USING SPECTRAL ANALYSIS TO EVALUATE FLUTE TONE QUALITY

A Thesis

presented to

the Faculty of California Polytechnic State University

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Computer Science

by

Ron Yorita

December 2014

COMMITTEE MEMBERSHIP

TITLE: Using Spectral Analysis to Evaluate Flute Tone Quality

AUTHOR: Ron Yorita

DATE SUBMITTED: December 2014

COMMITTEE CHAIR: John Clements, Ph.D.
Associate Professor of Computer Science

COMMITTEE MEMBER: Phillip Nico, Ph.D. .
Associate Professor of Computer Science

COMMITTEE MEMBER: Clark Turner, Ph.D.
Professor of Computer Science

ABSTRACT

Using Spectral Analysis to Evaluate Flute Tone Quality

Ron Yorita

Many skilled flutists place a high priority on “good” tone quality, or timbre. Timbre can be defined as the audible difference in character that a listener perceives for two notes played at the same pitch [13, 29]. Different timbres are determined by the combination and balance of harmonics that comprise a note [2, 38, 47]. Unlike pitch and rhythm, timbre is difficult to objectively quantify. This project explores (1) how tone quality is described by skilled flutists, (2) whether the harmonic spectrum has some correlation with tone quality, (3) whether certain harmonic spectra are preferred, or considered “good”.

Thirty-one flutists ranging from high school students to professionals were recorded. A set of samples was used in surveys and interviews to capture descriptors and ratings of tone quality. All of the recorded samples were analyzed using application programs, Harmonic Analysis Tools (HAT), created for this study. HAT uses digital signal processing techniques to produce “spectral signatures”. The signatures consist of the harmonic content, pitch, and amplitude of a sample. In the future, with further development, HAT may be a useful tool for musicians for tone development in the practice room.

The outcome of this research is a baseline set of some often used descriptors. In addition, results showed some correlation between harmonic spectra and descriptors. There were also trends in preferences with respect to certain spectral characteristics. An unexpected finding was that University students showed divergent timbre preferences compared to highly experienced flutists.

ACKNOWLEDGMENTS

Many thanks to the musicians that took part in the recording process, and to the individuals that participated in the surveys.

Thank you to faculty, staff, students, family, and friends who have supported my efforts during my time at Cal Poly, Southern Oregon University, and Fort Lewis College.

My heartfelt appreciation to my wife, Randi, who has been my partner through all the adventures.

TABLE OF CONTENTS

List of Tables ix

List of Figures x

1 Introduction 1

2 Domain Specific Background 4
- 2.1 Flute Tone and Harmonics 4
- 2.2 Flute Construction Material and Tone Quality 7
- 2.3 The Flutist and Tone 8
- 2.4 Dynamics and Tone 9
- 2.5 Vibrato 9
- 2.6 Acoustics and Psychoacoustics 10
- 2.7 The Flutists Perspective 12

3 Technical Background 13
- 3.1 Digital Signal Processing 13
- 3.2 Spectral Analysis 15
  - 3.2.1 Fourier Transform Family 15
  - 3.2.2 Windowing Functions 17
  - 3.2.3 Overlapping Frames 18
  - 3.2.4 Frame Size and FFT Bins 18
- 3.3 Pitch Detection 19

4 Related Work 21
- 4.1 Flute Tone Analysis Using Spectra 21
- 4.2 Spectral Analysis For Musicians 23

5 HAT Implementation 26
- 5.1 Overview 26
- 5.2 Functionality and Displays 27

5.2.1 Real-Time Analysis . . . 28
5.2.2 Ratio . . . 31
5.3 HAT Implementation . . . 33
5.3.1 HAT Components . . . 33
5.3.2 Implementation Alternatives . . . 35
6 Methodology 37
6.1 Gathering Flute Tones . . . 37
6.2 The Survey . . . 40
6.2.1 Selecting Flute Tones for the Survey . . . 40
6.2.2 Designing the Survey . . . 42
7 Results and Analysis 44
7.1 General Observations . . . 44
7.1.1 Recording Sessions . . . 44
7.1.2 Spectra Variability . . . 45
7.1.3 Spectra and Pitch . . . 46
7.1.4 Non-flute Spectra . . . 48
7.2 Survey Results . . . 48
7.2.1 "Skilled Flutist" Demographics . . . 48
7.2.2 FL10s Descriptors . . . 49
7.2.3 FL10s Ratings . . . 53
7.2.3.1 G4 Spectra . . . 54
7.2.3.2 D4 Spectra . . . 56
7.2.3.3 B4 and D5 Spectra . . . 57
7.2.3.4 Trends in Ratings and Harmonic Signatures . . . 59
7.2.4 Cal Poly Ratings . . . 59
7.3 Vibrato and Dynamics Analysis . . . 62
7.3.1 Vibrato . . . 62
7.3.2 Dynamics . . . 65
7.3.3 Syrinx . . . 68
8 Conclusions 71

8.1 FL10s 71
8.2 CPM 72
8.3 Other Observations 73
8.4 Future Work 74
    8.4.1 Acquisition of Flute Tones 74
    8.4.2 Additional Survey Data 74
    8.4.3 Other Analysis 75
    8.4.4 Final Words 75
Bibliography 76

Appendices

A Recording Equipment 82
B Recording Details 85
C Survey Questions 88
D Demographics 91
D.1 Full Flute List 91
D.2 FL10s 92
D.3 Favorite Flutists 92
E Spectra 94
E.1 Flute Spectra - Repeated Notes 94
E.2 Flute Spectra - Range Of Pitches 96
E.3 Oboe, Trumpet, and Clarinet Spectra 97
F Rating Distributions 98
G Flute List Descriptor Correlation 99

LIST OF TABLES

7.1 FL10s Descriptor Usage 50
7.2 FL10s Descriptors 51
7.3 FL10s Descriptors for “great” or “poor” ratings 52
7.4 FL10s summary of ratings and descriptors 54
7.5 FL10s and CPM ratings and p-values 61

LIST OF FIGURES

2.1 Air jet across embouchure hole [54] 5
2.2 Flute standing waves [12] 5
2.3 Fletcher-Munson equal loudness curves [32] 11

3.1 Analog to digital signal conversion [32] 13
3.2 Sampling rates [25] 14
3.3 Sine waves comprising a sound wave [32] 16
3.4 Windowing function [43] 17
3.5 Overlapping window frames [43] 18

4.1 Oscilloscope graphs [49] 21
4.2 Lip to embouchure hole gap effect on harmonics [9] 22
4.3 Harmonics at different pitches and loudnesses [14] 24

5.1 Audio file or microphone to spectral signature 27
5.2 HAT RTA display 28
5.3 HAT Harmonics Color Coding 29
5.4 HAT RTA Line display 30
5.5 HAT RTA Line Chart - correlating harmonics, pitch, and loudness 31
5.6 HAT Ratio display 32
5.7 RTA 32
5.8 Ratio 32
5.9 Ratio for flutist C8 33
5.10 Ratio Compare 33
5.11 HAT components 34

6.1 Notes for recording 37

7.1 Flutist P3 playing B4 46
7.2 Flutist U3 playing B4 46
7.3 Flutist U7 and P10 playing D4 47
7.4 Flutist U7 and P10 playing B5 47
7.5 Flutist P2 playing C4 48
7.6 Flutist U5 playing C4 48
7.7 FL10s - teaching experience 49
7.8 FL10s rating distribution 51
7.9 FF10s and Flute List ratings 53
7.10 FL10s g4 spectra ratings 55
7.11 FL10s d4 spectra ratings 57
7.12 FL10s survey results for pitch B4 58
7.13 FL10s survey results for pitch D4 58
7.14 Cal Poly ratings 60
7.15 FL10s vs. Calpoly sorted ratings 61
7.16 Straight-tone and vibrato 63
7.17 Pitch and vibrato 64
7.18 Dynamics, straight-tone D4 and D5 66
7.19 Dynamics, straight-tone and vibrato 67
7.20 Syrinx by Claude Debussy [10] 68
7.21 Syrinx spectra 69

A.1 ASUS built-in 82
A.2 MacBook built-in 82
A.3 ASUS with Audio-technica 83
A.4 Macbook with Audio-technica 83
A.5 ASUS with H2 zoom 84
A.6 MacBook with H2 zoom 84

B.1 Sheet music for recording 85

D.1 Flute List - teaching experience 91

E.1 G4 repeated note (part 1 of 2) 94
E.2 G4 repeated note (part 2 of 2) 95

CHAPTER 1

Introduction

For the purposes of this paper, “tone quality”, “tone color”, and “timbre” will be used interchangeably. These terms will be defined as: the audible difference in character that a listener perceives for two notes played at the same pitch [13, 29]. For example, an oboe playing the note at a given pitch has a different timbre than a flute playing the same pitch. Moreover, two different flutists playing the same pitch can have different timbres or tone qualities.

Flutists value good tone quality, and some regularly invest time practicing tone exercises. Geoffrey Gilbert suggested spending one-third of practice time on tone studies [18]. Some of Gilbert’s former students include James Galway, William Bennett, and Trevor Wye. There is a substantial corpus of information devoted to flute tone quality. These sources include: books, periodicals, websites, fora, and scholarly papers. Browsing the bibliographies of these, dissertations, or quality websites on the topic of flute timbre reveal a wealth of literature [3, 6, 54].

When flutists talk about tone quality, there is often an implicit understanding that different musical contexts require particular tone colors [18, 28, 49, 51]. For example, what is appropriate for an early Baroque chamber music piece might be different than an Ian Anderson (Jethro Tull) improvised solo. Further, within a given genre or even within a single composition, circumstances may call for a variety of subtleties of timbres. Appropriate tonal variations are applied to enhance musical interpretation during a performance. This might be akin to a painter using a palette with a variety of subtle hues to shade his art.

Aside from musical circumstances, personal taste plays a role in evaluating tone quality. Some people prefer Jean-Pierre Rampal, some James Galway, and others Emmanuel Pahud. Interestingly, there have been studies that show personality types influence timbre preferences [39,50]. Another aspect of preference may be the level of ear-training and the aural acuity of the listener. Cultural background is another factor that influences tone perception [54]. Although musical context, personal taste, as well as other factors influence each listeners assessment of "good" flute tone quality there may be some consensus on standards for tone. For example, a beginning elementary school flutist will probably produce a sound that is less desirable than the principal flutist from an elite symphony orchestra.

Flute students may find it challenging to develop their sound while navigating through the aforementioned subjectivity. This is further exacerbated by perplexing descriptions of tone color like: bright, dark, dull, edgy, hollow, round, fuzzy, pure, reedy, etc. It is also not uncommon to find flute literature indicating more harmonics enriches tone quality. However, the specifics of which harmonics and the appropriate balance is is seldom specified. Even when data like harmonic spectra are used to illustrate differences in tone quality, readers must often rely on only written descriptions of timbre differences without the benefit of aural input [36]. To quote Roger Stevens: "Verbal terms describing tone colors are quite inadequate, and as such descriptions are, for the most part, purely subjective." [49]

This project explores (1) how tone quality is described by skilled flutists, (2) whether the harmonic content or spectral signature has some correlation with tone quality, (3) whether certain acoustic signatures are preferred, or considered "good". The hope is that there are some measurable aspects of timbre that can be associated with desirable qualities.

In order to accomplish these objectives the first step was to obtain a range of flute tone samples. These samples were analyzed by application programs written specifically for this project. Finally, descriptors and ratings from experienced flutists were procured via online surveys and one-on-one interviews.

The remainder of this paper is organized as follows: chapter 2 introduces domain-specific background information; technical background information is in chapter 3; chapter 4 covers related flute tone research that involves harmonics; the implementation of application programs is discussed in chapter 5; chapter 6 outlines the methodology and experimental setup; chapter 7 details the results and analysis; chapter 8 contains conclusions and future work.

Selected audio clips for this thesis are playable from a web browser at:

http://flutetone2.atwebpages.com/

CHAPTER 2

Domain Specific Background

> A researcher in timbre perception must simultaneously be a musicologist, psychologist, physicist and, perhaps, a computer programmer and engineer. [21]

This section introduces aspects of flute tone, like how the instrument produces sound and aspects that influence timbre and listener perception.

2.1 Flute Tone and Harmonics

A flute sound is produced when the flutist blows air across the embouchure hole (see figure 2.1). The “embouchure hole” is the hole in the flute headjoint located close to the flutist’s lips. The column of air from the flutist is sometimes referred to as an “air jet” [55]. When the air jet strikes the edge of the embouchure hole the stream oscillates into and above the hole, acting as an air reed [49, 51].

The air reed oscillations can be explained by the Bernoulli effect from fluid dynamics [4, 22]. When the air jet strikes the edge of the embouchure hole there are eddies or instabilities. When the air jet bends into the embouchure hole the air jet moves faster above the hole resulting in a pressure differential. Since there is less pressure above the embouchure hole, the airstream then bends in the opposite direction and moves above the hole. Under the proper circumstances the air jet rapidly oscillates into and above the hole.

Figure 2.1: Air jet across embouchure hole [54]

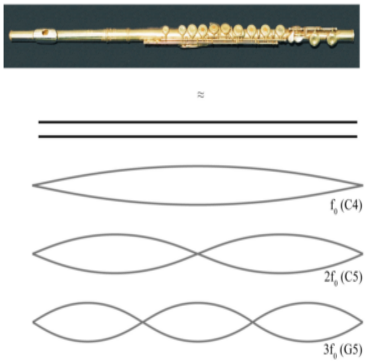


Figure 2.2: Flute standing waves [12]

When the air column inside the flute vibrates in a uniform manner the flute becomes a resonator [29]. In this situation, there is something called a "standing wave" inside of the flute. The wave appears to be stationary, but actually has energy moving from the headjoint towards the far end of the flute. The energy reflects back towards the headjoint, but at equal amplitude along the same wave pattern [12].

A given flute tone oscillates at a periodic frequency [4]. This frequency generally determines the pitch listeners perceive, and this is the fundamental frequency or the first harmonic. One factor controlling the fundamental frequency for a flute tone is how the flutist configures or presses the keys on the instrument. The configuration alters the length of the instrument causing the flute to vibrate at different frequencies [2]. Figure 2.2 depicts the standing waves in a flute with all of the tone holes closed. The longest wave is the fundamental frequency. It is also possible to use the airspeed/pressure to alter the pitch. This is referred to as "overblowing", and can produce higher pitches which are depicted by the shorter waves [55].

There are other naturally occurring frequencies that accompany the fundamental, and together these comprise a set of harmonics. Each harmonic vibrates at a frequency that is an integer multiple of the fundamental. For example, many people know that the note A440 vibrates at 440 cycles per second, or 440 Hz. The fundamental frequency for this note is 440 Hz. There is a harmonic that vibrates at 440 Hz x 2 = 880 Hz, or the second harmonic. Harmonic 3 has frequency 440 Hz x 3 = 1,320 Hz, and so forth.

Sometimes the harmonics above the fundamental are referred to as "overtones" [2]. The fundamental along with the overtones comprise the harmonics, or harmonic spectrum for a tone. The set of harmonics are also sometimes referred to as partials. The combination and balance of these harmonics determine the tone quality or timbre [2, 38, 47]. For clarity, the terms "harmonics" and "fundamental" will be used from here on. The terms "overtones" and "partials" will be avoided. The exception will be situations where quotations from other sources include these terms.

A computer-generated sine wave tone is an example of a tone with a strong fundamental and lacking in other harmonics [4]. This is often described as a "pure" or "simple sound" [34, 49]. Although not as simple as a sine wave tone, flute tone is often described as having a "pure" tone quality compared to other instruments. In contrast, oboe tone contains a rich mix of upper harmonics and has a more "complex" quality [22].

There is some debate on certain factors influencing timbre discrimination. The onset of a musical note is called the "attack", and can be characterized by a short burst of energy. This is followed by a steady-state phase. There is some disagreement on the relative importance of the steady-state versus the initial attack transient for timbre recognition. In Dr. John Hajda's dissertation, he cites several studies on each side of the debate. Hajda quoted a study where the initial attack was removed from recorded sounds:

> A tuning fork was mistaken for a flute, a trumpet for a cornet, an oboe for a clarinet, a cello for a bassoon; but even more contrasting tone colors could not be differentiated, such as cornet and violin, or French horn and flute. [21]

One of the conclusions from Hajda's research was: "The steady state was the most salient segment for the identification of sustained single continuant tones of the Western Orchestra." However, this may not sway proponents that contend the attack plays the greatest role. This research project focuses on the steady-state harmonics of flute tones.

2.2 Flute Construction Material and Tone Quality

Since the flute is the primary resonator, a frequent topic of discussion is how much impact the material of a flute's construction has on tone quality. Sir James Galway has a Youtube video where he plays a short excerpt on 16 different flutes [19]. The flutes were built by six different manufactures from silver, gold, and platinum with varying alloys. Although there are subtle differences between the timbre produced on each instrument, many fans and skilled flutists would recognize that it is Galway playing regardless of which flute was used. Verne Q. Powell, a builder of elite flutes, once said: "As far as tone is concerned, I contend that 90% of it is the man behind the flute." [51]

One scientifically controlled study by John Coltman showed that skilled players and experienced listeners could not distinguish between flutes constructed from silver, copper, and wood [7]. In a subsequent anecdotal test, Coltman played a flute constructed from cherry wood and then another constructed from concrete, and produced tones that were indistinguishable to an audience [23].

A scientific study by Gregor Widholm used identical Muramatsu model flutes constructed with different alloys including silver, gold, and platinum. Widholm found that professional flutists and listeners could not detect differences in tone quality [53]. An article by Neville Fletcher cites experiments where listeners could not distinguish between silver, copper, or cardboard [15].

Despite these studies, there are very strong proponents that material does make a difference in tone quality [35, 51]. Whether tone quality differences are real or imagined, demands for flutes constructed with a wide range of materials and alloys persist. This is evident as many quality flute manufacturers continue building and selling flutes constructed from a variety of materials.

