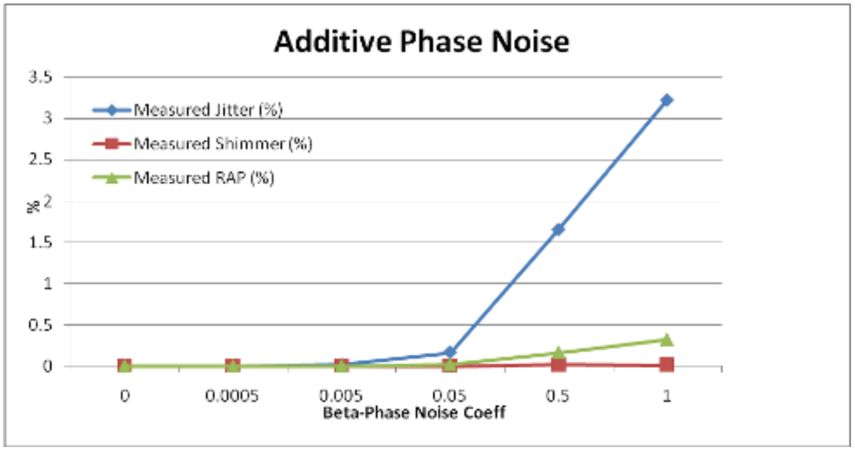


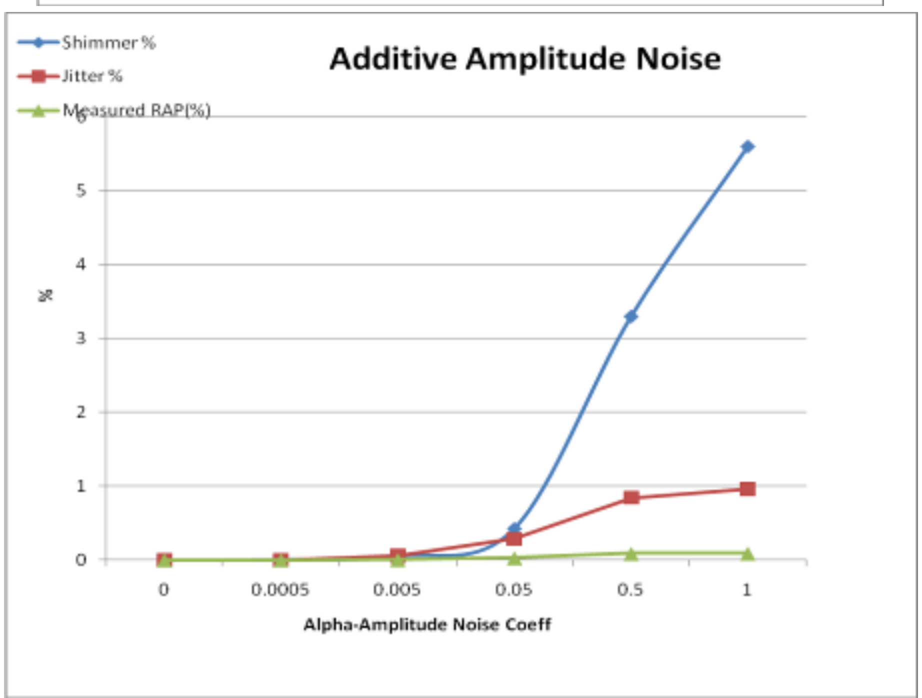
Software Testing Strategy II & Results

- Testing with human voice data
- Accurate pitch(F0) measurements
  - Low percent error
  - Known parameter
- Different results for time-based parameters than MDVP
  - Un-calibrated parameters
  - Highly dependent on implementation of algorithm
    - Results will vary for different systems

	Average Fundamental Frequency			Shimmer %			Relative Amplitude Perturbation			Jitter%
Data filename	**MDVP**	**Measured**	**Error**	**MDVP**	**Measure**	**% Error**	**MDVP**	**Measured**	**Error**	**Measured**
Hancock Sept 25 ah WAV.WAV	239.394	238.9075	0.20%	1.956	2.312	18%	0.243	0.1465	40%	0.6965
Hancock Sept 25 hoarse.WAV	311.762	331.5317	6.34%	11.007	12.9312	17%	4.754	0.9242	81%	4.5591
Hancock Sept 25 ah strained.WAV	287.225	288.4448	0.43%	4.545	5.9019	30%	1.169	0.6071	48%	2.8596
Kachi Sept 25 ah WAV.WAV	107.791	107.4438	0.32%	4.055	6.632	64%	0.472	0.0623	87%	0.3212
Kachi Sept 25 hoarse.WAV	141.143	141.5573	0.29%	13.907	21.0657	51%	7.227	1.7486	76%	8.7198
Kachi Sept 25 strained WAV.WAV	166.919	165.4212	0.90%	12.664	15.1608	20%	2.32	1.6032	31%	7.7763