2.3 The Flutist and Tone

Flutists can influence the timbre with their embouchure and air. "Embouchure" refers to the configuration of the flutist's lips. The shape, size, angle, airspeed/pressure affect how the flute resonates [18, 24, 49, 51]. Additionally, the length of the air-reed, or the distance from the lips to the edge of the embouchure hole, plays a role [9, 17]. While it is clear that the flute is a resonator, it turns out that vibrations from the instrument feedback into the players mouth so the flutist becomes an upstream resonator [5, 30]. Robert Dick sums it up nicely:

> Firstly, the tone of the flute is not just the tone made in the instrument, it is a complex combination of the flutist and the flute. The sound we hear is that of the air vibrating within the flute, but resonated within the body of the flutist!. [11]

The physical configuration of the players anatomy contribute to the tone. In particular, the mouth cavity (tongue and soft palate position) and throat configuration influence timbre [18, 24, 28, 54]. There are even further refinements to optimizing a flutist's resonance. Robert Dick teaches a technique for "throat tuning" based on the pitch of a note which allows the vocal chords to resonate appropriately [11]. Robert Aitkins has a slightly different approach, advocating always "setting the body resonance for the lowest notes of the flute" to achieve good tone [3].

Flutists understand there is an intrinsic relationship between the flutist and the flute. Provided the instrument is of reasonable quality, a particular flutist produces similar sounds when playing different flutes. However, it is often very easy to distinguish between two flutists when they play a particular flute.

2.4 Dynamics and Tone

In a musical context, the term "dynamics" can be used in various ways. For the purposes of this study, "dynamics" will refer to the perceived loudness, or volume, of a note. The flute has a small range in terms of dynamics [49] and is not very efficient at producing sound. In one case, John Coltman estimated that only 2.4% of the airstream energy was converted to the standing-wave vibrations inside the flute [22]. One way to achieve louder dynamics is by using more air. Another means of changing dynamics is altering the tone color. Geoffrey Gilbert suggested using fewer harmonics for softer dynamics, and adding more harmonics for louder dynamics [18]. While changing tone color may not actually alter the amplitude, it may alter the listener's perception of the loudness of a note.

2.5 Vibrato

Flute vibrato is produced by varying the pitch, amplitude, and/or timbre [40, 46, 51]. Flutists achieve vibrato by pulsing the air pressure [51]. There is some debate on the optimal method of producing flute vibrato. Some advocate it originates from the diaphragm. However, Gartner points out that it is actually the abdominal muscles that produce this type of vibrato [20]. Others contend that vibrato is primarily controlled in the throat, larynx, or chest. Some suggest it is a combination of the factors, and others believe it varies depending upon circumstances. Understanding the mechanics that drive vibrato is not essential to this project. However, the effects of vibrato on the the harmonic signature is of interest. Some observations about vibrato are in the analysis section (see 7.3.1).

2.6 Acoustics and Psychoacoustics

In our context, sound is a vibration produced by a flute that travels through air. These vibrations can be recorded, measured, and analyzed. Acoustics, or "the science of sound" is a matter of physics [13]. However, the manner in which an individual perceives the sound is related to psychoacoustics.

Musical notes produced by a flute oscillate at a measurable frequency. The pitch that an individual perceives is related to the frequency but not identical [13, 34]. F. Alton Everest states:

> As intensity increases, the pitch of a low-frequency tone goes down, while the pitch of a high-frequency tone goes up. Fletcher found that playing pure tones of 168 and 318 Hz at a modest level produces a very discordant sound. At a high intensity, however, the ear hears the pure tones in the 150- to 300-Hz octave relationship as a pleasant sound. We cannot equate frequency and pitch, but they are analogous. [13]

Similarly, sound intensity versus loudness has an acoustic/psychoacoustic duality [13]. Loudness perception varies depending on the frequency of a sound and it is not completely understood [34]. Figure 2.3 shows the sound-pressure level (vertical scale) required for different frequencies (horizontal scale) to sound equally loud. The left side of the chart shows that at lower frequencies, the sound must have greater intensity to be perceived at the same loudness as frequencies near 5 KHz.

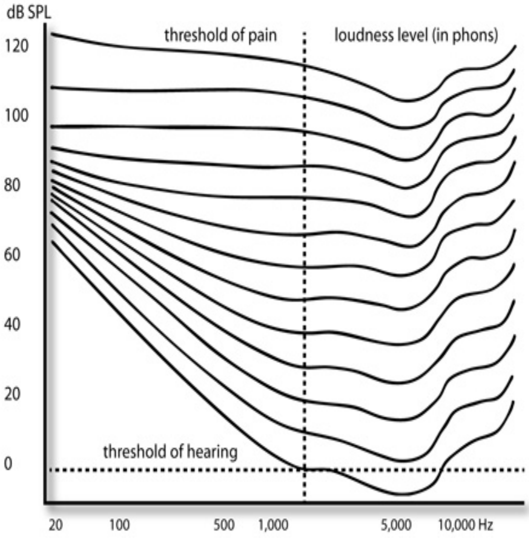


Figure 2.3: Fletcher-Munson equal loudness curves [32]

There is complex relationship between the frequency-pitch of various harmonics along with the intensity/loudness and listener perception of tone quality. Curtis Roads mentions: "... timbre is at least as concerned with perception as it is with sound signals." [43] Some of the factors that inform timbre perception enumerated by Roads include: amplitude, undulations due to vibrato and tremolo, perceived loudness, duration, and spectral content over time.

There is a duality between: frequency and pitch perception, sound intensity and loudness perception, and harmonics and timbre perception. While each of these pairs are closely related, they are not exactly the same. In each case, the first is directly measurable while the second is somewhat subjective depending upon the listener.

2.7 The Flutists Perspective

An informal survey of flute teachers and online flute forum members [27] produced a variety of descriptions for good flute tone: depth of sound, resonance, flexibility and colors, fullness of sound, full rich sound, variety of color, projection, mostly solid, mellow, mixture of brassy sizzle and pleasant, warm, clear, focused, centered, vibrant, bright, dark. It is not always easy to apply these descriptors objectively, and it is clear that not everyone interprets or applies them in the same way.

Most serious flute students and professionals will tell you that "good tone" is important, and many will indicate that it is their highest priority (see section 7.2.1). While there are many sources of information and instruction to help a flutist, defining good tone quality is problematic. Even if a common standard of good tone were established, flutists face other challenges. Musicians have a unique perspective of their own tone because of the close proximity to the sound source (the flute). Further, a flutist's perception of their own tone differs from nearby listeners because their head is resonating the sound. Flutists therefore have a nebulous target for tone quality, and often rely on self-feedback that may differ from an audience.

CHAPTER 3

Technical Background

This section introduces technical information relevant to implementation of the application programs. These programs are an application of Digital Signal Processing (DSP) for spectral and pitch analysis.

3.1 Digital Signal Processing

Sound waves from a flute are analog signals that can be measured by a device like a microphone. A microphone converts the acoustic vibrations to electrical signals [37]. An Analog-to-Digital Converter (ADC) changes the analog signal into a digital representation. The sound card on most contemporary computers has an ADC.

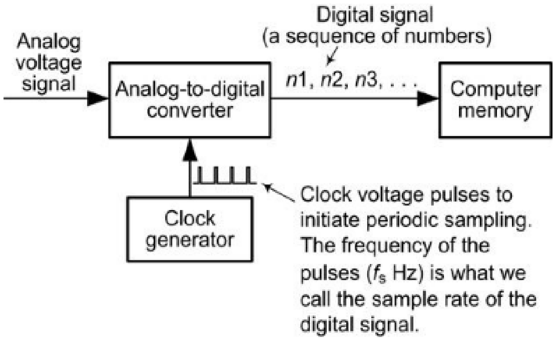


Figure 3.1: Analog to digital signal conversion [32]

Converters use a sampling process to approximate analog signals. The sampling rate affects the accuracy of a digital representation. In particular the sampling rate must be sufficient to capture the highest frequency of interest. 20,000 Hz is generally considered the upper frequency that humans can hear. The Nyquist theorem states that the sampling rate must be twice the highest frequency, or 40,000 Hz. The music industry standard sampling rate for music Compact Disks (CDs) is 44,100 Hz [32].

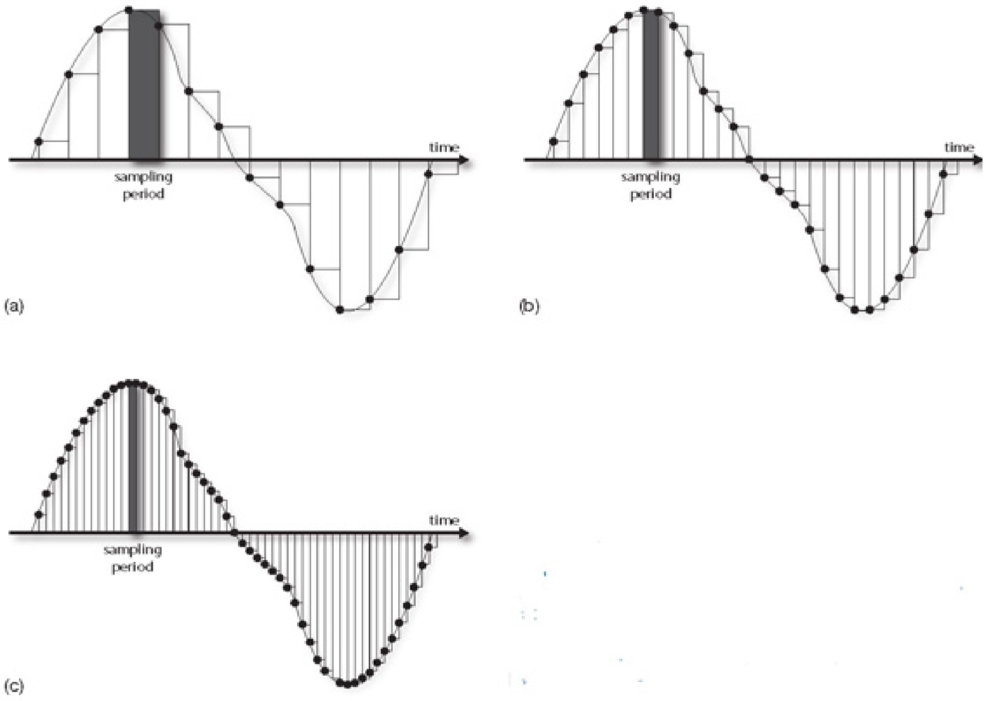


Figure 3.2: Sampling rates [25]

Figure 3.2 shows the effect of using different sampling rates. Increments along the horizontal axis depict the frequency of sampling. As the rate increases the accuracy of the signal representation increases. Example (a) in the diagram has the lowest sampling rate, and hence the least accurate representation. Example (c) has the highest sampling rate and has the closest approximation to the original analog signal. The vertical axis reflects the amplitude of the analog signal. The number of bits used to represent each samples amplitude is the bit rate. CDs use 16 bits so there are 2 to the 16th distinct possible values for each sample's amplitude. Higher bit rates have greater accuracy representing the amplitude. The process of assigning digital values is sometimes referred to as "quantization".

3.2 Spectral Analysis

A digital signal consists of a sequence of discrete values that have been sampled at fixed time intervals. This representation is in the time-domain [32, 43]. The digital signal can be transformed to the frequency domain to show the spectrum, or the frequencies, within the signal. This is also known as spectral analysis or spectrum analysis. Curtis Roads aptly states:

> Except for isolated test cases, the practice of spectrum analysis is not an exact science. The results are typically an approximation of the actual spectrum, so spectrum analysis is perhaps more precisely called spectrum estimation. [43]

One approach for spectral analysis is based upon Fourier analysis. In the 1800s Jean Baptiste Joseph, Baron de Fourier concluded that vibrations can be analyzed as a sum of simple sine waves. This theory eventually was implemented into an algorithm known as the Fourier Transform (FT). It is possible to see the harmonics that comprise a musical tone by applying a FT which converts a musical signal from the time-domain to the frequency-domain.

3.2.1 Fourier Transform Family

Fourier Transforms (FT) consist of a family of techniques for analyzing signals [43]. The mathematical foundations behind FT can be rigorous and are bypassed here in favor of a high level conceptual overview. The FT was designed for

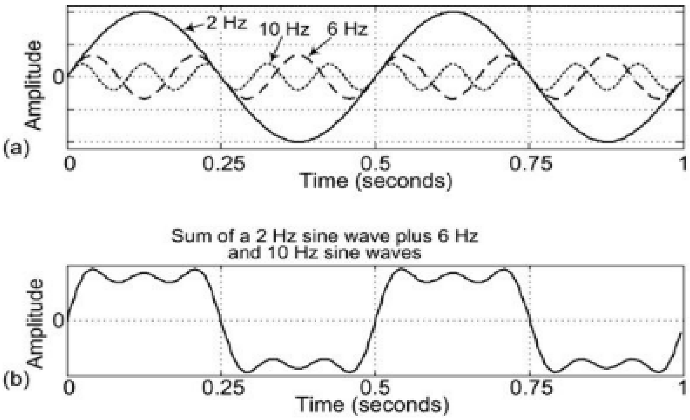


Figure 3.3: Sine waves comprising a sound wave [32]

analog signals. Short-time Fourier Transforms (STFT), Discrete Fourier Transforms (DFT) and Fast Fourier Transforms (FFT) are implementations of FT for digitized signals.

STFT process digitized signals in small equal sized increments, or “windows”. These windows are also referred to as “analysis frames” or “snapshots” [31]. DFT is an implementation of FT for digitized signals and can be applied to each frame producing the spectral components for that time increment. If the snapshots are short, the spectrum for each increment can be analyzed and displayed in near real-time.

The DFT implementation is computationally intensive with a complexity of N^2, or $O(N^2)$. Even with modern computers this can be prohibitive. A more efficient and commonly used algorithm is the Fast Fourier Transform (FFT) which has complexity $O(N \log_2 N)$ [31]. FFT produces exactly the same result as a DFT. As an illustration of the performance difference, given the same data an FFT might take less than 2 seconds while DFT would require more than 2 hours [31].

3.2.2 Windowing Functions

An underlying assumption in the FFT algorithm is that the signal is periodic and continues indefinitely [43]. Conceptually, to apply this algorithm to a snapshot, a analysis window can be replicated and appended together. However, if the signal in a given analysis window is not at exactly the period as the window, there will be discontinuities at the frame boundaries. To minimize these discontinuities a windowing function can be applied which tapers data at the edges of a frame in a bell-shaped manner (see figure 3.4). Examples of windowing functions are Hamming, Gaussian, and Blackman.

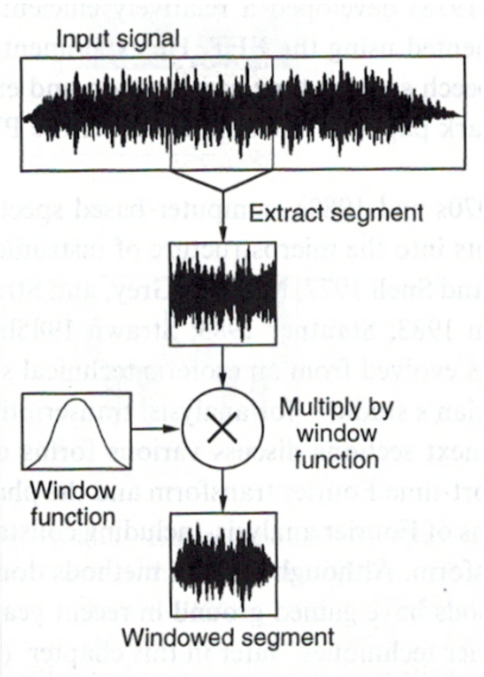


Figure 3.4: Windowing function [43]

While windowing functions reduce discontinuities, there is also some data loss at the edges of each frame. If the analysis frame is small, this may be significant. Another shortcoming of small frames is reduced resolution in the spectrum data (which is discussed further in section 3.2.4).

3.2.3 Overlapping Frames

Using overlapping frames to process input data mitigates some of the data loss caused by windowing functions. In figure 3.5 the bottom portion represents the digitized input signal. Each box, or segment, is a portion of the input. The segment length is known as the “hopsize”. The analysis frame in this example is 2 X hopsize. There is a 50% overlap between each analysis frame. Because the data is interleaved, portions of the data that were lost in a non-overlapping implementation are now present.

In this example, with 50% overlap, the spectral plots (the top portion of figure 3.5) can be produced twice as often versus using non-overlapping frames.

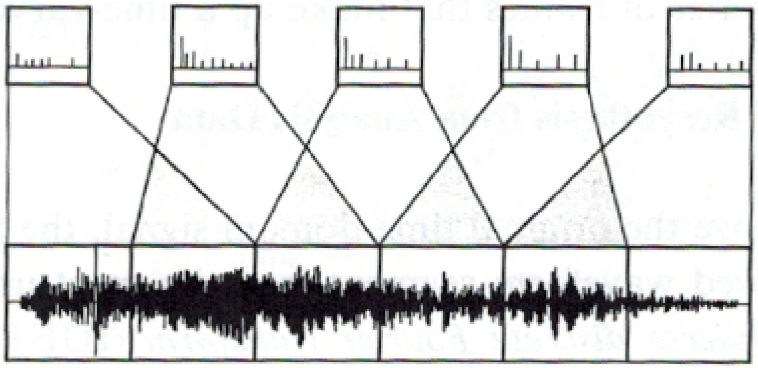


Figure 3.5: Overlapping window frames [43]

3.2.4 Frame Size and FFT Bins

FFTs divide the audible frequency space into frequency bins [43]. Each bin covers a range of frequencies. The number of bins is one-half the analysis frame size. For example, if the frame size is 2048, there will be 2048/2 = 1024 bins to represent the frequencies within a sound. With a sampling rate of 44,100 Hz, the highest frequency would be 20,050 Hz based on Nyquist. Each bin will cover a range of

20,050 Hz / 1024 bins = 19.5 Hz/bin.

There are trade-offs in processing rates vs. frequency resolution. With smaller frames, results can be displayed more frequently. A 2048 frame size using a non-overlapping frame implementation would produce results approximately 22 times per second with 1024 bins to represent frequencies. Using a frame size of 8192 would reduce the display rate, but increase the number of bins to 4096. An implementation of non-overlapping frames of size 8192 would have a noticeably sluggish screen display rate.

3.3 Pitch Detection

This research project is not primarily concerned with DSP pitch detection algorithms. However, some notion of the fundamental frequency is required to accurately determine which FFT bins contain relevant harmonic data. There are several pitch detection algorithms that can be used in either the time-domain, or the frequency-domain. For example, two time-domain approaches are zero-crossing and autocorrelation [43]. There are also frequency-domain algorithms like spectrum peak methods, phase vocoder, and harmonic product spectrum [33]. The time-domain autocorrelation approach was arbitrarily selected as the pitch detection algorithm.