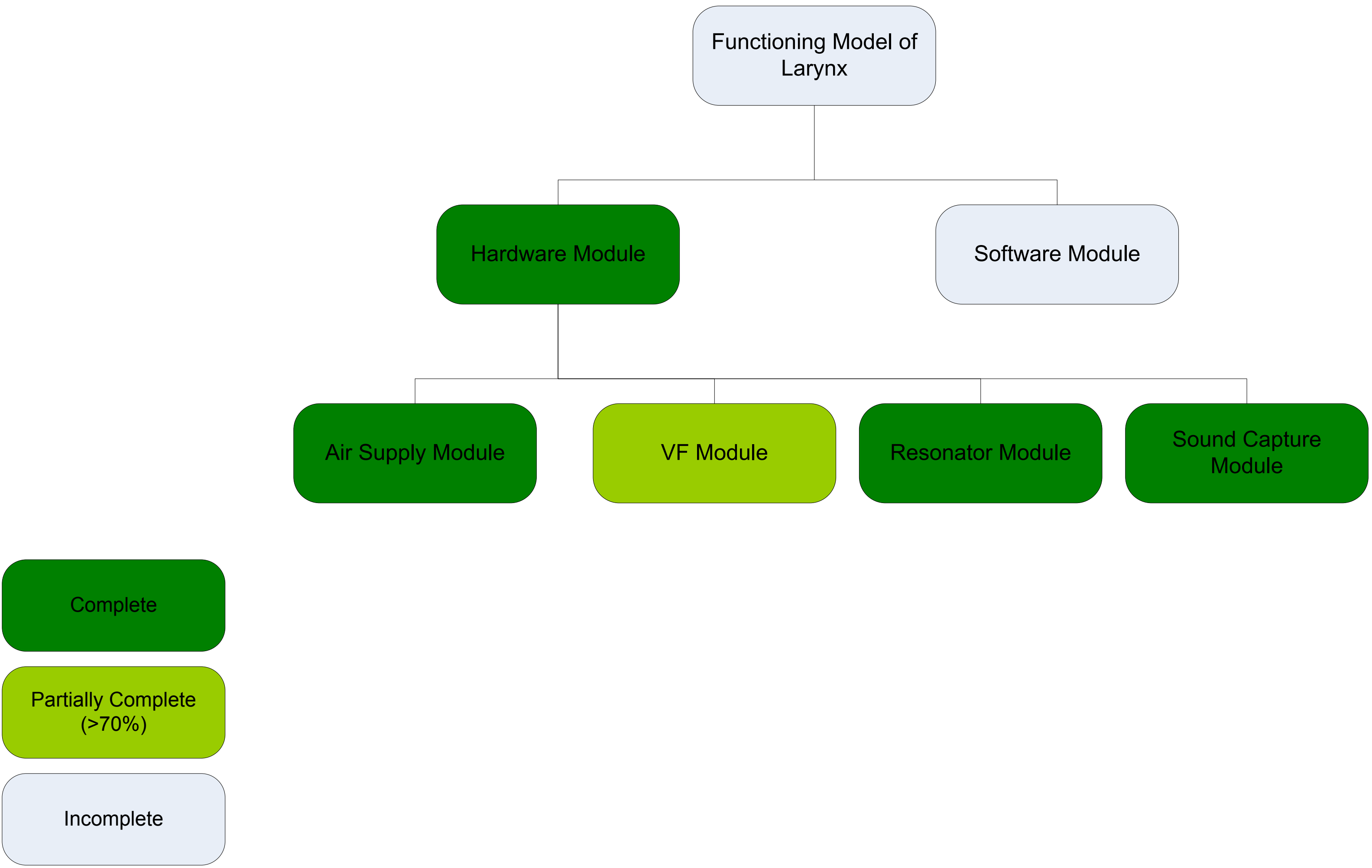
Software Testing Strategy II Results Cont’d



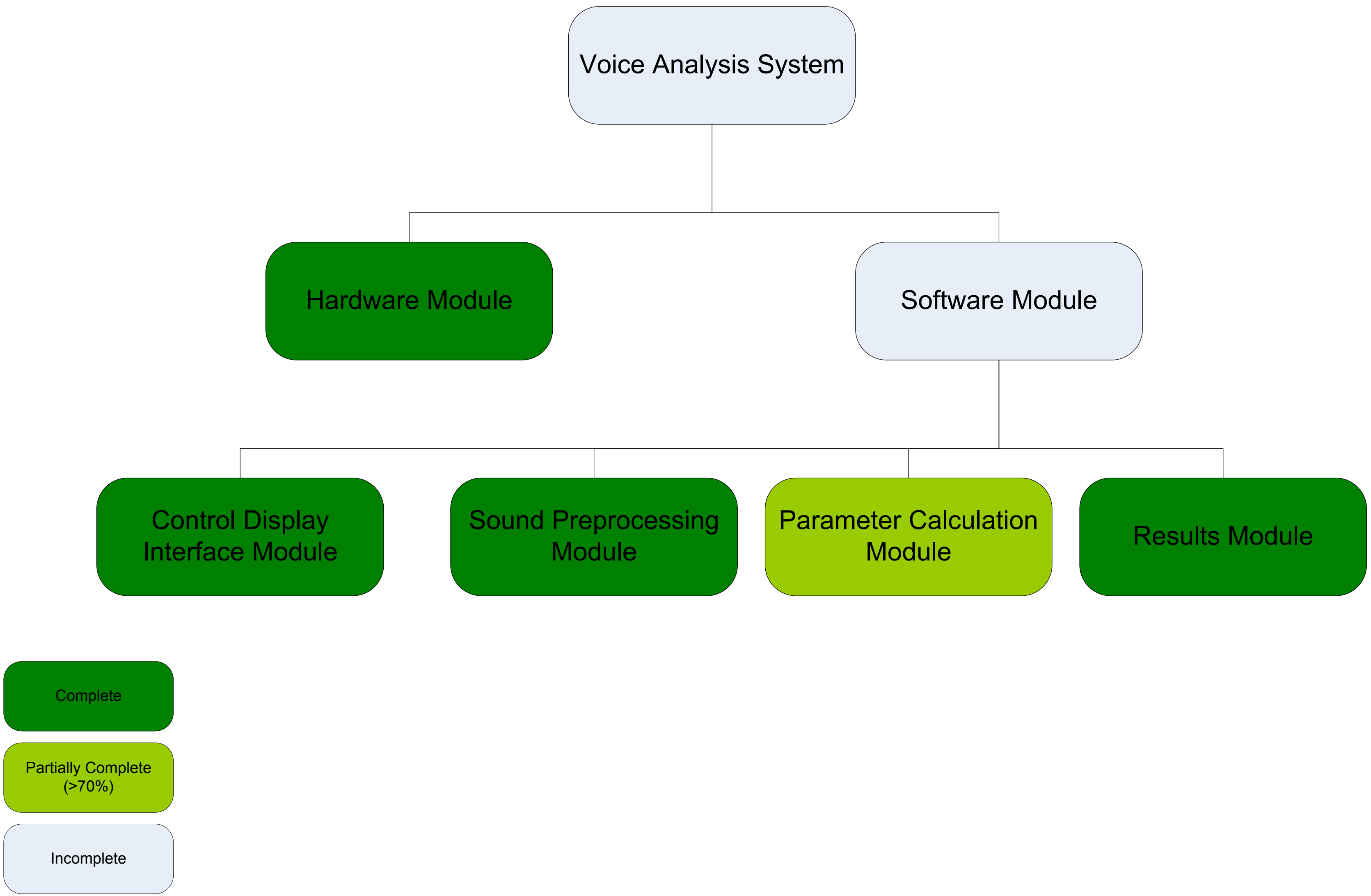


- Tested Software response to additive noise
- Calculated parameters responsive to additive noise

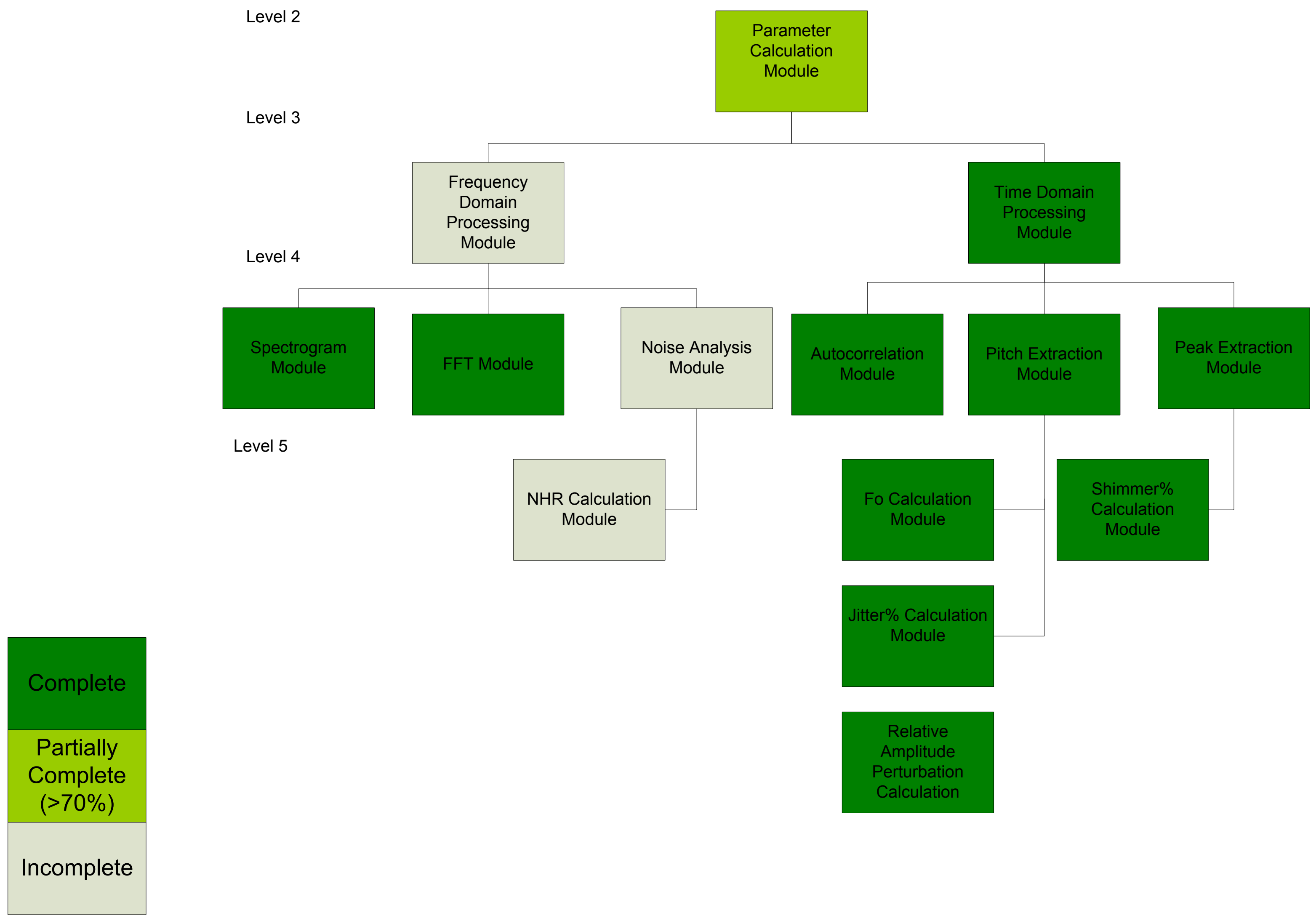
Work Accomplished: Hardware



Work Accomplished Software



Work Accomplished Software Cont' d



Conclusions

- Completed core components of project
  - VF model
  - Parameter Calculation Module
- Solid foundation established for future student(s)
- Project covered many areas of science and engineering: mechanical engineering, biology (anatomy), speech pathology and physiology, physics, electrical engineering, digital signal processing, software engineering among others

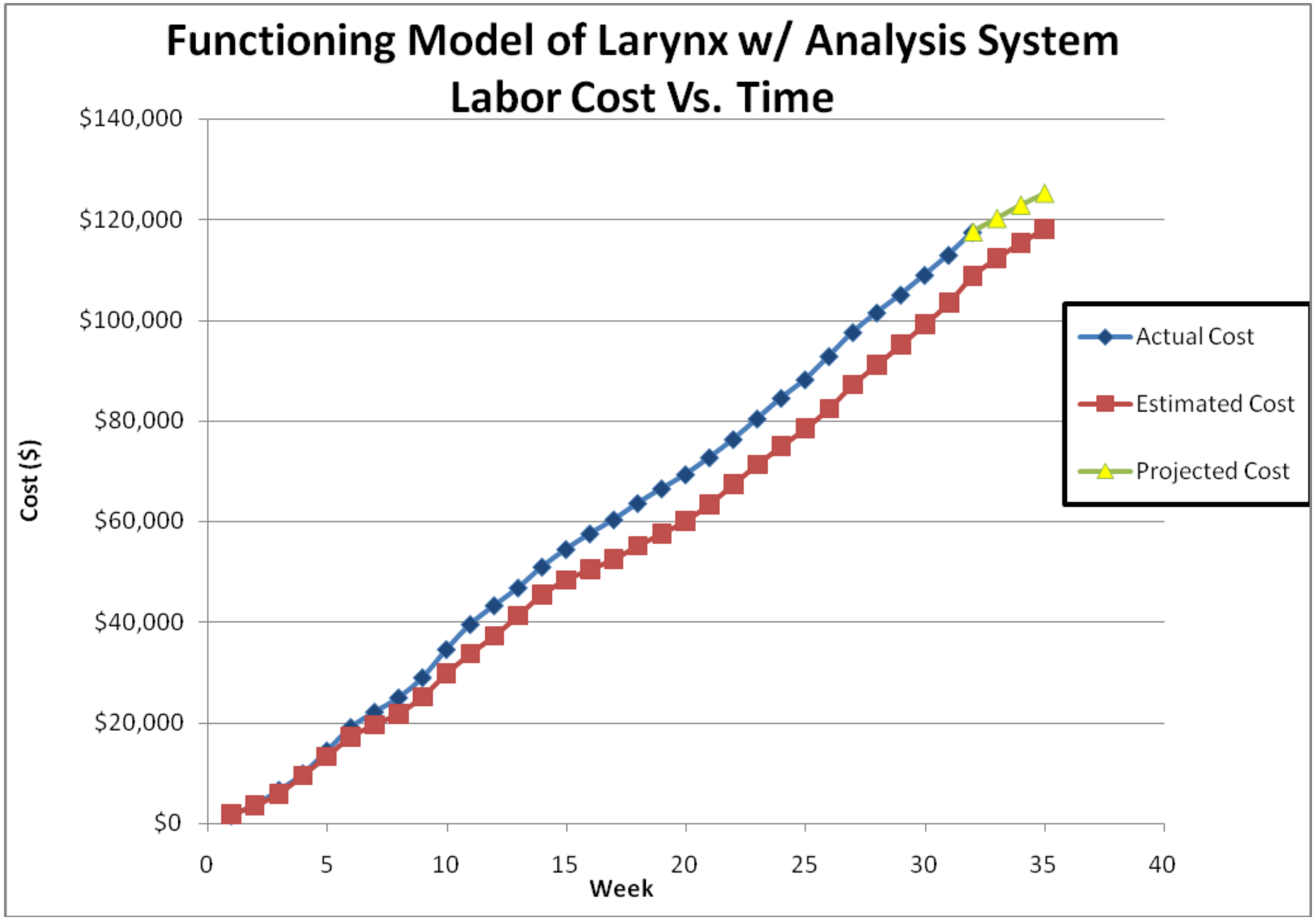
If I (or someone else) were to continue this project…