Like STFT, time-domain autocorrelation can use data from analysis frames to produce results in near real-time. Autocorrelation operates under the assumption that the frequency is relatively stable within an analysis frame. Essentially, the algorithm compares the data in a window to itself. The comparisons are done by shifting the data by successive intervals. If the shifted data has a high correlation with the original data, it is indicative that the period of a signal has been identified. The period can be used to calculate the frequency [43].

Autocorrelation can be computationally expensive if it is implemented to detect a wide range of frequencies, e.g., 20 Hz to 20 KHz. For the purposes of this study, this cost was minimized by restricting the frequency range (see section 5.3.2).

CHAPTER 4

Related Work

4.1 Flute Tone Analysis Using Spectra

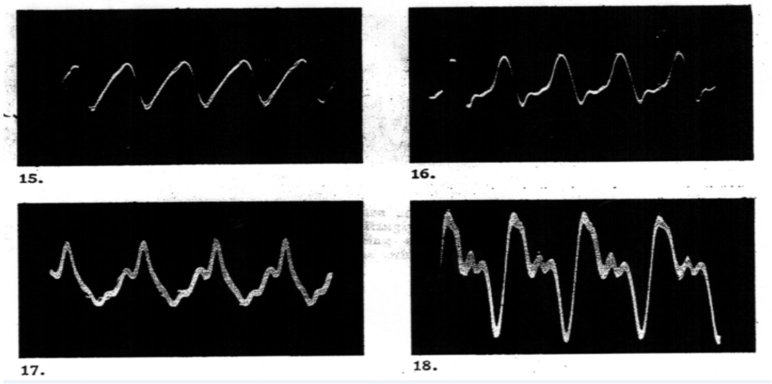


Figure 4.1: Oscilloscope graphs [49]

As early as 1967 spectral analysis was used by Roger Stevens [49] to analyze flute tone quality. He published a book showing graphs from an oscilloscope (figure 4.1). There are series of graphs as a flutist experimented with various embouchure, tongue, jaw and throat configurations. Stevens describes the tone qualities and shows the corresponding time-domain amplitude waves. These waveforms contain a composite of all the frequencies so the relative strengths of each harmonic is difficult to discern. Stevens carefully describes how to interpret each chart to quantify the strength of various harmonics. Although this may be the earliest published analysis showing harmonics, it is very thorough.

Dr. John Coltman was a physicist and avid flute hobbyist and published seminal works using the scientific method on flute timbre. One of his earliest studies from 1971 about flute construction material and tone quality was mentioned in section 2.2. In 1996 his paper on Just Noticeable Difference (JND) for timbre shows harmonic amplitude [8]. In his JND research Coltman conducted an experiment using synthesized sounds based on flute spectra. A single harmonic was incrementally modified to determine the JND in timbre perceived by trained flutists. Initially the test subjects were asked "whether the second tone sounded brighter (increased harmonic content) or duller (decreased harmonic content)". It soon became obvious that brighter and duller were interpreted differently, and in some cases meant the complete opposite for some. The methodology was modified to ask whether listeners perceived an increase or decrease in harmonic content. Another interesting observation was that lowering a pitch without altering the strengths of the harmonics is often heard as a decrease in harmonic content. However, some individuals perceive the opposite and hear an increase in harmonic content.

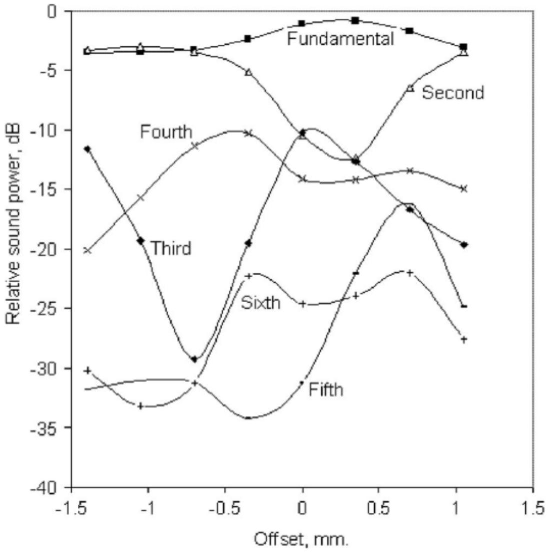


Figure 4.2: Lip to embouchure hole gap effect on harmonics [9]

A more recent paper by Coltman published in 2006, investigates how characteristics of the flutists air stream determine harmonic content of the tone [9]. Coltman investigated the "jet offset", which is the angle of the air stream into the flute embouchure hole. Additionally, he studied the effect of the gap size between the flutists lips and the embouchure hole. Figure 4.2 shows one of the graphics from his research. Harmonics spectra for experiments varying these factors was presented. The study also shows that blowing pressure alters the harmonic mix.

Dr. Neville Fletcher is a physicist and professor at The University of New South Wales in the Research School of Physical Sciences and Engineering with an extensive list of publications. He is also a musician who plays the flute, bassoon, and organ. A paper he published in 1975 shows that in the lowest flute range the fundamental can be weaker than some of the upper harmonics [14]. Figure 4.3 is from his paper showing spectral charts for four different flutists playing various pitches at different loudnesses. He notes "the harmonic development is quite considerable for low notes, though the higher harmonics are much weaker than for the reed woodwinds." Fletcher discusses flute performance techniques like blowing pressure and lip opening size and shape in his book, *The Physics of Music* [17]. In a more recent paper [16], Fletcher analyzes vibrato and discusses amplitude (loudness), pitch, and timbre vibrato. He posits that the relative mix of overtones can vary at different phases of the vibrato.

4.2 Spectral Analysis For Musicians

From a flutists and musicians perspective, Dr. Robert Billington used Cool Edit 2000, a PC software program, to look at harmonics [3]. In his dissertation, Billington describes Robert Aitken's approach of having flutists configure their

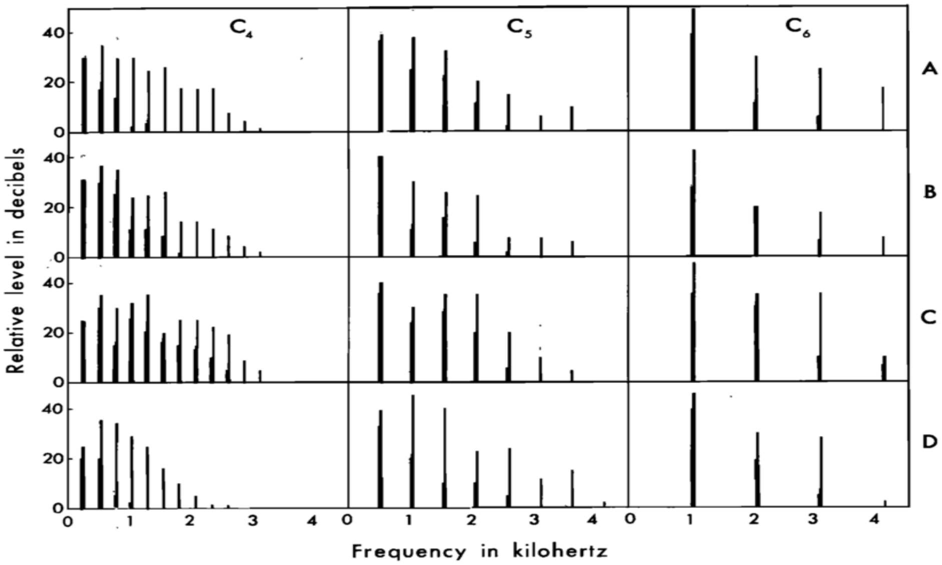


Figure 4.3: Harmonics at different pitches and loudnesses [14]

anatomy for maximum resonance for the lowest notes. Aitken believes that this configuration optimizes flute tone. Billington's study focuses on notes without vibrato. This isolates the "true quality of tone" and avoids weaknesses that can be masked by vibrato. There are spectral charts showing the effects of various changes in the flutists physical configuration. He contends that "Brightness will be associated with a sound whose second partial is of near equal or greater volume than the fundamental." He also indicates that "fullness will be associated with sound that exhibit a minimum of six partials and whose upper partials remain relatively loud with regard to the fundamental."

In 2002 Dr. Katharin Rundus investigated applying spectral analysis as a teaching tool for vocal instructors [44]. Rundus was looking for real-time visual quantitative measures for voice quality beyond the traditional subjective feedback. She used a program from Tiger DRS. The software was originally developed for Speech Pathologists and clinical use, however, it is suitable for spectral analysis of the singing voice. The use of this tool was not only helpful in identifying vocal problems, but was also valuable for ear training. The dissertation examines various factors influencing vocal quality and how to apply the information provided by the spectral analysis including: onsets and releases, breathing for singing, resonance, focus of the tone, articulation, and musical expression.

Daniel Jones conducted research with middle school trumpet players [26]. By using a real-time spectrogram for visual feedback he observed an improvement in 8th grade students tone quality. The oral cavity configuration is mentioned as having "perhaps the greatest impact on resonance" and thus was a focus item during weekly lessons. His metric for quantifying improvement was an increase in harmonics as a measure of improved resonance. Jones observed less improvement in the 7th graders and virtually no change for the 6th graders. He speculates that musicians with more experience can derive greater benefit from this approach. The paper cites several other studies using spectrograms to facilitate tone quality improvement.

CHAPTER 5

HAT Implementation

5.1 Overview

There are many Real-Time Analyzer (RTA) or spectrum analyzer apps available for both Android and iOS devices. At the onset of this project, none of these tools met all of the functionality requirements. For example, most of the RTA are truly real-time so once the digital signal ends the display goes blank. Comparing two different notes using these apps would be problematic. Of equal significance was having the ability to enhance or modify functionality as the project progressed. For example, requirements to visualize information in new ways, or to experiment with different data analysis implementations might arise. Building a set of tools and owning the code seemed a reasonable approach. HAT, or Harmonic Analysis Tools, consists of several application programs to analyze and visualize tone quality. For the purposes of this paper, only the two most relevant applications will be discussed.

The harmonics will be referred to using the notation H_n, where n is 1-7. For example, H1 denotes harmonic 1, and H7 is harmonic 7.

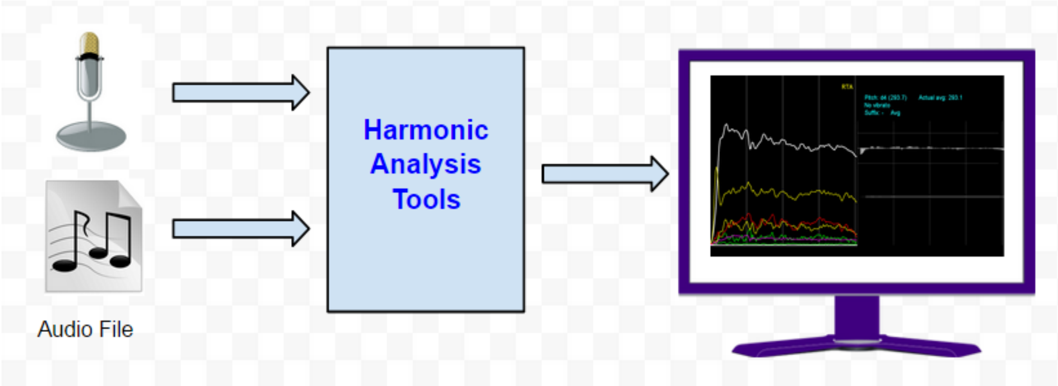
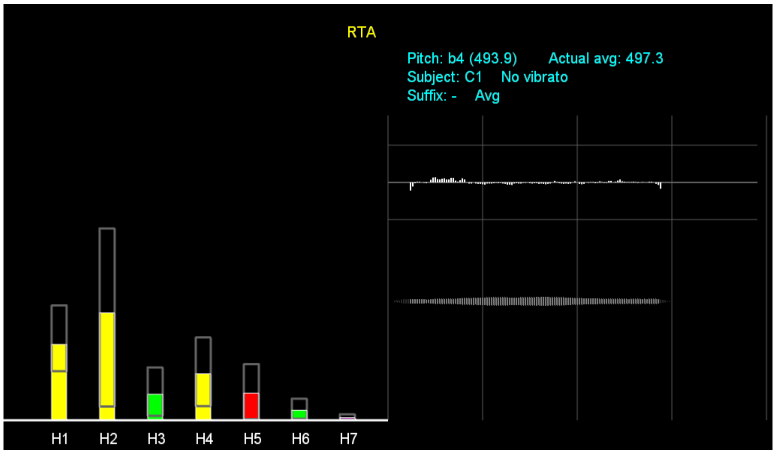


Figure 5.1: Audio file or microphone to spectral signature

5.2 Functionality and Displays

All of the analysis for this study was done using the recorded flute samples. However, there is some limited microphone support. Microphone input is processed as the note is played by the musician.

The HAT applications can be used to analyze spectra for any instrument, but were tailored to the flute. Flute tone is characterized as relatively pure, with few and weak upper harmonics [45]. Several of the studies in the related works section indeed confirm this fact. Further, tests using RTA applications indicate that the harmonics above H7 are generally very weak, or nearly absent. Most of the HAT displays therefore show only the H1-H7 This balances the need to visualize flute tone in a relatively accurate manner, yet keeps the amount of display information reasonable. In the long term, if HAT becomes a practice room tool for flutists, minimizing cognitive overload is an important design requirement.



Pitch: b4 (493.9)   Actual avg: 497.3
Subject: C1   No vibrato
Suffix: -   Avg

Figure 5.2: HAT RTA display

5.2.1 Real-Time Analysis

HAT supports basic RTA style data display. In figure 5.2 the left side of the display shows bars with the vertical height, indicating the strength of H1 to H7. These harmonics are displayed in real-time as the audio clip is playing. Depending upon the flutist, and also whether vibrato was used, the bars height can bounce up and down considerably during playback.

Figure 5.3 shows the colors used to represent each harmonic. The convention is that the fundamental or octaves of the fundamental are yellow. Any harmonic that is a fifth higher than the fundamental or octave is green. Thirds are red, and minor sevenths are purple.

Unlike most RTAs, when the note completes, the screen does not go blank. Upon termination of the note, averages for each harmonic are displayed. To eliminate the instability of the beginning (attack) and end (decay), the average consists of the entire note less the first and last 10 windows (analysis frames). This will arbitrarily be referred to here as the steady state portion of the note for the HAT applications. Each harmonic bar also has a hollow grey box. The

Harmonic #	Description	Graph color
H1	Fundamental (perceived pitch)	Yellow
H2	Octave above fundamental	Yellow
H3	Octave + fifth above fundamental	Green
H4	Two octaves above fundamental	Yellow
H5	Two octaves + third above fundamental	Red
H6	Two octaves + fifth	Green
H7	Two octaves + minor seventh above fundamental	Purple

Figure 5.3: HAT Harmonics Color Coding

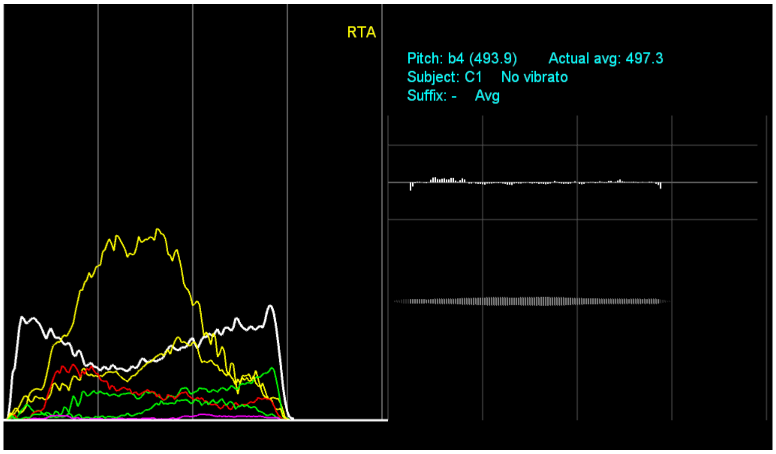
top of the box is the maximum level of that harmonic during the steady state. Conversely, the bottom of the hollow box shows the minimum.

The text in the upper right provides information about the note. For example, the target pitch frequency and actual measured frequency are listed. Also, a unique subject identifier for the flutist is displayed.

The bottom right-side display shows the amplitude (loudness) of the note over time. The long vertical lines demarcate one second intervals. Each short vertical lines indicate the loudness measured for a window (analysis frame).

The middle portion of the right-side display shows the pitch variation. The horizontal line represents the average measured steady state pitch. Using the average pitch as a reference was useful when examining pitch variations resulting from vibrato. The length of the short vertical bars originating from the horizontal average pitch line indicate how much the pitch deviated from the average for each window. Descending lines show some degree of flatness and ascending lines show sharpness. The horizontal lines above the average pitch indicate a quarter-tone above the average. Similarly, the horizontal below the average line indicates a quarter-tone below the average.

Averaging the steady state for each harmonic is a convenient method of summarizing the spectral information. However, if there is much fluctuation over time, useful information may be lost. The height of the hollow grey box for H2 in figure 5.2 is very large indicating considerable variation over time.

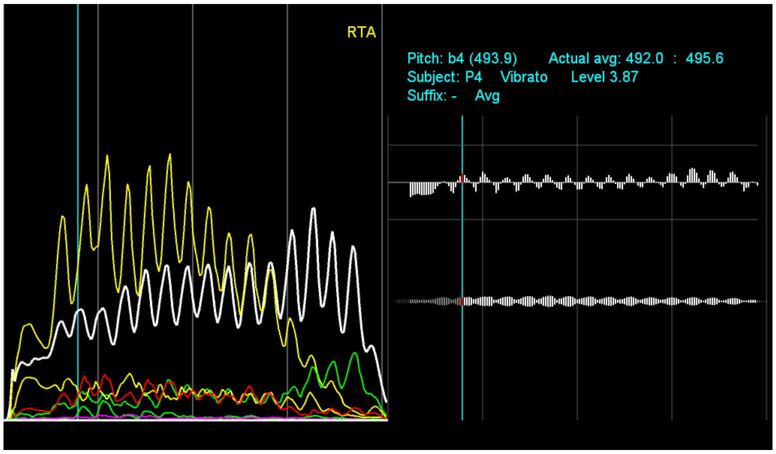


Pitch: b4 (493.9) Actual avg: 497.3
Subject: C1 No vibrato
Suffix: - Avg

Figure 5.4: HAT RTA Line display

A second view of the RTA uses a line chart to show the harmonics which preserves temporal information. Line chart views were used in this study as the “spectral signature”. Figure 5.4 shows an RTA line chart from the same audio file used in the previous figure. When displaying spectral data in this manner, long vertical lines appear on the left-side display indicating one second intervals. The fundamental is represented with a white line. All of the other harmonics follow the same convention as the bar chart. This view of the data clearly shows that the tone quality goes through different phases. Approximately one second into the note, the second harmonic (yellow) spikes as the fundamental (white) dips.