- Simulate common voice disorders
- Manipulate vocal folds to:
  - Adjust tension to change pitch of sound
  - Adduct (close) and Abduct (open)
- Enhance capability of user interface
- Compute Noise to Harmonic Ratio

Thank You

- Prof. Manuccia
- Dr. Hancock
- GW Machine Shop
- Maya Samuel for the idea for this project

Labor Cost



Functioning Model of Larynx w/ Analysis System
Labor Cost Vs. Time
$140,000
$120,000
$100,000
$80,000
$60,000
$40,000
$20,000
$0
Cost ($)
0
5
10
15
20
25
30
35
40
Week
Actual Cost
Estimated Cost
Projected Cost

Economic Analysis

PROTOTYPE COSTS:			
	Cost ($/hr)	Hours	Total Cost ($)
Project Manager	$66	350	$23,100
Design Engineer	$57	470	$26,790
Hardware Engineer	$48	519	$24,912
Software Engineer	$40	435	$17,400
Test Engineer	$36	411	$14,796
Technical Writer	$30	375	$11,250
Speech Pathologist (Consultant)	$100	25	$2,500
Overall Hours:	2585		
Overall Cost:	$120,748		
Multiplier:		2.8	
Labor Charged to Contract:		$338,094	
Parts Total:		$100.00	
Machine Shop, PCB fab. and pop.:		$50	
Sub-Total:		$150	
Pass-Through Fee:		1.05	
Total:		$158	
Cost of Labor:		$338,094	
Cost of parts & external services:		$158	
Total prototype cost:		$338,252	

PRODUCTION COSTS:		
Manufacturing Process:		
Salary ($/hr)	20	
Hours	200	
Total:	$4,000	
Software Testing:		
Salary ($/hr)	15	
Hours	200	
Total:	$3,000	
Total Engineering	$7,000	
Multipler:		2
Production Labor Costs:		$14,000
Number of Units:		1,000
Parts Cost:		$120,000
Printing and Packaging Cost:		$120,000

Production Non-labor Costs:	$240,000
Overhead Cost Multiplier:	1.4
Profit Multiplier:	1.2
Total Production Cost:	$426,720
Production Cost per Unit:	$426.72

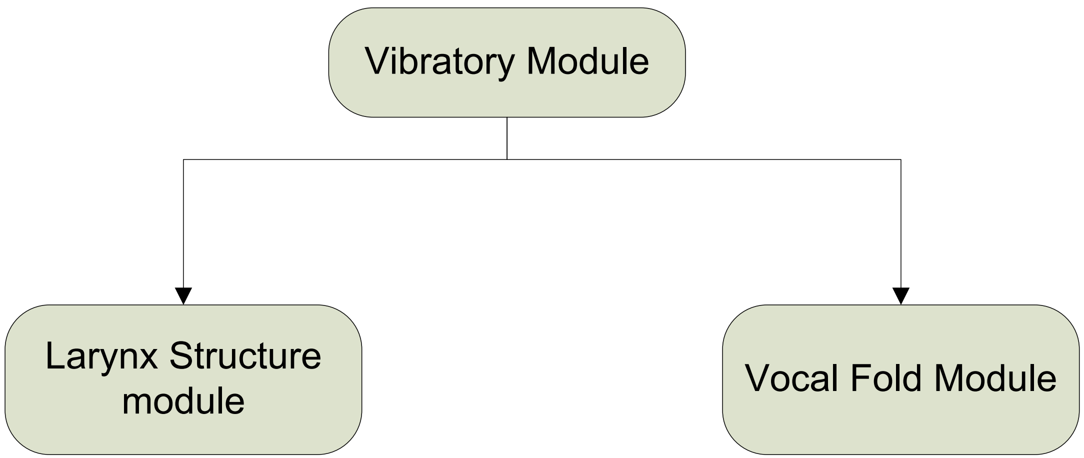
Table 4: Production Cost for 1000 units

PROJECT COST:		
Prototype Cost:	$338,252	
Production Cost:	$393,120	
Total Project Cost:	$731,372	
Total Project Cost per Unit:	$731.37	

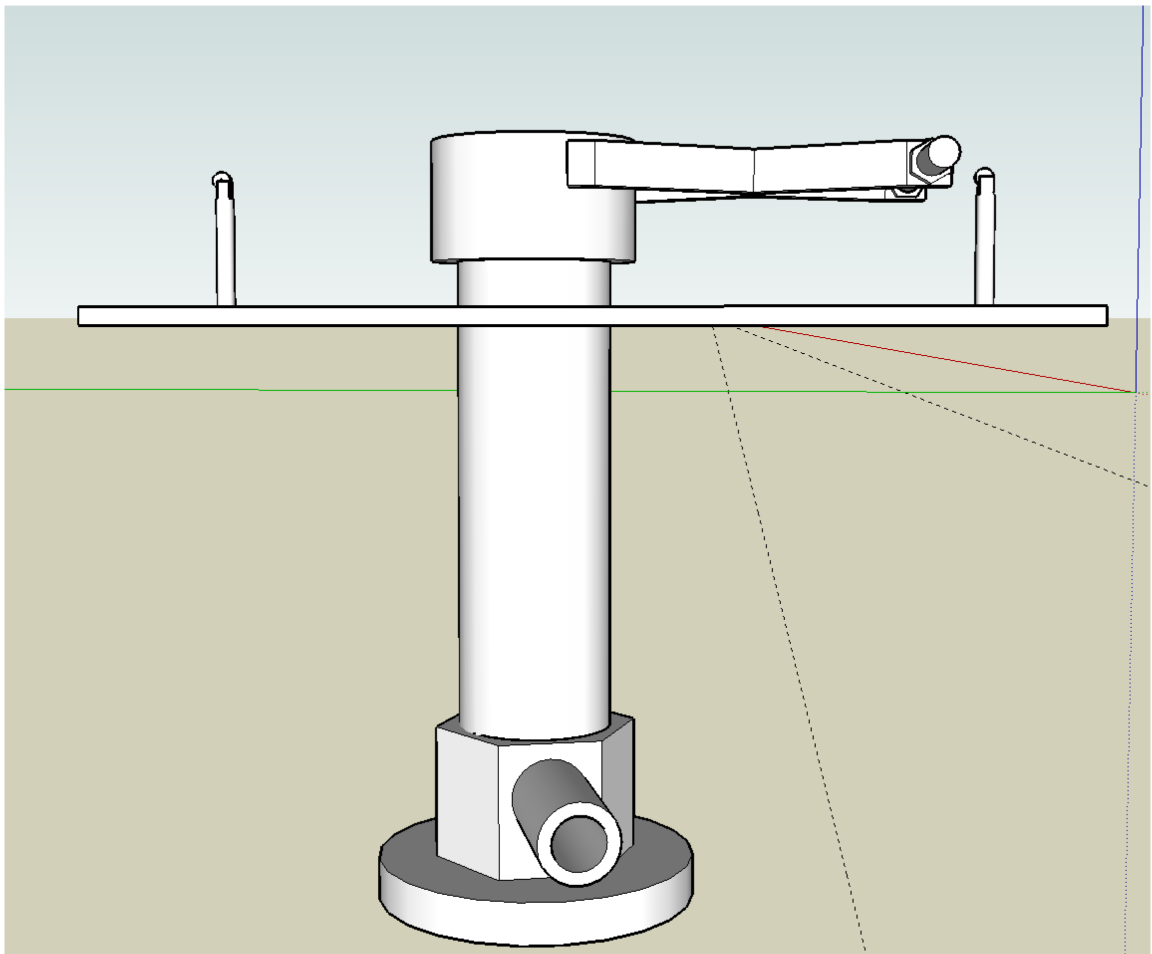
Table 5: Project Cost for 1000 units

COST OF DISTRIBUTION:		
Cost per Unit:	$762.51	
Wholesaler Price:	$915.01	1.2 Multiplier
Retailer Price:	$1,372.51	1.5 Multiplier