There are situations where correlating the behavior of the harmonics, pitch, and amplitude are of interest. Figure 5.5 shows an example of a note played with vibrato. Using the left and right keyboard keys, HAT will show a blue vertical



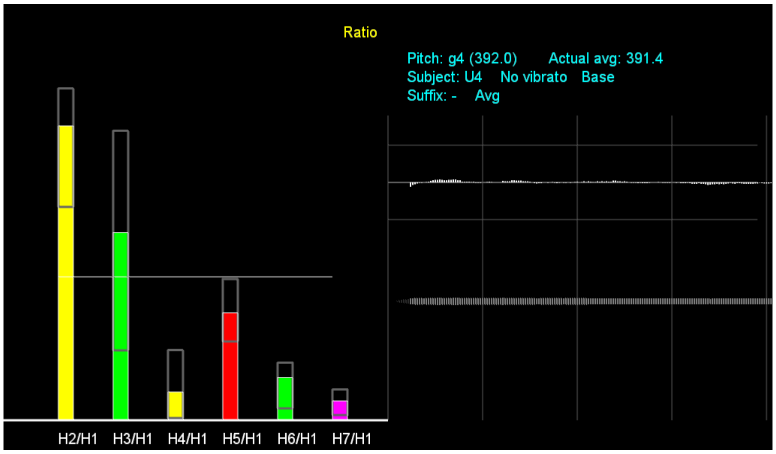
Pitch: b4 (493.9) Actual avg: 492.0 : 495.6
Subject: P4 Vibrato Level 3.87
Suffix: - Avg

Figure 5.5: HAT RTA Line Chart - correlating harmonics, pitch, and loudness

line. Depressing the arrow key will move the blue line one window (analysis frame) in the corresponding direction. The figure shows a point in the note where the fundamental (white) is peaking, the pitch is also peaking and the amplitude is at a plateau. The other harmonics do not necessarily peak at the same point as the fundamental.

5.2.2 Ratio

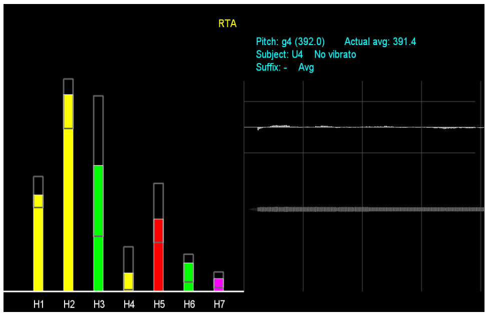
Since the relative strength of harmonics plays a role in tone quality, the HAT Ratio display uses a slightly different view of the spectra than traditional RTAs. The HAT Ratio application normalizes spectral information by calculating the ratio of each harmonic relative to the fundamental. Each harmonic is divided by the fundamental. Since the fundamental divided by itself is always equal to one, it is not shown as a vertical bar. In figure 5.6 the white horizontal line represents a ratio of 1. Harmonics with bars below the horizontal line are weaker than the fundamental. Conversely bars that extend above the horizontal line are stronger than the fundamental.



Pitch: g4 (392.0) Actual avg: 391.4
Subject: U4 No vibrato Base
Suffix: - Avg

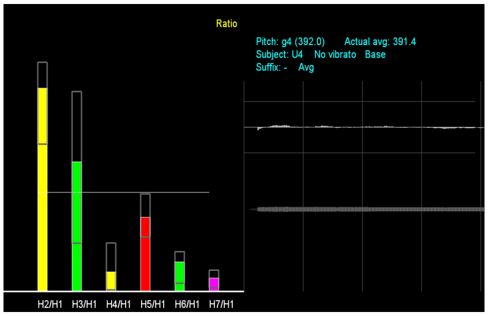
Figure 5.6: HAT Ratio display

Figure 5.7 and figure 5.8 show results for flutist U4 using the HAT RTA and the Ratio displays. In this example, the Ratio display clearly shows that the ratios for harmonic 2 (H2/H1) and harmonic 3 (H3/H1) extend above the horizontal white line and are therefore larger than the fundamental.



Pitch: g4 (392.0) Actual avg: 391.4
Subject: U4 No vibrato
Suffix: - Avg

Figure 5.7: RTA



Pitch: g4 (392.0) Actual avg: 391.4
Subject: U4 No vibrato Base
Suffix: - Avg

Figure 5.8: Ratio

Figure 5.9 shows the Ratio display for a different flutist, C8. In this case all of the ratios are very small relative to the fundamental. Flute literature advocating tones with greater presence of harmonics are “better” suggest that flutist U4s timbre is preferable over C8. Figure 5.10 shows another Ratio display that compares two different flutists. This view shows a substantial difference between the spectras for U4 and C8.

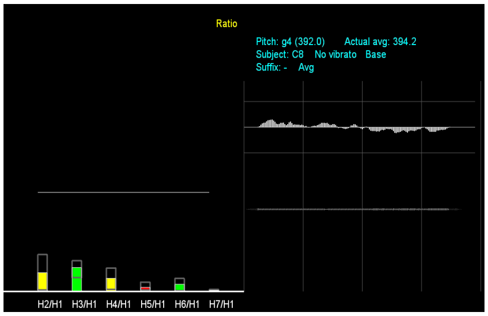


Figure 5.9: Ratio for flutist C8

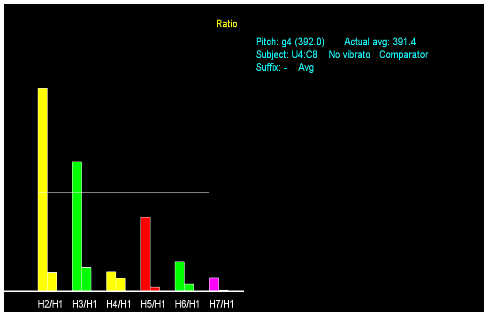


Figure 5.10: Ratio Compare

Although the ratio view was not used for analysis in this paper, it is presented here because there is potential utility beyond this study. For instance, a student may want to develop a timbre resembling their teacher. Using audio files of their teacher as a reference, the student can experiment and attempt to match the instructor's spectral signature.

5.3 HAT Implementation

HAT was implemented using the Processing programming language which originated to help make interactive graphics programming easier [41]. Processing is a dialect of Java, and is open-source with a community of developers that contribute libraries [42]. The Minim sound library is included with Processing and was written by Damien DiFede [52].

5.3.1 HAT Components

The processing runtime has a built-in loop, the draw loop, that renders information to an output window at a user defined rate. This loop runs automatically and continuously. For HAT, the output window is the RTA or Ratio screen, and the rate is set at 22 times per second. This rate is rapid enough so screen images appear reasonably in-sync with audio playback.

The HAT user selects a particular flutist to analyze. The user can also specify which style of a note to play. For example, the user can choose samples with: straight-tone (the default), vibrato, or dynamics. The user then picks a particular pitch for playback and analysis. Ultimately these user specified criteria determine a sound file for HAT to process. Internally, HAT uses a Minim library API call to load and play the user specified sound file. A HAT listener function receives callbacks as the sound file is played back. Callbacks occur whenever a hopsize increment of the sound file has been processed. HAT uses a 50% overlapping window, so the hopsize is one-half the window size. The default window size for HAT is 2K with a 1K hopsize. The HAT user can choose window sizes of 2K, 4K, and 8K.

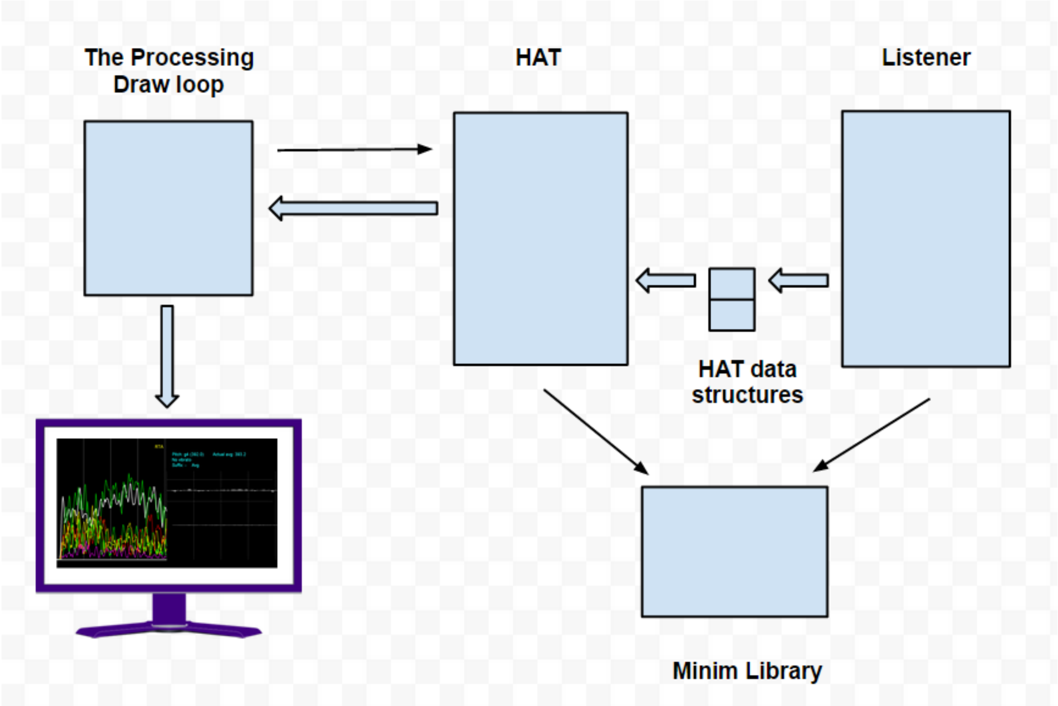


Figure 5.11: HAT components

Whenever HAT begins playing a sound file, it resets all the data structures that contain analysis data. Subsequently, whenever the listener function receives a callback, it first uses autocorrelation to find the fundamental pitch for the current analysis window. Then, an FFT is performed on the window by using a Minim API. The Minim library call automatically applies a Hamming window function prior to performing the FFT. The listener then uses the fundamental pitch frequency to locate the appropriate FFT bins for the seven harmonics. The window's frequency and harmonic data are inserted into the data structures.

Since the callbacks occur asynchronously during sound file playback, analysis data continues to accumulate in the structures. Meanwhile, each iteration of the processing draw loop retrieves all of this data, and renders it to the display.

5.3.2 Implementation Alternatives

There are three areas of the HAT implementation where alternative approaches were tested. The first area was non-overlapping versus overlapping windows. The second area surrounded different approaches calculating the harmonics from FFT bins. The last item was reducing the computations for autocorrelation.

Both non-overlapping and overlapping windows (50% overlap) were implemented. However, for a given audio file, both implementations produced spectral signatures that were nearly identical. Also, the overall characteristics between distinct samples were clearly visible with either implementation. Since either approach was sufficient for the project, overlapping windows were used. Although the level of overlapping can be adjusted, HAT simply uses 50% overlap. No testing was conducted with varying levels of overlap.

A given frequency maps to a bin, which will be referred to as the "target bin".

Some effort was applied investigating alternatives for FFT bin selection. Since each FFT bin covers a range of frequencies, there is the possibility that a note may be at a frequency close to the boundary of its target bin. Another consideration that is a note's frequency may fluctuate within a window (for example with pitch vibrato). In either case, using a single target bin for a frequency can provide misleading results if the pitch moves outside of that bin for an analysis window.

Analysis and testing showed that using the target bin along with the neighboring bins produced good results. Generally, the target bin has the greatest amplitude and the neighbor bins have the second and third greatest. Bins farther away than the two neighbor bins drop off precipitously. Implementations summing the three bins versus averaging the bins were tested. The Minim library also has an API, calcAvg, that allows callers to specify a frequency range and Minim returns the average. Other than the scale of the harmonics, the overall geometry and proportions were similar for summing, averaging, and Minim calcAvg. HAT allows the user to choose among these implementations. All of the spectra shown here used the Minim calcAvg.

The last item concerns reducing the computational complexity of autocorrelation. Searching the entire audible range from 20 Hz to 20K Hz is unnecessary. Since most flutists have a range from B3 to B6, the frequency range can be narrowed to approximately 247 Hz to 1,976 Hz. However, HAT requires users to specify the target pitch. This allows HAT to restrict the autocorrelation search to 2 half-steps of the target pitch. This greatly reduces the computations required for autocorrelation.

CHAPTER 6

Methodology

This section outlines the how the flute tone samples were acquired and analyzed.

6.1 Gathering Flute Tones

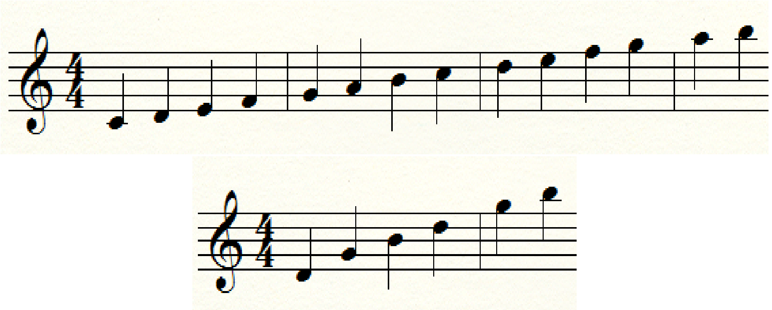



Figure 6.1: Notes for recording

In order to measure and evaluate flute tone quality, the first step was to establish a repository of flute long-tones. Thirty-one flutists of various skill levels were recorded. The flutists included: four high school students, 16 university students (both music majors and non-majors), and 11 professionals. Details for the recording process and instructions for the participants are in Appendix B.

Recruiting professional flutists proved challenging. Approximately 30 professionals were contacted via email. These individuals were generally instructors at universities and/or active members of orchestras or ensembles. In many cases, there was no response to the message. In a few situations they responded, but decided not to participate. Fortunately, a local flutist took interest in the project and helped secure professionals in the Central Coast Region. The four out-of-state professionals had prior interactions with the researcher and graciously agreed to participate. Although a larger set of skilled flutists could have provided useful data, time constraints precluded further recruiting efforts.

Scheduling recording sessions required some flexibility. Four of the professionals resided in either Colorado or New Mexico. Seven professionals were located within a 70 miles radius of San Luis Obispo. Student from three institutions participated: the University of California at Santa Barbara, California Polytechnic State University, and San Luis Obispo High School. It was not reasonable to expect all these flutists to commute to one location and use a controlled recording studio. Circumstances necessitated conducting recording sessions at locations convenient for the subjects. These venues included offices, class rooms, practice rooms, churches, and private homes. Individuals were recorded from January 2014 to May 2014.

The variability of recording environments between flutists is not a fatal issue for this study. While some level of audio fidelity is necessary, rigid control and exacting duplication of the environment is not required. The flutists, flute, and venue are all taken together to produce a digital recording. Any given digital recording represents a particular timbre that is a product of the musician and circumstances. The focus of this project is to take a digital recording, analyze its spectral characteristics, and use that recording for a listener to describe and rate the quality. The essential relationship is that the recording used to produce the spectral signature is what a listener uses to describe the quality.

For example, one specific recording venue may enhance the second harmonic but attenuate the upper harmonics (harmonics 3 to 7) for a particular note.

The spectral analysis might show a strong second harmonic and weak upper harmonics. The listener will hear and evaluate the recording with these identical characteristics. This project attempts to understand what descriptors accompany a particular recording. The listener does not need to know if these characteristics are caused by the room, flutist, or both. In this example, the commonly occurring descriptors might be: mellow, pure, thin, airy. Given these descriptors and the harmonic profile, the objective is to determine whether these are qualities that are preferred.

That being said, where possible, the factors that could be controlled during recording sessions were. Microphone placement was always approximately 5 feet directly in front of the musician. The identical microphone and recording equipment were used for all the sessions. Appendix A contains details about the recording equipment. The recording level was identical for all sessions. A rough rule of thumb for factors influencing recordings is: 50% musician, 20% room, 20% microphone position, 10% microphone choice [37].

The recording sessions produced more than 1,600 samples. Each sample was analyzed by the Harmonic Analysis Tool application. HAT has a screenshot save capability. For each audio sample there is a corresponding image of the spectral analysis. To facilitate viewing sets of spectra, rudimentary javascript programs were implemented. Using these javascript programs in conjunction with a simple web browser form, it was possible to group samples by pitch or flutist. This enabled visual spectra comparisons: between flutists, among different pitches for a flutist, between straight-tone and vibrato, etc. The associated sound file for each spectra could be played for aural comparisons as well.

6.2 The Survey

The survey's primary purpose was to establish how skilled flutists describe and judge tone quality. The following sections explain how the survey was designed. The process of administering the survey to the target demographic is also covered.

6.2.1 Selecting Flute Tones for the Survey

Only straight-tone samples were used for the survey. There were two factors behind this decision. The first is that vibrato adds significant complexity to the harmonic characteristics of a long-tone. Some observations about vibrato and harmonics are discussed in section 7.3.1. The second factor is that tone samples with vibrato might draw evaluators to focus on the vibrato quality rather than the underlying tone quality. One survey comment said it well, "... non-vibrato dis-serves the better players". The corollary might be, "vibrato can betray lack of mastery and refinement in less accomplished flutists". Using straight-tone samples circumvented these issues.

A set of notes was selected based on their acoustic signatures. An important criteria was to use samples with with relatively stable harmonics over the duration of the note. It is surprising how much the harmonic mix can fluctuate for a straight-tone note. The other selection criteria was based on visual and aural distinctiveness within a given pair of notes. Along these lines, tone-pairs had:

- visually different harmonic signatures
- timbre differences that most skilled listeners could easily hear
- the same pitch or note name

Concurrent with the recording sessions, seven professional flutists and one university band director were individually interviewed. They listened to pairs of samples and described the tone quality and rated them. The interviewees were not given any criteria for “good” or “bad” tone quality. As the interviews progressed some samples were eliminated. For instance, one pair of samples with visually different signatures was often judged as sounding the same. Since one of the objectives was to capture descriptors for tone quality, feedback like “they sound the same” or “I can’t hear any difference” was not useful. As the interviews progressed, newly acquired samples with appropriate characteristics replaced samples that were deemed less effective.