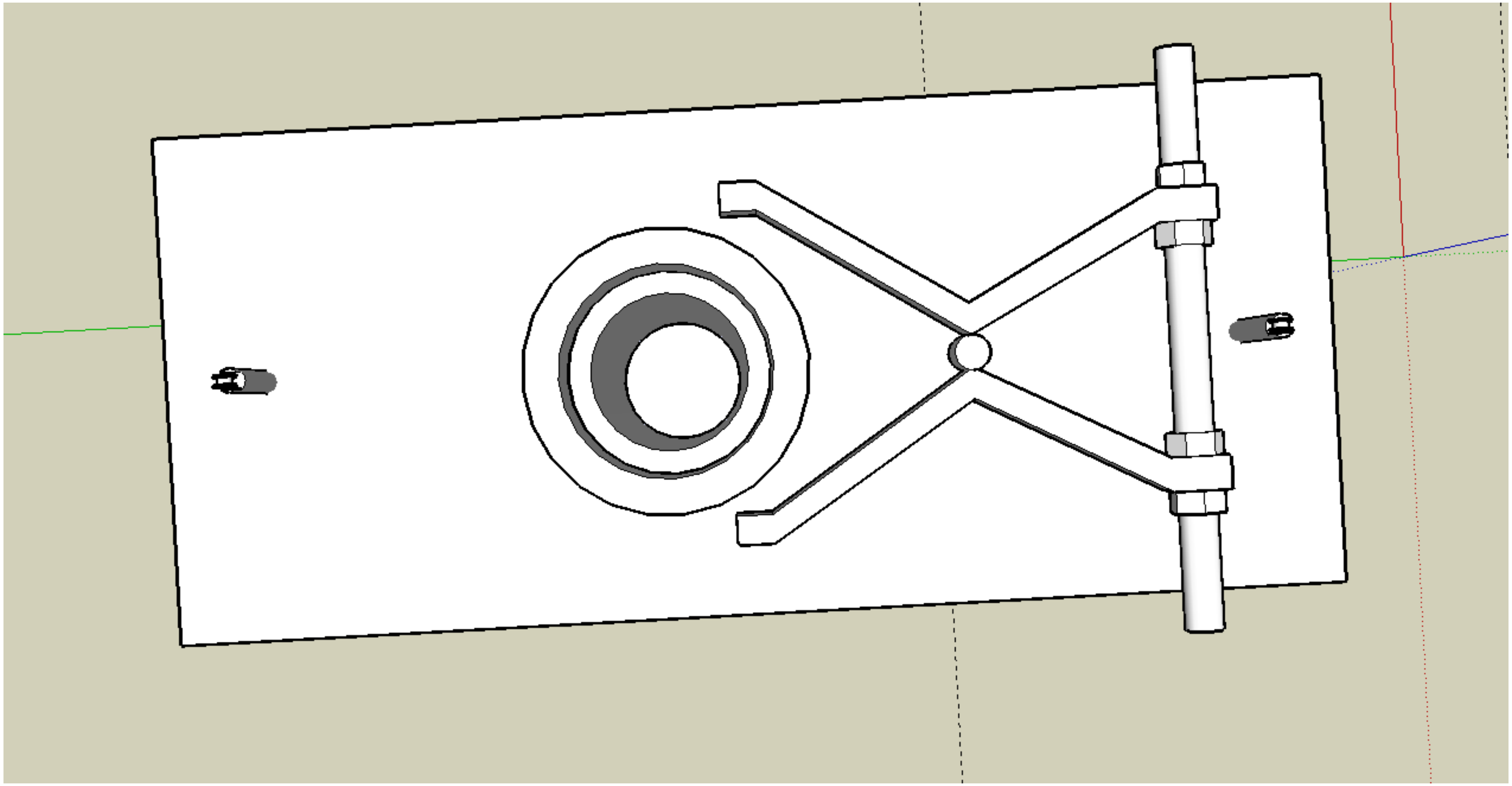
Vibratory Module Decomposition



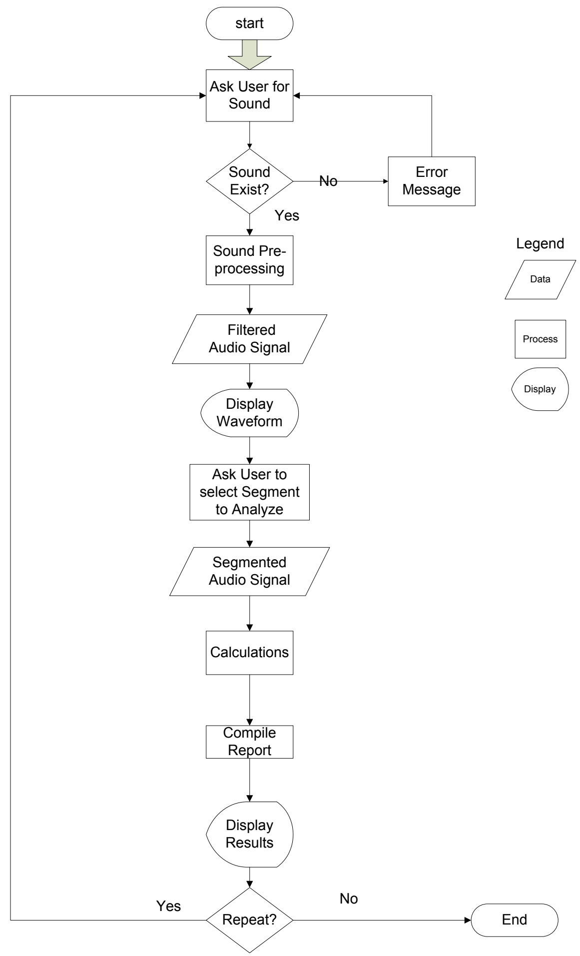
Vocal Fold Manipulation



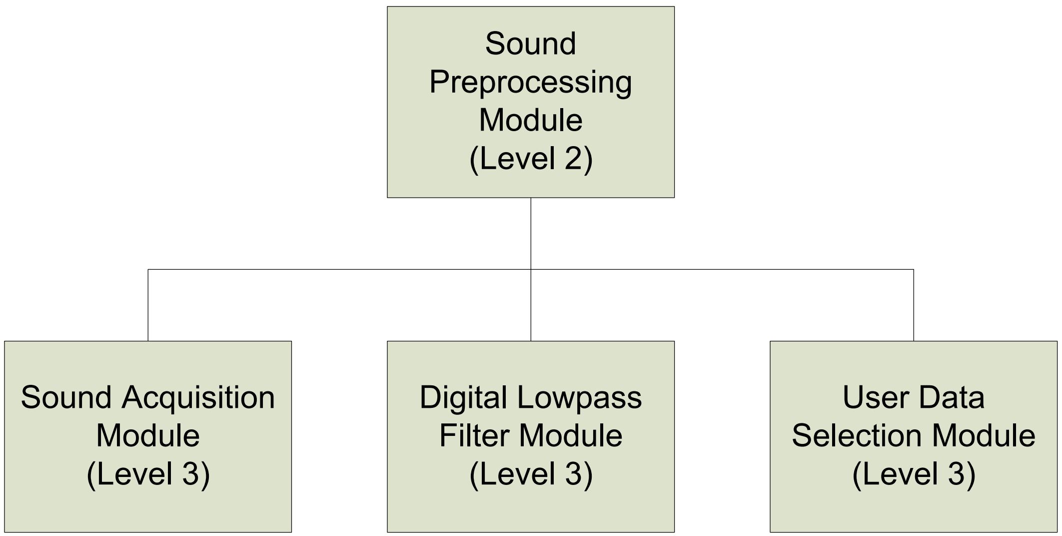
Vocal Fold Manipulation



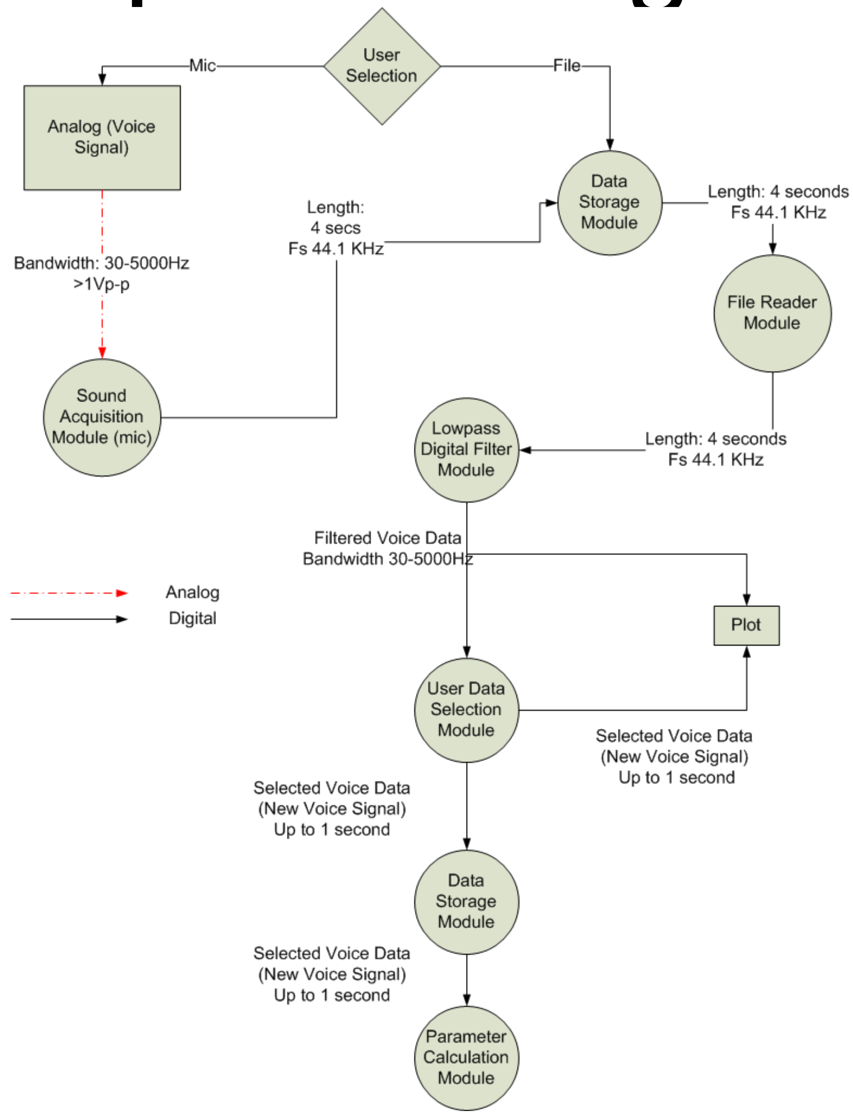
Software Execution Flow



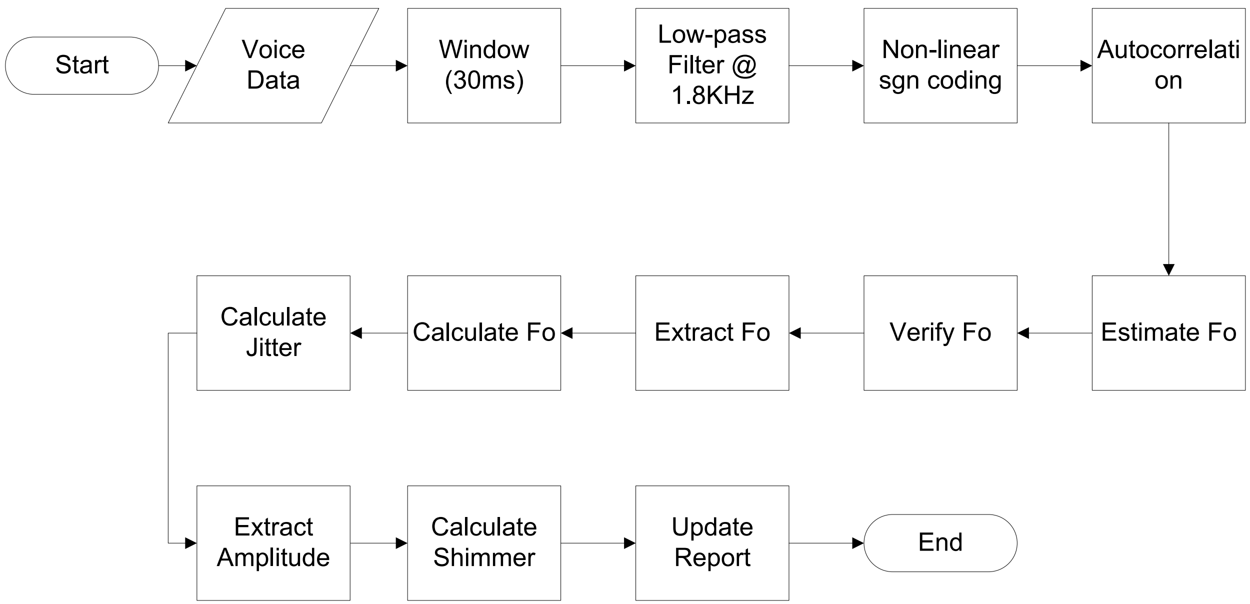
Sound Preprocessing Decomposition



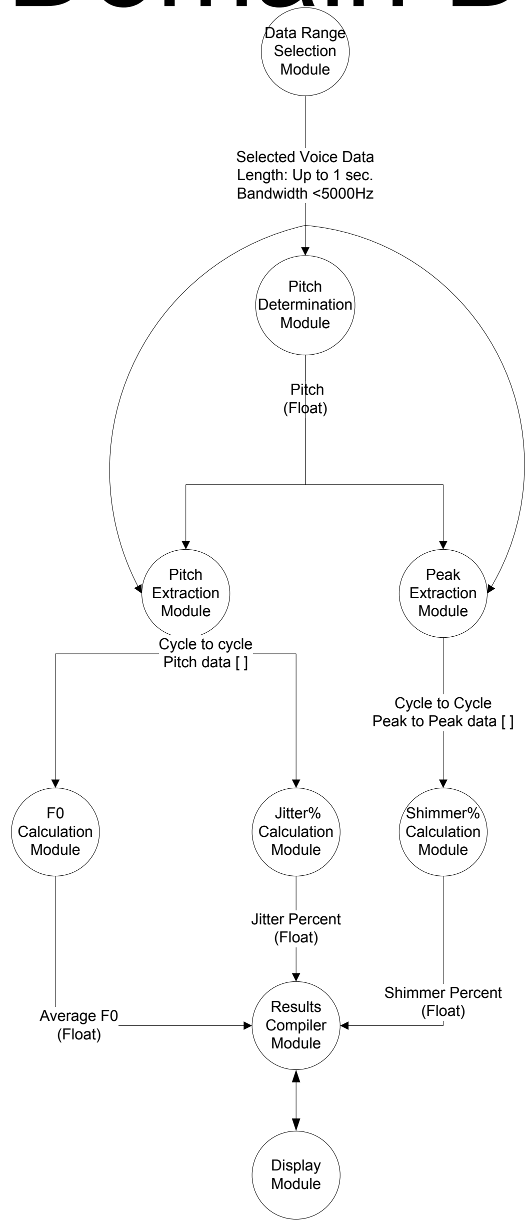
Sound Preprocessing Data Flow



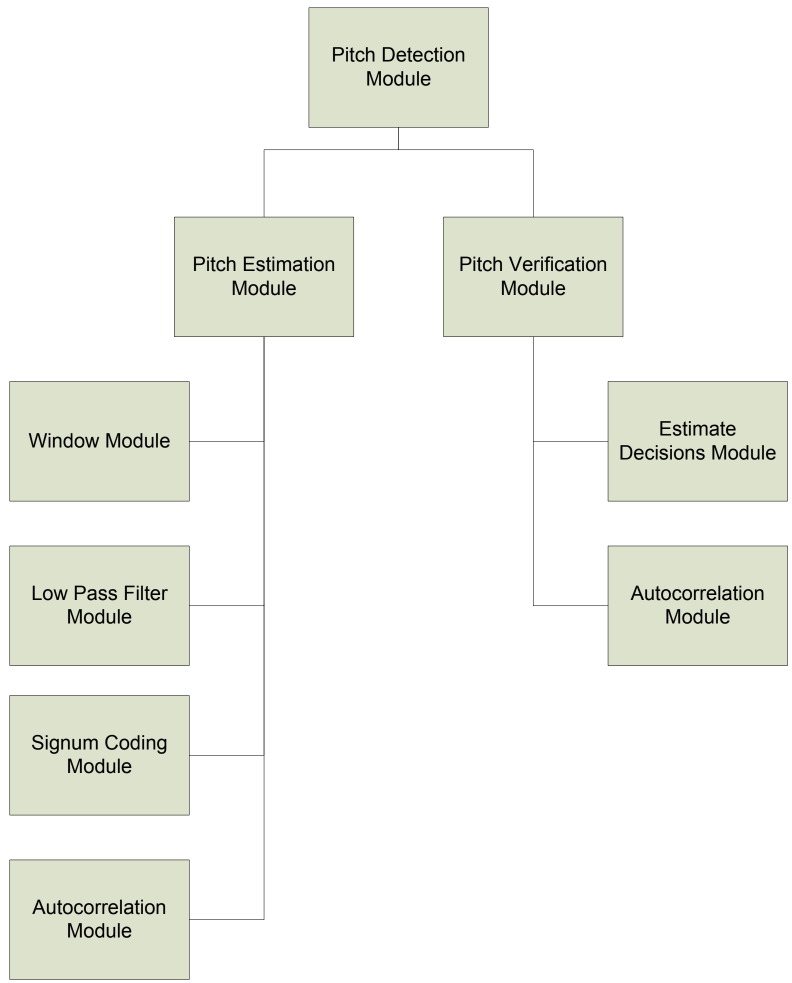
Time Domain Processing Execution Flow



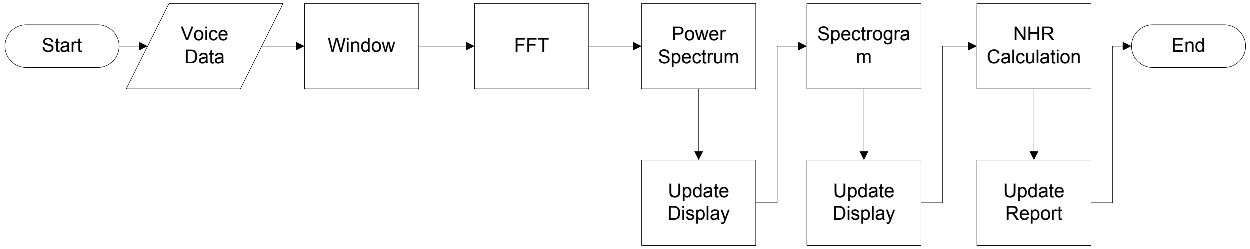
Time Domain Data Flow



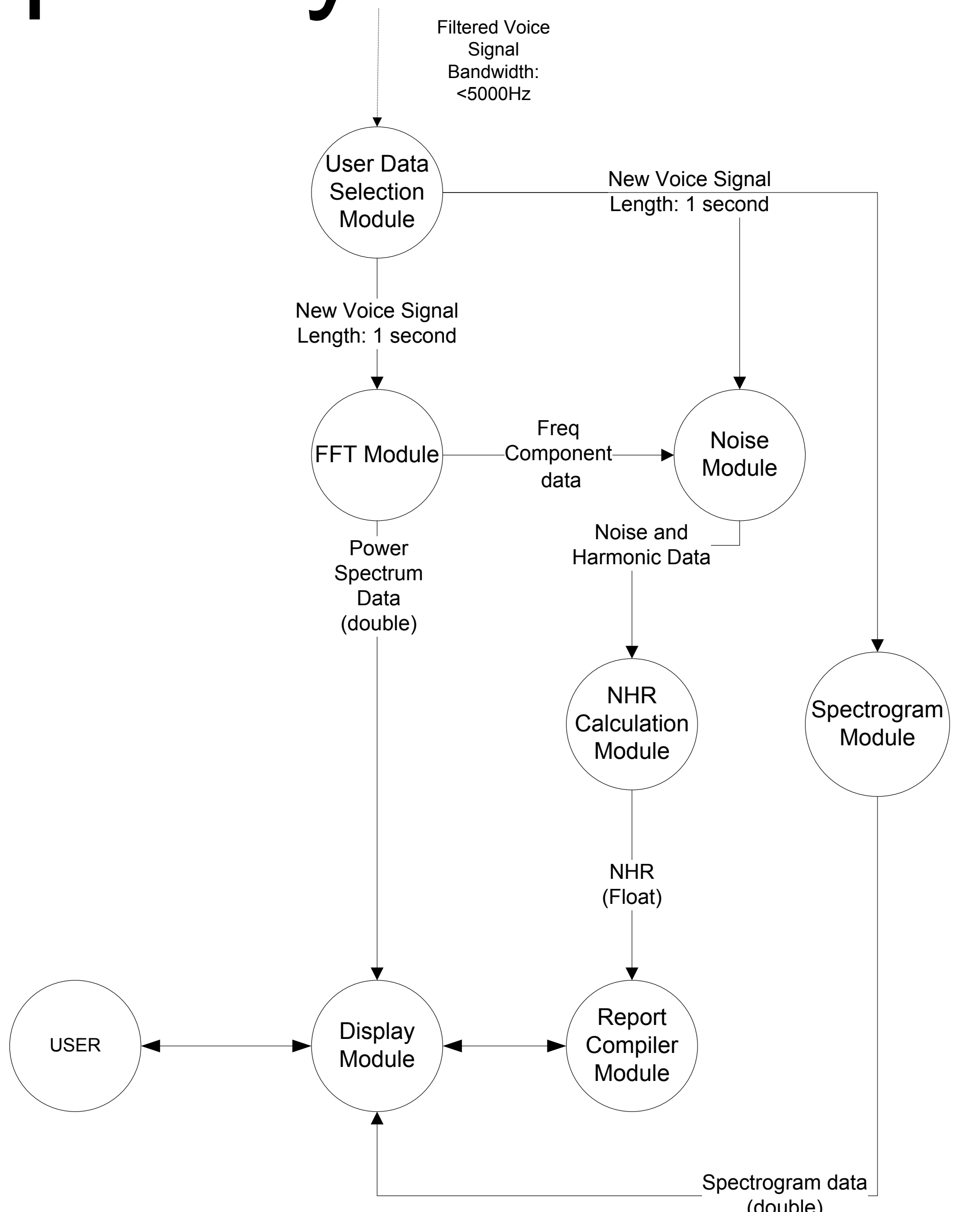
Pitch Determination Module Decomposition



Frequency Domain Execution Flow



Frequency Domain Data Flow



Calculated Parameter Equations

$$Shimmer = 100\% * \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|A^{(i)} + A^{(i+1)}|}{\frac{1}{N}\sum_{i=1}^{N} A^{(i)}}$$

$$Jitter = 100\% * \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|T^{(i)}+T^{(i+1)}|}{\frac{1}{N}\sum_{i=1}^{N} T^{(i)}}$$

$$F_0 = \frac{1}{N}\sum_{i=1}^{N} F_0^{(i)}$$

$$RAP = \frac{\frac{1}{N-2}\sum_{i=2}^{N-1}\left|\frac{To^{(i-1)} + To^{(i)} + To^{(i+1)}}{3} - To^{(i)}\right|}{\frac{1}{N}\sum_{i=1}^{N} To^{(i)}}$$

where: $To^{(i)}$, $i=1,2...N$ - extracted pitch period data,

$N = PER$ - number of extracted pitch periods.