During the interview process, when listeners were given a single note to evaluate, they were often unable to make any judgements. Generally they needed at least two notes so they could establish some frame of reference. Alternatively, when listeners were given a set of three different notes to compare, they often needed to listen to the audio clips several times. With sets of three notes, listeners frequently asked to hear various pairings within the set. Using three note comparisons markedly increased the complexity of the process. Based on the interview experience, the decision was made to use pairs of samples for survey.

An important observation was that listener fatigue degraded feedback quality. To minimize the impact of listener fatigue, the number of comparisons needed to be limited. Restricting the comparisons to 6 pairs allowed evaluators to complete all the process in approximately 10-15 minutes. Keeping the survey short ensured a high survey completion with attentive responses throughout. The obvious trade-off is data could only be gathered for a small set of tone samples.

6.2.2 Designing the Survey

The primary target for survey input from skilled flutists was the Flute List. This is the "longest-established internet mailing list relating to the flute" [27]. It is an online forum that continues to be very active. There are informative exchanges between enthusiastic hobbyists, students, and professionals. It is not unusual to see contributions from prominent and influential flutists. Administering a survey to the Flute List required a mechanism that members could easily access. The survey also required that participants could play audio clips, and subsequently rate and describe the tone quality. Many free online survey tools do not support embedding audio files into their surveys. SoGoSurvey provides audio support by upgrading from the complementary basic version [48]. Since SoGoSurveys upgrade cost was reasonable, it was selected for implementing the survey.

Each pair of tones consisted of notes played by different flutists at a similar pitch. The pitch-pairs were ordered in ascending pitch in the hope that this might maintain a higher level of participant interest during the survey process. The 6 pitch-pairs were: D4, D4, G4, G4, B4, D5. The survey questions are in Appendix C.

The survey was intentionally designed to be somewhat ambiguous. Specifically, participants were asked to listen to, and then describe/rate tone quality for notes without any guidance. The musical context, as well as any other criteria for judging timbre were omitted from the instructions. One of the objectives was to observe whether there is some level of consensus about flute tone quality independent of context. No examples of adjectives or phrases for tone quality were provided. This avoided biasing evaluators, and required them to use their own descriptors. The instructions encouraged participants to freely express their opinions as this is subjective and therefore there are no correct or incorrect answers.

Prior to launching the Flute List survey, a pilot study was run with Cal Poly student musicians. An interesting trend emerged from the pilot study. There were some noticeable differences in tone preferences based on the professional interviews versus the student pilot. Since the number of interviewed subjects as well as the student pilot participants was small, it was premature to draw any conclusions. As a result of this observation, additional student surveys were conducted.

CHAPTER 7

Results and Analysis

This chapter is organized into three sections. The first part covers some general observations from the recording sessions and the subsequent spectral analysis. The second section is an analysis of the survey results. The next two sections contains analysis of spectra for long-tones with vibrato and with dynamics. The last section looks at long-tones from some prominent flutists.

7.1 General Observations

7.1.1 Recording Sessions

Playing a series of long-tones without musical context, in a singing or sweet manner, can be challenging. The task can be more difficult for some musicians if they become self-conscious playing in front of a researcher and a recording device. In other cases, foreknowledge that the recordings would be analyzed for tone quality introduced anxiety or nervousness. Tell-tale signs of stress included:

- Decreasing duration of long-tones. Although instructed to play notes for 4-5 seconds, they were shorter. In some cases, the duration continued to decrease as the session progressed.
- Notes were “cracked” and replayed a few times before the subject continued to the next note.

- Some subjects spontaneously voiced confessions of feeling nervous.
- Some of the more experienced musicians made statements like: “it is difficult to sound pretty playing long-tones”, or “playing without vibrato is not normal”.
- For straight-tone notes, varying degrees of vibrato crept in. For vibrato notes, there was unevenness or some loss of control.
- A few individuals had difficulty playing low notes (in some cases the condition of flute may have contributed to the situation).

Manifestations of nervousness occurred for a subset of flutists across all levels of experience and skill. This may have precluded optimal tone production. The recorded notes may not be a true reflection of the flutist’s overall abilities. Rather, they are examples of tones produced in circumstances less than ideal for artistic expression.

All the subjects were instructed to play at a mezzo forte dynamic level. There was a considerable difference of loudness between musicians. Some of the less experienced musicians played quite softly. Some of the more experienced played quite loudly. However, there was no clear pattern, as some highly trained and experienced flutists played softly.

7.1.2 Spectra Variability

One unexpected outcome of the spectral analysis was the level of instability in some straight-tone notes. Figure 5.4, in section 5.2.1 (HAT Implementation), shows one example of this phenomena. Figures 7.1 and 7.2 show two additional examples. Notes like these, with large changes in the spectra, were not used for the survey. It would have been difficult for a listener to describe and rate the timbre since it is volatile.

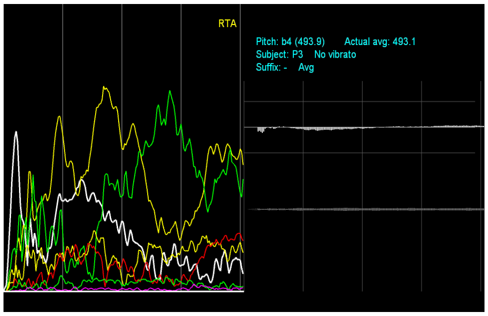


Figure 7.1: Flutist P3 playing B4

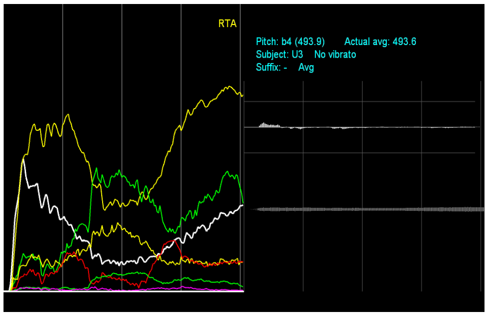


Figure 7.2: Flutist U3 playing B4

Part of the recording process required the subjects to play certain notes more than once. Specifically, the notes in a G major triad (D4, G4, B4, D5, G5, B5) were played multiple times. Appendix E.1 contains some examples. For any given flutist, there are noticeable timbre differences between notes played at the same pitch. Experienced flutists might not find this surprising. It is not uncommon for individuals to make adjustments as they attempt to find optimal tone quality. Further, flutists exercise flexibility and sometimes alter their tone color depending upon performance circumstances. However, in this situation, the changes in timbre were not intentional.

7.1.3 Spectra and Pitch

HAT spectral analysis confirms that as the pitch goes higher, the harmonic content tends to decrease. Figure 7.3 shows the spectra for two flutists playing a D4. Figure 7.4 show the spectra for each flutist playing a B5. The charts clearly show that B5, the higher pitch, has much weaker upper harmonics compared to each flutist's corresponding D4. Appendix E.2 shows the spectra for different flutists playing a variety of different pitches. Each flutist's timbre changes depending upon the pitch.

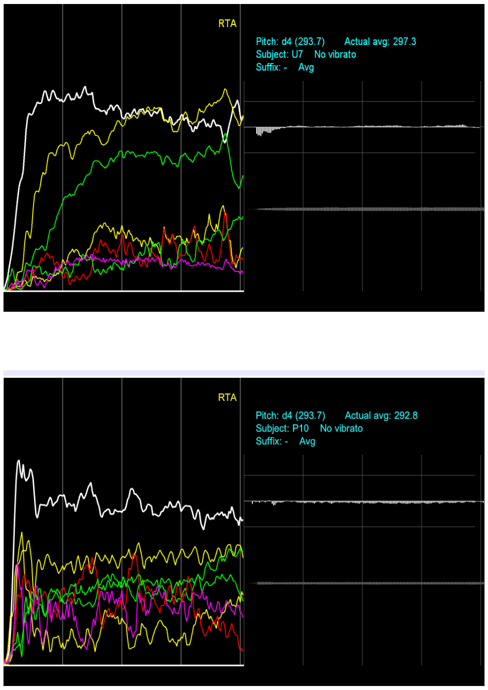


Figure 7.3: Flutist U7 and P10 playing D4

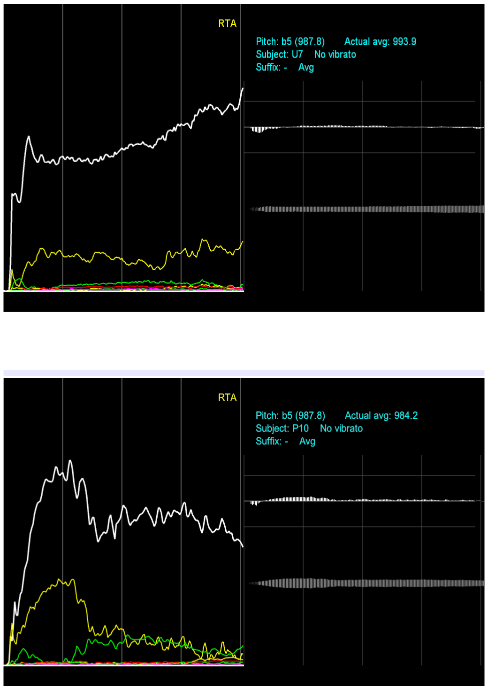


Figure 7.4: Flutist U7 and P10 playing B5

Harmonic signatures comparisons will be restricted to samples at the same "target" pitch levels. The distinction between a "target" pitch and the "same" pitch is an important one. Although a tuner was provided during the recording sessions, it is unreasonable to expect musicians to play each note at an exact frequency. Notes were generally close to the "target" pitch, but may have wavered in sharpness or flatness.

It is not uncommon for the low pitches of the flute to have spectra in which the fundamental frequency is not the strongest harmonic. Figures 7.5 and 7.6 show the spectra for two different flutists playing C4 (middle C). In both cases, the fundamental (white) harmonic is not the strongest.

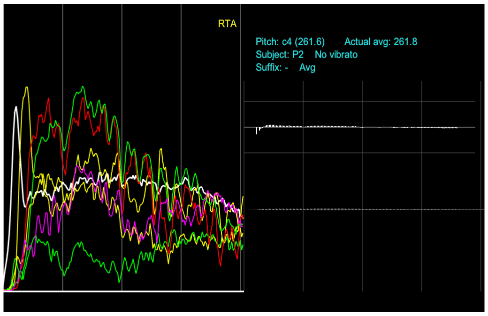


Figure 7.5: Flutist P2 playing C4

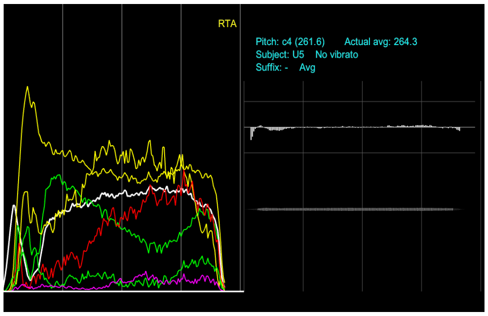


Figure 7.6: Flutist U5 playing C4

7.1.4 Non-flute Spectra

Some flute timbres have been described as sounding reedy, oboe-like, or brassy/trumpety. For comparative purposes, a professional trumpet player, oboist, and clarinettist were recorded. Spectra for some of their samples are included in Appendix E.3 for reference.

7.2 Survey Results

This section first covers the demographics of the survey participants that comprise the "skilled flutists". Then the timbre descriptors and ratings are explored. Finally, survey results from University musicians are presented.

7.2.1 "Skilled Flutist" Demographics

The Methodology chapter articulated the reasoning behind using the Flute List for survey input. There were a total of 121 participants from the Flute List, and the overall experience level is substantial. Demographics are included in Appendix D. A subset of these participants with the most experience and training was used for most of the analysis. For convenience, this subset will be referred to as the FL10s. The selection criteria for the FL10s was Flute List participants with: 10 or more years of teaching experience, 10 or more years of private lessons, and play/practice/rehearse 10 or more hours per week. Essentially, they are seasoned flute instructors with substantial private training that actively maintain their performance skills.

The FL10s comprise the “skilled flutists” and consists of 41 individuals. On the average they have been teaching for 26 years. and have studied privately for an average of 14.8 years. FL10s play, practice, and rehearse an average of 19 hours per week. 80% of the FL10s rated tone quality as highest priority (5 on a scale of 0-5). The balance rated this as 4, so all of the FL10s indicated tone quality as a high priority.

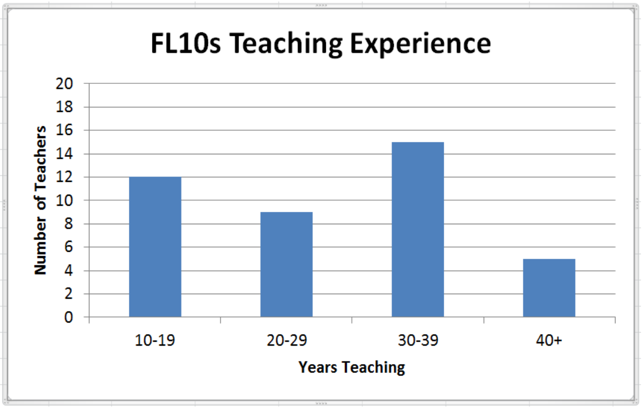


Figure 7.7: FL10s - teaching experience

7.2.2 FL10s Descriptors

The nouns “descriptors”, “adjectives”, and “terms” will be used interchangeably here. A set of frequently used descriptors was extracted from the FL10s survey responses to understand how “skilled flutists” describe tone quality. These terms were then categorized with their ratings to determine if they are considered favorable or unfavorable qualities. Any patterns revealed here apply to the survey samples, and may not be generally applicable.

The ten most frequently used descriptors are shown in table 7.1. The rating scale range was from 1 (“poor”) to 5 (“great”). A 3 would be considered a *neutral* rating. *Favorable* descriptors accompanied ratings of 4 or 5; *Neutral or favorable* descriptors were used with ratings of 3 or higher; *Unfavorable* descriptors were used with ratings of 2 or lower. The *Across all ratings* category is for descriptors accompanying all ratings (rating from 1 to 5). The entries are ordered from most to least frequently used.

Columns 1 and 2 are self explanatory. Column 3, *Count*, shows the number of times each descriptor was used by the FL10s. Column 4, *#People*, indicates the number of distinct FL10s participants who used that descriptor. Together, column 2 and 4 provide some indication of whether some individuals repeatedly used a particular descriptor across the survey samples. The last column gives a sense of whether the descriptor might be considered positive, negative, or non-determinant.

Table 7.1: FL10s Descriptor Usage

Ranking	**Descriptor**	**Count**	**#People**	**Category**
1	focused	52	22	neutral/favorable
2	airy	37	17	neutral/unfavorable
3	unfocused	31	19	unfavorable
4	edgy	29	16	across all ratings
5	clear	22	12	neutral/favorable
6	rich	20	12	favorable
7	weak	20	9	unfavorable
8	full	18	14	favorable
9	diffuse	15	9	across all ratings
10	open	14	8	across all ratings

Figure 7.8 provides some additional context by showing the FL10s rating

distribution. There were somewhat more unfavorable ratings resulting in an average of 2.76. The "great", or 5, was given sparingly relative to the other ratings.

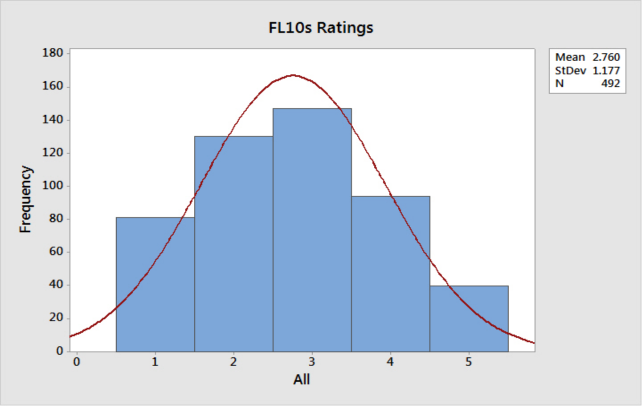


Figure 7.8: FL10s rating distribution

Another way to organize tone descriptors is based upon ratings. Table 7.2 shows sets of adjectives for various rating categories. The parenthetical numbers indicate the number of times each term was used across all the survey samples. This table contains 34 descriptors, and all of them were used at least 5 times.

Table 7.2: FL10s Descriptors

Circumstances	**Descriptors**
Favorable	rich(20), full(18), resonant(5), colorful(4)
Neutral or favorable	focused(52), clear(22), round(8), bright(8), dark(7), buzz(5)
Neutral or unfavorable	airy(37), forced(14), harsh(13), hollow(12), soft(12), dull(12), overblown(10), brassy/trumpety(6), lacking-core(6)
Unfavorable	unfocused(31), weak(20), thin(14), unsupported(11), sharp(9), breathy(8), muffled(6), nasal(6), uncontrolled(6), uncentered(5)
Across all ratings	edgy(29), diffuse(15), open(14), loud(11), warm(5)

Although the terms "bright" and "dark" seem to have opposite meaning, they appear to be used interchangeably describing samples rich in harmonics. Terms like "sharp" (pitch vs. point/edge) and the survey comments did not always provide sufficient context to clearly interpret what was intended. The *Across all ratings* category may reflect personal taste. For example, some individuals may or may not prefer an "edgy" or "diffuse" tone quality.

Another way to view the descriptors is to focus on the terms associated either "great" or "poor" ratings. Since these ratings are at the extreme ends of the scale, they indicate either very positive or very negative reaction to a timbre. For the FL10s, the "great" rating occurred 40 times, and the top three descriptors were "focused", "clear", and "rich". The "poor" rating occurred 81 times, and the most frequently used descriptors were "airy", "unfocused", and "weak". Table 7.3 shows the frequency of these terms. The percentage indicates how often the term was used. For example, there were 22 occurrences of the term "focused" within the 40 "great" ratings; $22 \div 40 = 55\%$.

Table 7.3: FL10s Descriptors for "great" or "poor" ratings

Descriptor	**Occurrence**	**Percent**
focused	22	55%
clear	8	20%
rich	8	20%

Descriptor	**Occurrence**	**Percent**
airy	13	16%
unfocused	12	15%
weak	10	12%

Within the 12 survey samples, an important criteria for the FL10s is whether the timbre sounded focused or unfocused. Samples perceived as focused, clear, or rich were rated positively. Samples that were airy, unfocused, or weak were rated negatively.

This closing paragraph of the descriptor analysis takes a brief interlude from the FL10s. Appendix G contains correlation analysis of descriptors from the complete Flute List survey responses. It is a programmatic approach that examines a larger data space. It shows a high correlation between the term focused with: clear, strong, and supported. The term unfocused was highly correlated with:

breathy, diffuse, fuzzy, and airy. It also contains uncorrelated terms; for example the descriptor edgy is not correlated with: unfocused, dull, soft, weak, airy, unsupported, diffuse, breathy, open, mellow, hollow, and fuzzy. See the appendix for more details.

7.2.3 FL10s Ratings

The average ratings for each sample provides a high level view of tone quality preference. These are shown in figure 7.9. Both the complete Flute List and the FL10s averages are included, and the results are similar. Some of the ratings, like 1B, 2A, 3B, and 4A, are slightly lower for the FL10s. A few of the others are slightly higher.

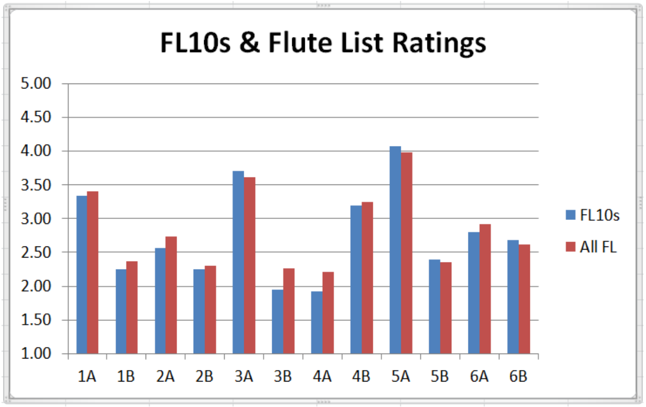


Figure 7.9: FF10s and Flute List ratings

The FL10s results ranged from 1.93 to 4.07. If 3 is considered average, 8 of the 12 samples were below average, and the remaining 4 were above average. Charts showing the rating distributions for each sample are in Appendix F.

The FL10s ratings are summarized in table 7.4. The results are sorted by lowest to highest rating. There are two categories of timbre that FL10s rated unfavorably:

1. Tones perceived as weak, airy, or unfocused (4A, 3B, 1B, 5B, 2A, 6A).
2. Tones perceived as excessive in some form: overblown, edgy, or brassy/trumpety (2B, 6B).

Table 7.4: FL10s summary of ratings and descriptors

Rating	**Sample**	**Pitch**	**Great**	**Poor**	**Descriptors**
1.93	4A	G4	0	16	airy, weak, uncontrolled
1.95	3B	G4	0	14	unfocused, breathy/airy, thin
2.24	1B	D4	1	12	slightly-diffuse vs. unfocused, weak, airy
2.24	2B	D4	0	11	sharp, overblown, edgy
2.39	5B	B4	2	6	unfocused, weak, airy
2.56	2A	D4	2	8	unsupported, airy, thin
2.68	6B	D5	1	8	refined vs. nasal, brassy/trumpety, edgy
2.80	6A	D5	1	3	bright vs. unfocused, weak, airy
3.20	4B	G4	4	3	rich, focused, full vs. forced, nasal, trumpety
3.34	1A	D4	4	0	focused, dark/bright, rich
3.71	3A	G4	10	0	focused, clear, rich
4.07	5A	B4	15	0	focused, clear, full

Sample 4B had a range of responses and was rated favorably. Although it was sometimes described as forced or trumpety, it also received favorable remarks like rich or focused. The remaining three samples (1A, 3A, 5A), were rated favorably and had descriptors like: focused, dark/bright, rich, or clear.

The following subsections examine the spectra, ratings, and descriptors in greater detail. The first subsection looks at the pitch G4 which has the two lowest rated samples. Then pitch D4 is analyzed since it has the next two lowest rated samples. The last subsection continues to B4 and D5 in a similar vein.

7.2.3.1 G4 Spectra

The two lowest rated samples have a target pitch of G4. Figure 7.10 shows the harmonic signatures for all G4 samples. The images are ordered left to right, top to bottom based on rating. 4A and 3B were rated the lowest and neither received any "great" ratings. They share some common descriptors for the unfavorable ratings (2-ratings plus "poor" ratings): unsupported, weak, and unfocused. They were never described as dark or bright. Visually, both have a strong H1 (white). 3B also has a very strong H2 (yellow). Both have very little H3 (green) or H5 (red). There is a gap between the strongest harmonics and the upper harmonics.

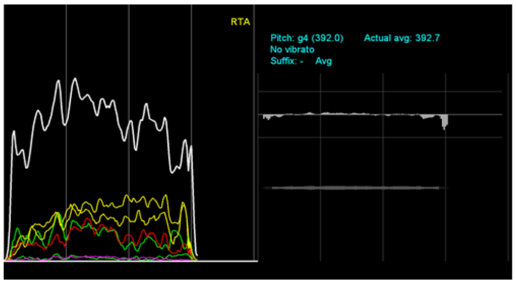
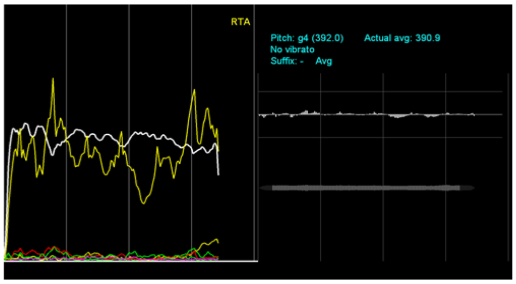
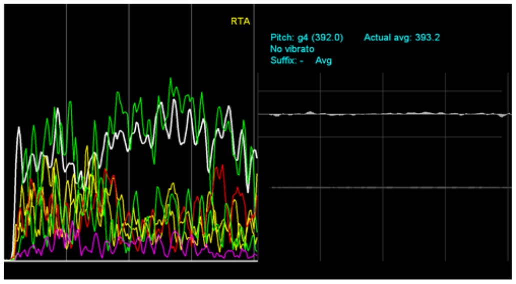
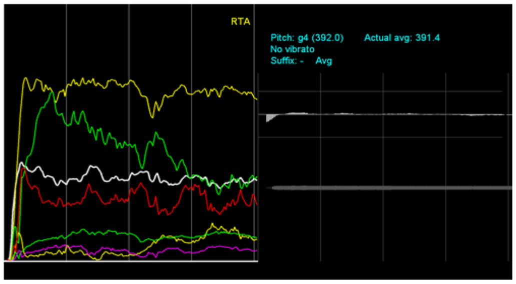
Pair 4A (G4) - Rating 1.93	**Pair 3B (G4) - Rating 1.95**
Great (0): n/a **Poor (16): airy, weak, uncontrolled**	**Great (0): n/a** **Poor (14): unfocused, breathy/airy, thin**
	
Pair 4B (G4) - Rating 3.20	**Pair 3A (G4) - Rating 3.71**
Great (4): rich, focused, full **Poor (3): forced, nasal, trumpety**	**Great (10): focused, clear, rich** **Poor (0): n/a**
	

Figure 7.10: FL10s g4 spectra ratings

In contrast, 3A was the second most favored sample in the survey. The harmonic signature is very rich in harmonics. This is a case where the H2 and H3 are stronger than the fundamental. Some of the "great" descriptors were: focused, clear, and rich. Some of the other favorable terms include full, round, and bright/dark.

Interestingly, 4B has diverse results. Individuals that liked 4B used terms like: rich, focused, and full. Those that dislike 4B use adjectives like: forced, nasal, and trumpety. Another descriptor often used for 4B was edgy. H1 and H3 are very prominent in the harmonic signature, and all of the harmonics fluctuate considerably. The room used for recording 4B's sample had considerable reverb. It is not certain if this contributed to the fluctuating harmonics.

7.2.3.2 D4 Spectra

The third and fourth lowest rated samples have a target pitch of D4. Both 1B and 2A were rated unfavorably and share some common descriptors with the lowest rated G4 samples (airy, unfocused, thin). For 1B and 2A, either H1 or H2 is very strong, and then there is a gap. The upper harmonics comprise a relatively small portion of the mix for these two samples. D4 is near the bottom of the flute range where some flutists' spectra have a rich mix of upper harmonics. Neither of these two samples exhibit this characteristic.

2B was also rated unfavorably. Unlike the samples that were disliked for being airy or unfocused, 2B was judged as being sharp, overblown, and edgy. There is an unusually strong H3 dominating the signature. In this case, H1, H2, and H3 are all stronger than the fundamental.

1A was the only D4 sample that was rated somewhat positively. Some of the favorable descriptors included: focused, dark/bright, and rich. The term edgy was also used for this sample. Visually, the spectra is rich in harmonics. H7 (purple) is unusually strong. Some of the trumpet samples in Appendix E.3 show a high H7 content.

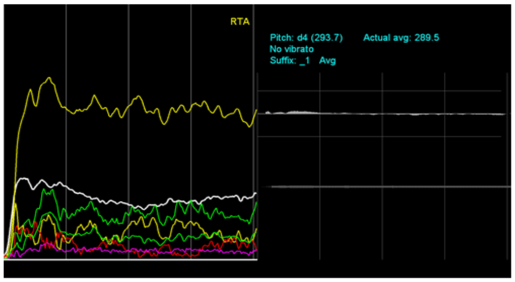
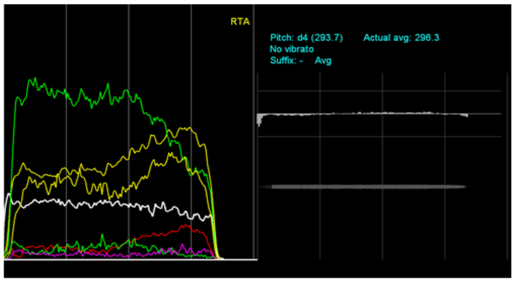
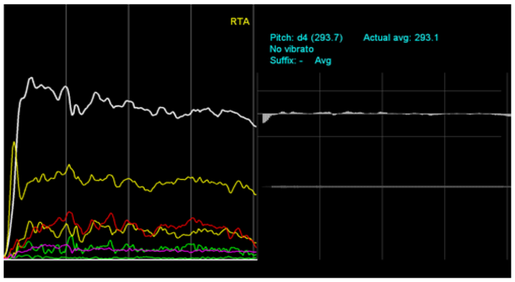
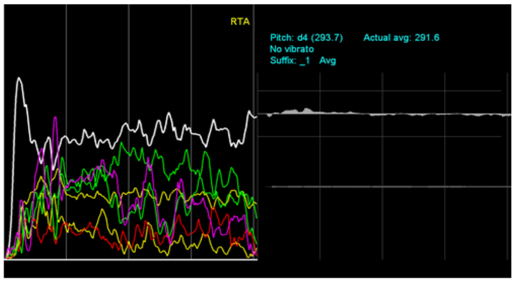
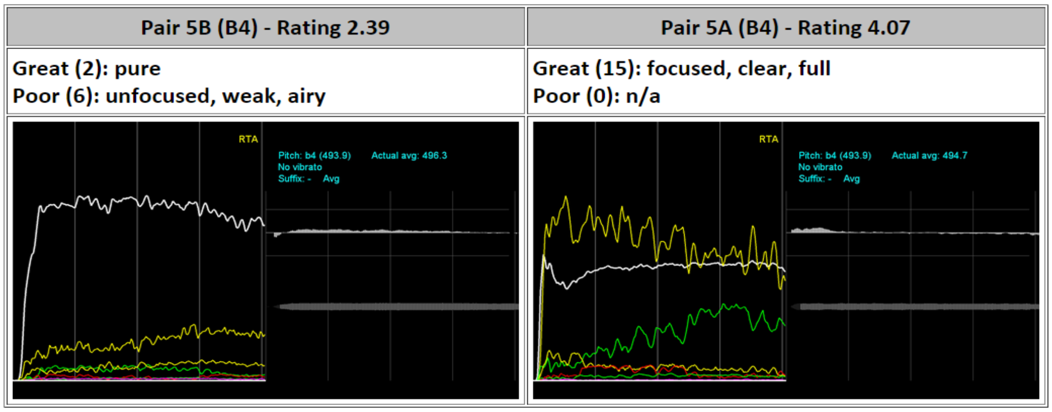
Pair 1B (D4) - Rating 2.24	**Pair 2B (D4) - Rating 2.24**
Great (1): slightly-diffuse **Poor (12): unfocused, weak, airy**	**Great (0): n/a** **Poor (11): sharp, overblown, edgy**
	
Pair 2A (D4) - Rating 2.56	**Pair 1A (D4) - Rating 3.34**
Great (2): centered **Poor (8): unsupported, airy, thin**	**Great (4): focused, dark/bright, rich** **Poor (0): n/a**
	

Figure 7.11: FL10s d4 spectra ratings

7.2.3.3 B4 and D5 Spectra

The target pitch B4 has one unfavorably rated sample and the other sample was the highest rated in the survey. 5B received mainly unfavorable or neutral ratings. Like some of the other samples with low ratings, 5B has descriptors like: unfocused, weak, and airy. It can be characterized as having a dominant H1 with little presence of upper harmonics.

5A had the highest ratings with “great” descriptors like: focused, clear, and full. Other adjectives include: vibrant, rich, and resonant. It is tempting to visually compare the spectra for 5A with 3B as they share some similarities, but 5A has a greater presence of H3. However, since 3B is a lower pitch it is probably

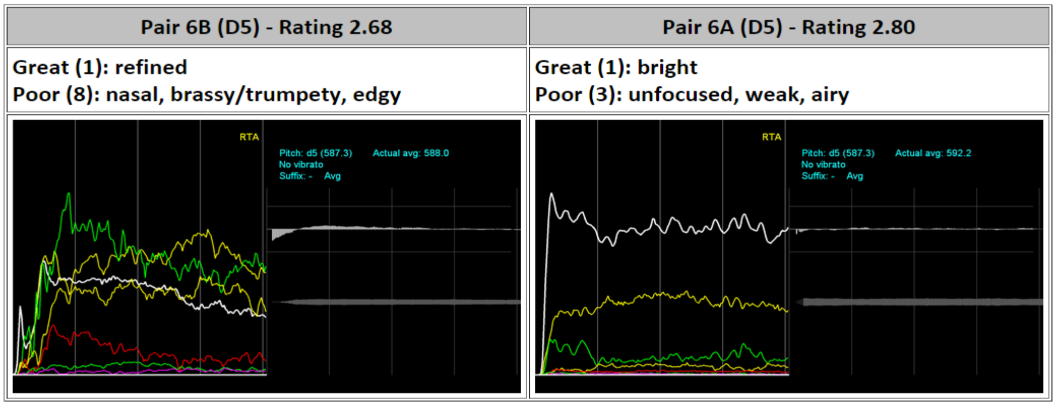


Pair 5B (B4) - Rating 2.39	Pair 5A (B4) - Rating 4.07
Great (2): pure Poor (6): unfocused, weak, airy	Great (15): focused, clear, full Poor (0): n/a

Figure 7.12: FL10s survey results for pitch B4

not appropriate to correlate the signatures.

The final pair of notes had D5 as the target pitch. 6B contains an unusually strong H3. Among the 31 flutists, it was very rare to see H3 as the dominant harmonic for pitches in the second flute octave. Although the ratings were generally neutral to unfavorable, more than 25% of the ratings were favorable. Those that liked 6B often used descriptors like: focused, and clear. The negative terms included: nasal, brassy/trumpety, and edgy. This is a case where some common descriptors accompanied both favorable and unfavorable ratings. The adjectives edgy and brassy/trumpety are examples of this type of descriptor.



Pair 6B (D5) - Rating 2.68	Pair 6A (D5) - Rating 2.80
Great (1): refined Poor (8): nasal, brassy/trumpety, edgy	Great (1): bright Poor (3): unfocused, weak, airy

Figure 7.13: FL10s survey results for pitch D4

6A was close to neutral in terms of ratings. There were some interesting contradictions in descriptors: focused versus unfocused, bright versus dull/pale. The bulk of the adjectives were: unfocused, weak, and airy.

7.2.3.4 Trends in Ratings and Harmonic Signatures

Since there are only 12 samples, and these are subdivided into 5 different target pitches, it is not prudent to make sweeping generalizations from these results. However, there are trends that can be observed within the survey data:

- Samples with strong H1 and/or H2 and relatively weak upper harmonics (H3-H7), were not rated highly by experienced flutists. Visually, there is a gap between the prominent harmonics and the upper harmonics. Samples of this type were seldom described as bright/dark, or edgy. Rather, these samples are often described as weak, airy, or unfocused.
- Samples with a strong H3 were described as edgy, nasal, or trumpety/brassy. When H3 was disproportionate, the samples received negative ratings.
- Samples with a balance of harmonics received favorable ratings.
- For the FL10s, the descriptors bright/dark were used mainly for the top 2 samples. However, the full Flute List used these terms more liberally for other samples containing a high level of H3 (2B, 6B, 4B, 5A, 1B).

7.2.4 Cal Poly Ratings

An early Cal Poly pilot survey showed divergent ratings from the professional interview sessions. In order to determine if this was an anomaly versus a trend, additional surveys were administered to Cal Poly students to gather more data.

A total of 103 students participated in the subsequent Cal Poly surveys. 37

students were non-musicians. Of the 66 musicians, 26 were actively enrolled in Cal Poly music ensembles. Figure 7.14 compares the ratings from musicians versus non-musicians. Overall, the ratings are fairly similar between these two categories of students.

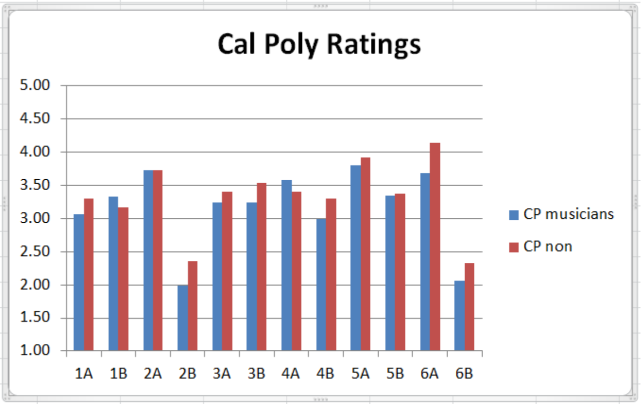


Figure 7.14: Cal Poly ratings

Figure 7.15 compares FL10s and Cal Poly musicians (CPM) ratings. These groups represent the more experienced and more highly trained subsets of each respective survey set. The samples are sorted according to FL10s rating results. There are clearly differences between FL10 and CPM preferences. The overall mean for CPM is 3.17 which is higher than the FL10s mean of 2.76. It is not that surprising that FL10s has a lower average; the FL10s are flute instructors and would have a critical ear when evaluating flute timbre.

Surprisingly, all the samples that the FL10s described as weak, airy, unfocused (4A, 3B, 1B, 5B, 2A, 6A) were more acceptable to CPM. In fact, all of these samples received unfavorable rating from FL10s, but received favorable ratings from the CMP. For example, the FL10s lowest rated sample, 4A, received a favorable rating by the CPM. The two-sample t-test was used to calculate the statistical significance of the rating difference beween the FL10s and CPM (see

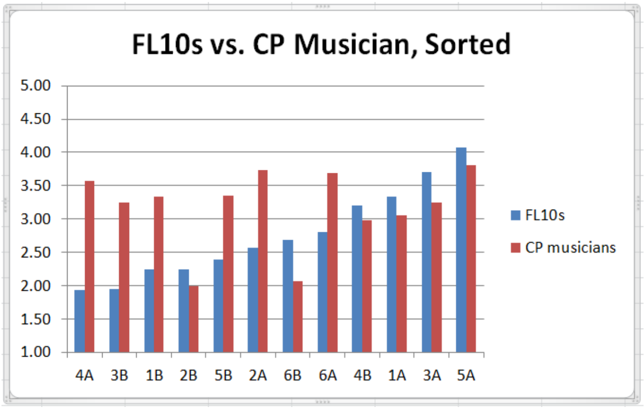


Figure 7.15: FL10s vs. Calpoly sorted ratings

Table 7.5: FL10s and CPM ratings and p-values

Sample	Pitch	FL10s	CPM	Difference	p-value
4A	G4	1.93	3.58	-1.65	0.000
3B	G4	1.95	3.24	-1.29	0.000
1B	D4	2.24	3.33	-1.09	0.000
2B	D4	2.24	1.98	0.26	0.173
5B	B4	2.39	3.35	-0.96	0.000
2A	D4	2.56	3.73	-1.17	0.000
6B	D5	2.68	2.06	0.62	0.005
6A	D5	2.80	3.68	-0.88	0.000
4B	G4	3.20	2.98	0.21	0.355
1A	D4	3.34	3.06	0.28	0.137
3A	G4	3.71	3.24	0.46	0.027
5A	B4	4.07	3.80	0.27	0.117

table 7.5). A rule-of-thumb is if the two-sample t-test p-value is less than 0.05, the difference can be considered significant. For this set of samples (4A, 3B, 1B, 5B, 2A, 6A), all of the p-values are 0.000, clearly indicating statistical significance.

In contrast, the two samples that the FL10s describe as overblown or edgy (2B, 6B), received lower scores from the CPM. The rating difference for 2B is not significant, but it is for 6B with a p-value of 0.005. All of the samples that FL10s rated favorably (4B, 1A, 3A, 5A) received slightly lower ratings from the

CPM. Of these, only sample 3A might be considered statistically significant with a p-value of 0.027.

The results indicate that CPM preferred the tones with less harmonic content than FL10s. Within the context of the survey results, highly trained and skilled flutists have a different criteria for tone quality than CPM. This raises some interesting questions:

- Do only highly trained/skilled flutists really appreciate the tone quality of accomplished flutists?
- Does the general public actually prefer flute tones with less harmonic content?

7.3 Vibrato and Dynamics Analysis

Although the primary focus of this research is how experts describe and rate tone quality, there are some other interesting aspects that emerged. The first area is vibrato analysis. The second topic is about flute timbre and dynamics.

7.3.1 Vibrato

The survey was based on straight-tone notes. This approach established some correlation between harmonic signatures and tone quality unencumbered by the complexity of vibrato. However, vibrato merits some analysis as it is commonly used by skilled flutists.

Figure 7.16 shows two different flutists, P2 and P4 playing the pitch G4. When flutist P2 adds vibrato, there is an increase in the H2 and H4 (yellow) and a decrease in H3 (green). For flutist P4, adding vibrato results in a substantial drop

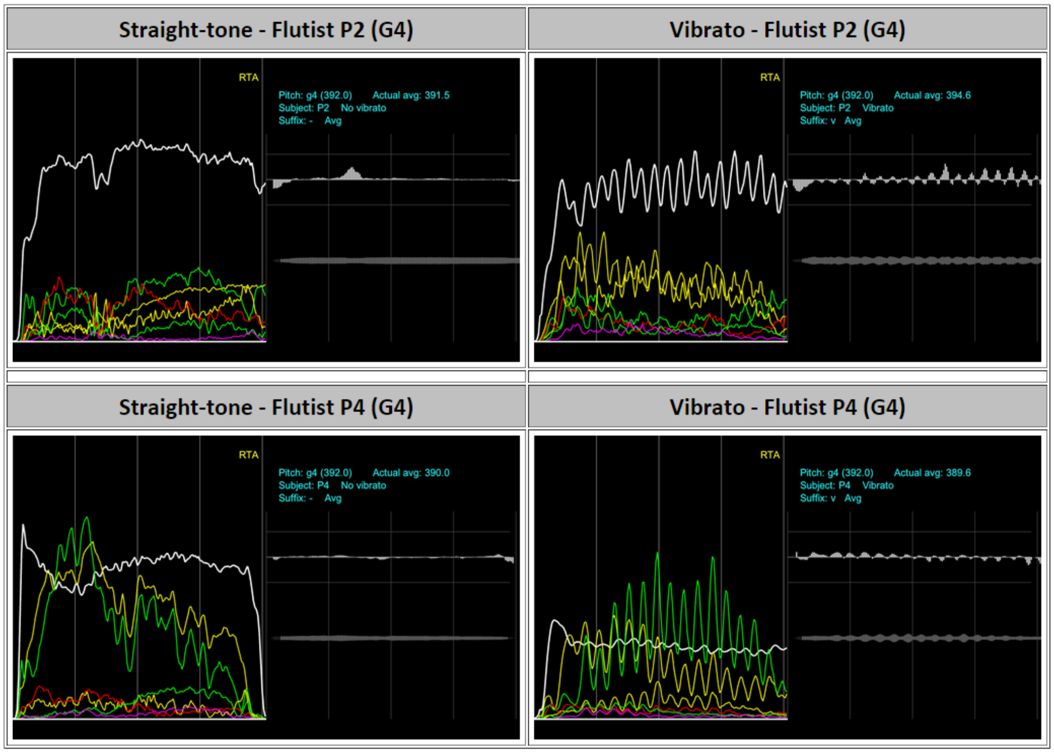


Figure 7.16: Straight-tone and vibrato

in H1, and a reduction in H2. P4's H3 harmonic has a large range of oscillation, but H1 is nearly stable. For these two examples, the harmonics oscillate in a similar pattern. In the HAT Implementation chapter, figure 5.5 shows an example where some of the harmonics oscillate at a different phase.

The pitch graph (upper right) shows that both flutists' pitch oscillates with vibrato. The range of pitch oscillation varies, as does the degree of sharpness or flatness. This observation applies, both within each sample, as well as between the two flutist's vibrato samples. The loudness/amplitude graph(lower right) shows both flutists' vibrato results in a pulsing in terms of loudness.

Looking at these two examples for one pitch (G4), there are several differences in:

- Which harmonics increase or decrease in strength when vibrato is introduced
- The degree that each harmonic oscillates
- The amount of pitch variation

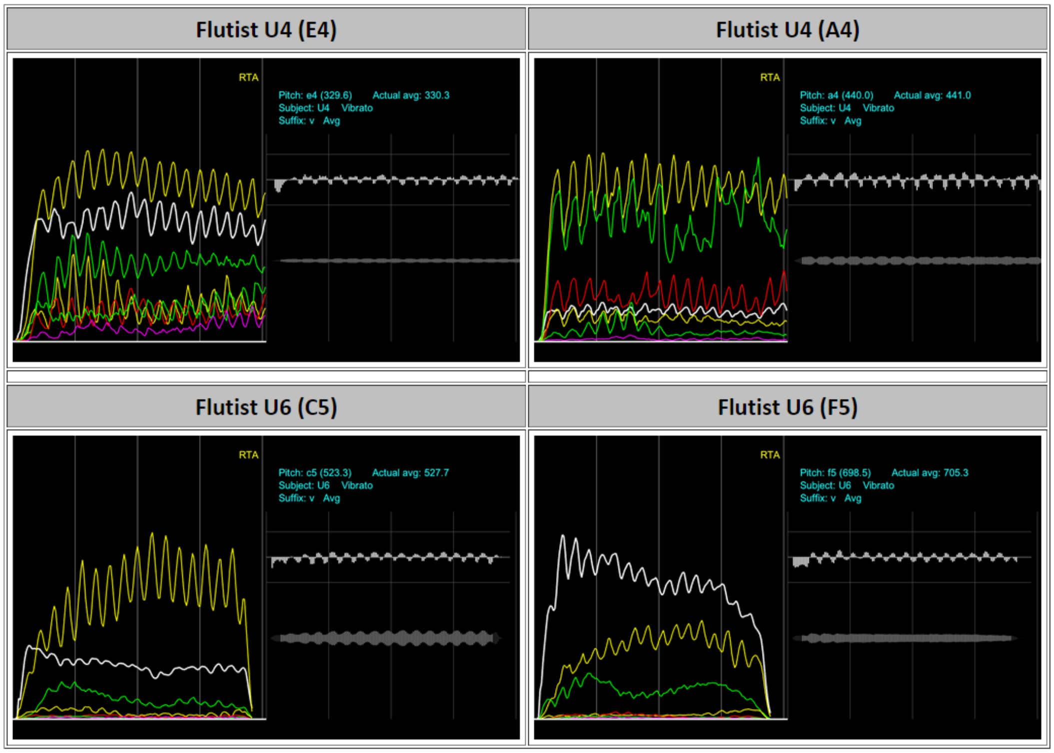


Figure 7.17: Pitch and vibrato

Figure 7.17 shows examples of two flutists' vibrato at different pitches. The first row show flutist U5 playing an E4 and and A4. Aside from the differences in signature and harmonic oscillation range, there is the the difference in the amount of pitch range variation. In this situation, U5 has a wider pitch variation at the higher pitch (A4). The higher pitch is also somewhat louder with a noticeable difference in amplitude variation.

Flutist U6 shows a situation in which the amplitude/loudness oscillations are relatively deep for the lower pitch (C6). The higher pitch (F5) uses mainly harmonic and pitch vibrato, but little amplitude vibrato. The amplitude/loudness graph (on the lower right) for C5 shows a fairly even pulsing throughout the duration of the note. In contrast, when U6 played F5, the note started out with some loudness pulsing, but then it is hard to visually detect any pulsing after 1 second.

It is not practical to show comparative examples for a broad range of pitches for all the flutists. Although not shown here:

- There are differences between the various flutists that were recorded.
- There are also differences within each flutists' vibrato at different pitches.
- There are even changes for a given flutist playing the same pitch multiple times with vibrato.

These differences include various combinations of harmonic (color), pitch, and amplitude (loudness) variations. The variations occur within a note, as well as between notes.

Using only a few samples, it is apparent that the impact of vibrato on harmonic content is quite complex. A comprehensive analysis of vibrato and tone quality is beyond the scope of this project.

7.3.2 Dynamics

All of the professional flutists, and most of the university flutists, were asked to play a set of long-tones using dynamics. The musicians were instructed to play selected notes for approximately 8 seconds each. The notes started at a pianissimo level with a gradual crescendo to fortissimo. Each note was played first with straight-tone, and then with vibrato.

Sustaining a note for 8 seconds with a gradual and smooth crescendo requires

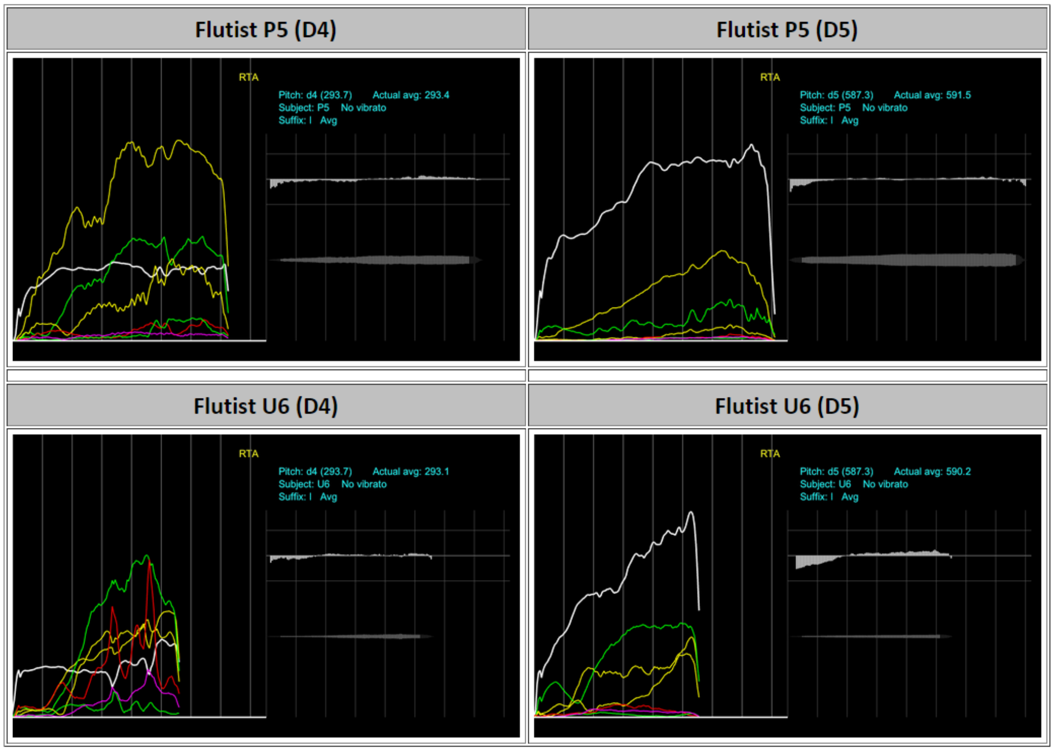


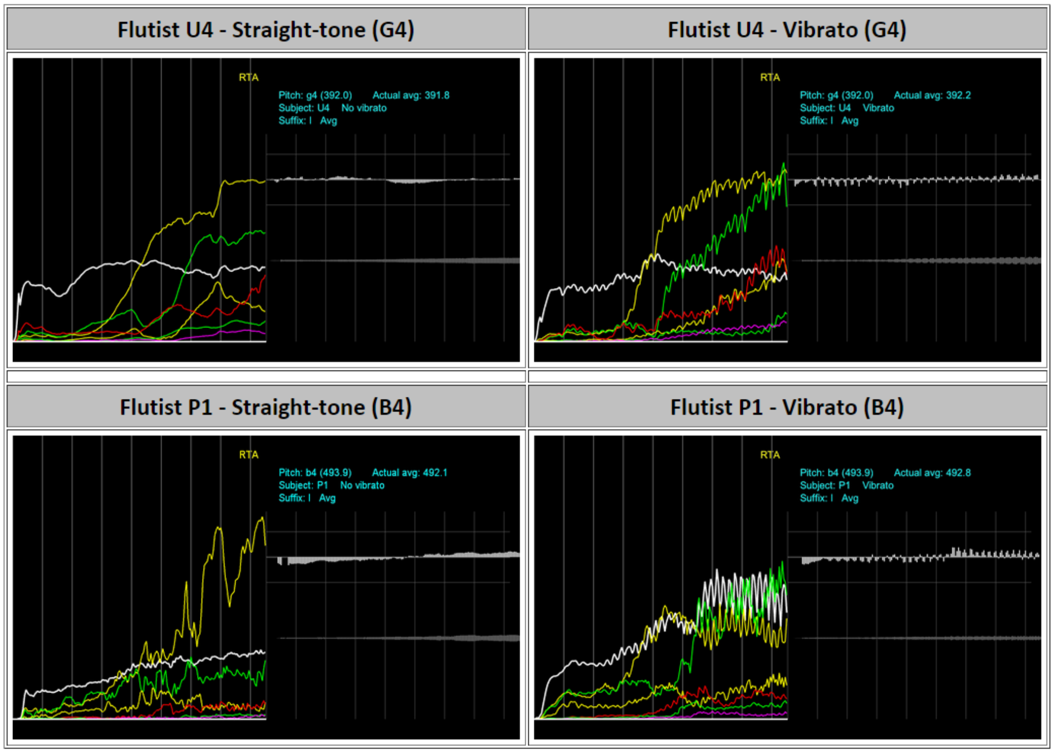
Figure 7.18: Dynamics, straight-tone D4 and D5

skill and practice, a task that proved demanding for many of the musicians. It was not usual to hear uneven changes in dynamics and pitch. Many of the flutists were only able to produce a narrow range of dynamics. In a few cases, there was little perceptible change in volume as the note supposedly progressed from soft to loud.

Figure 7.18 shows two flutists, P5 and U6, playing the pitches D4 and D5 using straight-tone. Flutist P5's lower pitch (D4) shows increasing H2, H3, and H4 as the note increases in loudness. In contrast, at the higher pitch (D5), harmonics H1 and H2 increase. Many of the samples with dynamics have a general trend where the pitch grows sharper as the note gets louder.

For flutist U6, the lower pitch note (D4) shows harmonics that are much different than P5. H2, H3, and H5 grow with the crescendo. At the higher pitch

(D5), U6's H1 and H5 grow as the note progresses. The amplitude/loudness indicator shows that P5 has a larger measured range of amplitude/loudness than U6. P5 is using both loudness and timbre to achieve the crescendo.


Flutist U4 - Straight-tone (G4)

Flutist U4 - Vibrato (G4)

Flutist P1 - Straight-tone (B4)

Flutist P1 - Vibrato (B4)

Figure 7.19: Dynamics, straight-tone and vibrato

Figure 7.19 shows two examples of using vibrato with dynamics. Flutist U4's straight-tone and vibrato harmonics are fairly similar. There are many samples where there is a considerable change in the harmonic mix when vibrato is combined with dynamics. Flutist P1 shows a situation where the vibrato introduces noticeable differences from the straight-tone harmonics. For P1, as the crescendo progresses, H1 and H3 play a larger role in the timbre than the straight-tone sample.

All of the spectra shown in this section are relatively organized. There are many samples, not shown here, where the harmonics change in a chaotic manner as the notes crescendo. The samples shown here show that the timbre often changes with dynamics. The easiest way to observe this is by comparing the relative mix of harmonics on left side of each line chart versus the right side. Like the vibrato analysis, a comprehensive exploration of dynamics is beyond the scope of this project.

7.3.3 Syrinx

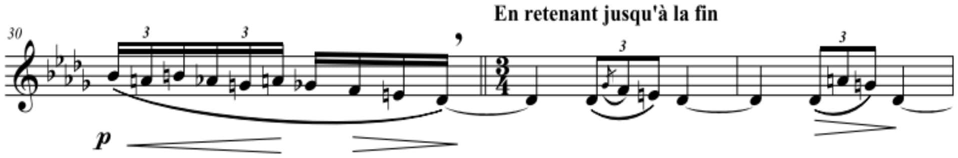



Figure 7.20: Syrinx by Claude Debussy [10]

This section applies spectral analysis for flute tones within a specific musical context. Syrinx is a flute piece composed by Claude Debussy in 1913 [10]. Since it is a solo work, extracting long-tones from recordings is relatively straightforward. Near the end of the work, there is a low D♭ that spans measure 31-32 (figure 7.20) that will be examined here.

During various phases of this research, participants were asked to name their favorite flutists. Appendix D contains the complete list. The three most frequently mentioned artists were Emmanuel Pahud, Jean-Pierre Rampal, and Sir James Galway. The harmonic signature of the D♭ for these top three flutists, plus an amateur, are shown in figure 7.21.

In terms of dynamics, Galway is applying a decrescendo and Pahud is using a crescendo. Rampal and the other musician maintained a fairly constant dynamic level. For vibrato, Pahud starts with a straight-tone and applies increasing vi-

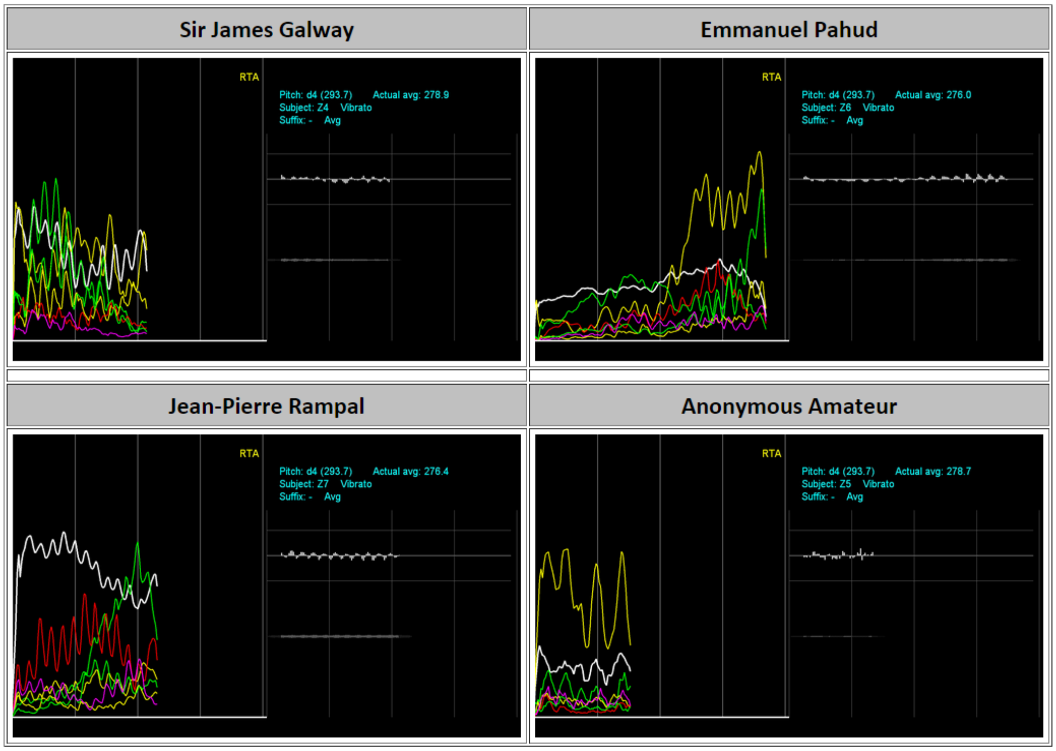


Figure 7.21: Syrinx spectra

brato (both depth and speed) as the note crescendos. The other musicians use a relatively constant vibrato throughout the note.

A visual inspection reveals differences in the harmonic signature, both within each note, as well as across musicians:

- As Pahud applies the crescendo, H2 and H3 increase dramatically. In contrast, the H1 remains relatively stable throughout note.
- Although Rampal maintains a relatively constant dynamic level, the timbre changes towards the end of the note with H3 spiking. As H3 increases, H1 and H5 taper off. Rampal's fundamental is prominent, and also shows a very high H5 content.
- Galway's harmonics are tightly intermingled, with H3 and H6 both tapering off towards the end of the note. During phases of the note, the lines are

interlaced making it difficult to trace individual harmonics.

- The amateur flutist has a dominant H2 with clear separation from the other harmonics.

It would be informative to conduct a follow up survey using samples like these.

CHAPTER 8

Conclusions

There are several outcomes from this project. The three thesis questions are addressed in the first section. The second section summarizes an interesting trend in flute tone preference by Cal Poly University musicians. Aside from the thesis questions, other observations from this study are itemized in the third section. The fourth section covers future work, and finally, the last section contains a few closing comments.

8.1 FL10s

The FL10s' survey results were used to address the thesis questions. These 41 flutists are a subset of the Flute List survey participants that have significant teaching experience, received considerable private training, and have maintained their playing skills. Within the context of the 12 survey samples, a set of descriptors emerged. The top ten descriptors, ordered by frequency, are:

1. focused
2. airy
3. unfocused
4. edgy
5. clear
6. rich
7. weak
8. full

9. diffuse
10. open

There are associations between these descriptors and tone preference. The most commonly used descriptors for favorably rated samples are: focused, clear, and rich. The unfavorably rated samples are generally accompanied with descriptors like: airy, unfocused and weak.

There are some adjectives that were used across all ratings: edgy, diffuse, open, loud, and warm. These terms may represent attributes that reflect personal taste.

There are associations between spectral characteristics and some descriptors:

- Samples that contain a balance of harmonics were described as: focused, rich, bright/dark, or clear.
- Samples lacking upper harmonics have descriptors like: airy, unfocused, weak,or thin.
- Samples with unusually strong H3 have adjectives like: edgy, nasal, brassy/trumpety, or forced.

There are two categories for the unfavorable tones:

- Samples with few upper harmonics (H3-H7)
- Samples with excessive or out-of-balance harmonics (generally excessive H3)

8.2 CPM

Cal Poly Musicians (CPM) show divergent preferences from the skilled flutists. In particular, there is a statistically significant difference for samples that the FL10s rated poorly. While FL10s disliked samples with weak upper harmonics, the CPM liked these samples. Further, although not statistically significant, all of the notes that the skilled flutists rated favorably were rated slightly lower by students.

This difference in preference was unexpected. It is unclear if this is indicative of differences in how highly skilled flutists perceive tone quality from the general public.

8.3 Other Observations

This study revealed that there can be surprising amounts of instabilities in straight-tone notes. Harmonic signatures can also have noticeable differences when a given flutist plays the same pitch multiple times. This project also verified that the harmonic signature changes with pitch level; as the pitch moves higher, the upper harmonics tend to decrease.

Vibrato is very complex, and it is not uncommon to see:

- Variations in the pitch, amplitude, and harmonic ratios
- Different oscillation phases for various harmonics
- Different amplitudes of oscillation for various harmonics
- Different levels of pitch, amplitude, and/or harmonic oscillation depending upon the pitch of a note; or even when the same pitched note is repeated

Behavior of the harmonics are also quite complex with dynamics. There can be very limited actual changes in amplitude/loudness. Rather, in many cases, the harmonic profile changes as a note *crescendos*, but the amplitude/loudness does not increase noticeably. The harmonic signature changes can be chaotic as the dynamics increase.

8.4 Future Work

Future work falls into three categories: acquisition of flute tones, gathering additional survey data, and other analysis.

8.4.1 Acquisition of Flute Tones

Chapter 6, covered some of the challenges recruiting professional flutists as recording subjects. Acquiring samples from additional highly skilled flutists, perhaps from the “favorite flutist” list, could yield valuable insights. Ideally, sessions would be conducted in a recording studio administered by sound engineers. This would help ensure high quality recordings. Also, the recording methodology could be expanded to include short musical phrases containing sustained notes. Using phrases from well known flute repertoire would establish a musical context. Musical context would benefit both the individuals recording the tones, as well as those evaluating the tone quality.

8.4.2 Additional Survey Data

The surveys for this study used a limited set of only 12 straight-tone samples that were arranged in pairs. It is unclear if pairing these samples introduced biases with respect to the ratings and descriptors. Although useful results were derived from this methodology, administering additional surveys to skilled flutists using a broader range of samples is desirable. Samples could include a wider range of pitches that include vibrato, dynamics, and short phrases. Further surveys could also organize audio samples differently than paired notes of the same pitch.

8.4.3 Other Analysis

Further analysis of the existing data could reveal more discoveries. There are more than 1,600 audio samples, and each sample consists of a wealth of information. Every sample can produce a vector of data for each analysis window. This includes the frequency, seven harmonics, and overall amplitude. All this data may be helpful understanding how these complex and interrelated factors impact perceived tone quality. Given the large amounts of data, there are opportunities to apply machine learning techniques. These techniques can be applied to not only the spectra data, but to survey results.

HAT shows the first seven harmonics, along with pitch and amplitude. There are potentially other factors affecting timbre that are outside of these measurements. For example, descriptors like airy or breathy may be describing background ambient noise, or wind noise from the flutist's air jet. These types of sound would not neatly fall into the FFT bins used to produce the spectral charts. They represent qualities not captured by the approach used for this study. It would be prudent to investigate factors that impact timbre perception which are not captured by the current HAT implementation.

8.4.4 Final Words

One of the objectives of this project was to understand if there are quantifiable aspects for "good" tone quality. While there appears to be some correlations, much more work is required towards this goal. The unstated, and underlying question motivating this project was: given some metric for "good" tone quality, can a tool like HAT help musicians improve their tone quality more efficiently? Further research along this avenue could benefit aspiring musicians.

BIBLIOGRAPHY

[1] Audacity. Free open source software for recording and editing sounds. `http://audacity.sourceforge.net`. Accessed: October 2014.

[2] John Backus and John W Coltman. The acoustical foundations of music. *Physics Today*, 23(5):69–70, 2008.

[3] Robert David Billington. A description and application of Robert Aitken's concept of the physical flute. 2000.

[4] Murray Campbell and Clive Greated. *The musician's guide to acoustics.* Oxford University Press, 1994.

[5] Jer-Ming Chen, John Smith, and Joe Wolfe. How players use their vocal tracts in advanced clarinet and saxophone performance. In *Proc. Int. Symp. Music Acoustics*, 2010.

[6] Jennifer Cluff. Flute teacher and blogger. `http://www.jennifercluff.com/reading.htm`. Accessed: August 2014.

[7] John W Coltman. Effect of material on flute tone quality. *The Journal of the Acoustical Society of America*, 49(2B):520–523, 1971.

[8] John W Coltman. Just noticeable differences in timbre of the flute. *Catgut Acoustical Society Journal*, 3(1):26–33, 1996.

[9] John W Coltman. Jet offset, harmonic content, and warble in the flute.

The Journal of the Acoustical Society of America, 120(4):2312–2319, 2006.

[10] Claude (arranged by Bernard Dewagtere) Debussy. Syrinx. `http://www.free-scores.com`. Accessed: August 2014.

[11] Robert Dick. *Tone development through extended techniques*. Multiple Breath Music Company, 1986.

[12] Paul A Dickens et al. *Flute acoustics: measurements, modelling and design*. PhD thesis, PhD Thesis, University of New South Wales, 2007.

[13] Frederick Alton Everest, Ken C Pohlmann, and Tab Books. *The master handbook of acoustics*, volume 4. McGraw-Hill New York, 2001.

[14] Neville Horner Fletcher. Acoustical correlates of flute performance technique. *The Journal of the Acoustical Society of America*, 57(1):233–237, 1975.

[15] Neville Horner Fletcher. Silver, gold or platinum? *The Flute*, 1996.

[16] Neville Horner Fletcher. Vibrato in music. *Acoustics Australia*, 29(3):97–102, 2001.

[17] Neville Horner Fletcher and Thomas D Rossing. *The physics of musical instruments*. Springer, 1998.

[18] Angeleita Stevens Floyd. *The Gilbert Legacy: Methods, Exercises, and Techniques for the Flutist*. Winzer Press Cedar Falls, Iowa, 1990.

[19] James Galway. Sir James Galway 16 flute demonstration (youtube). `https://www.youtube.com/watch?v=GOn3n3N3SOY`. Uploaded: June 19 2007.

[20] Jochen Gartner. The vibrato. *With Particular Consideration given to the Situation of the Flutist, Bosse, Regensburg*, pages 79–83, 1981.

[21] John Michael Hajda. The effect of time-variant acoustical properties on orchestral instrument timbres. 1999.

[22] Donald E Hall. *Musical acoustics: An introduction.* Wadsworth Publishing Company, 1980.

[23] Karla Harby. Unsound reasoning. *Scientific American*, 278:20–21, 1998.

[24] Howard Harrison and Neil. White. *How to play the flute.* St. Martin's Press, 1983.

[25] David Miles Huber and Robert E Runstein. *Modern recording techniques.* CRC Press, 2013.

[26] Daniel A Jones. Uses of spectrographic analyses to improve tone quality of middle school trumpet students. *Journal of Music, Technology & Education*, 5(3):241–256, 2013.

[27] Larry Krantz. Flute list pages. `http://www.flutelist.com`. Accessed: September 2014.

[28] John C Krell. *Kincaidiana: a flute player's notebook.* National Flute Association, 1997.

[29] Siegmund Levarie and Ernst Levy. *Tone: A study in musical acoustics.* Kent State University Press, 1980.

[30] Weicong Li, Jer-Ming Chen, John Smith, and Joe Wolfe. Vocal tract effects on the timbre of the saxophone. *SMAC13*, 2013.

[31] Gareth Loy. *Musimathics: the mathematical foundations of music*, volume 2. MIT Press, 2011.

[32] Richard G. Lyons and D. Lee Fugal. *The Essential Guide to Digital Signal Processing*. Pearson Education, 2014.

[33] Philip McLeod. Fast, accurate pitch detection tools for music analysis. *Academisch proefschrift, University of Otago. Department of Computer Science*, 2009.

[34] Brian CJ Moore. *An introduction to the psychology of hearing*. Brill, 2012.

[35] Karen Evans Moratz. *Flute for Dummies*. John Wiley & Sons, 2009.

[36] Thomas Nyfenger. Music and the flute. 1986.

[37] Bobby Owsinski. *The recording engineer's handbook*. Hal Leonard Corporation, 2005.

[38] Barry Parker. *Good vibrations: the physics of music*. JHU Press, 2010.

[39] Phillip David Payne. *An investigation of relationships between timbre preference, personality traits, gender, and music instrument selection of public school band students*. The University of Oklahoma, 2009.

[40] Edwin Putnik. *The art of flute playing*. Alfred Music Publishing, 1973.

[41] Casey Reas and Ben Fry. Processing 2 programming language. `http://www.processing.org`. Accessed: October 2014.

[42] Casey Reas and Ben Fry. *Getting Started with Processing*. " O'Reilly Media, Inc.", 2010.

[43] Curtis Roads. *The computer music tutorial*. MIT Press, 1996.

[44] Katharin Elaine Rundus. *Imagination in harmony with science: Spectral analysis as a practical pedagogic tool in the voice studio*. 2002.

[45] Stanley Sadie and Alison Latham. *The Norton/Grove concise encyclopedia of music*. WW Norton New York, 1994.

[46] Carl Emil Seashore. *Psychology of music*. Courier Dover Publications, 1967.

[47] William A Sethares. *Tuning, timbre, spectrum, scale*, volume 2. Springer, 2005.

[48] SoGoSurvey. Online survey tool. `http://www.sogosurvey.com`. Accessed: September 2014.

[49] Roger S Stevens and Ruth N Zwissler. *Artistic Flute: Technique and Study*. Highland Music Company, 1967.

[50] Timothy Glen Stewart. *The relationship between personality type and timbre preference*. UMI Dissertation Information Service, 1993.

[51] Nancy Toff. *The flute book: a complete guide for students and performers*. Oxford University Press, 2012.

[52] Jan Vantomme. *Processing 2: Creative Programming Cookbook*. Packt Publishing Ltd, 2012.

[53] Gregor Widholm, Renate Linortner, Wilfried Kausel, and Matthias Bertsch. Silver, gold, platinum-and the sound of the flute. In *Proc. Int. Symposium on Musical Acoustics, Perugia*, volume 1, pages p277–280, 2001.

[54] Gerda Reinette Wilcocks. *Improving tone production on the flute with regards to embouchure, lip flexibility, vibrato and tone colour, as seen from a classical music perspective.* PhD thesis, University of Pretoria, 2006.

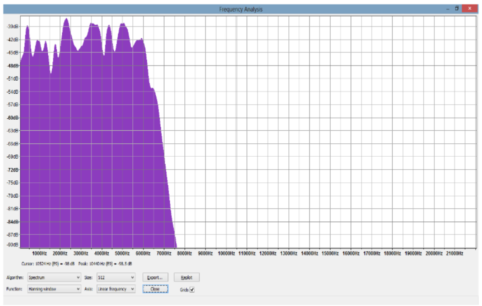
[55] Joe Wolfe, John Smith, John Tann, and Neville H Fletcher. Acoustic impedance spectra of classical and modern flutes. *Journal of Sound and Vibration*, 243(1):127–144, 2001.

APPENDIX A

Recording Equipment

All recording were captured using a MacBook Air (mid-2012 model) with an Audio-Technica AT2020 USB microphone and the Audacity 2.0.3 digital audio editor [1].

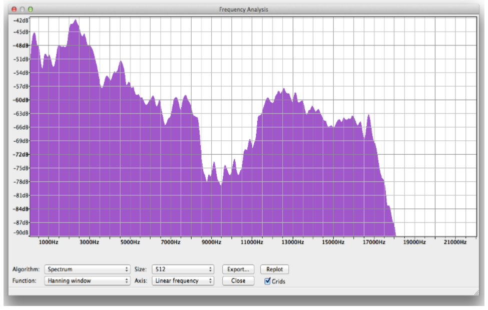
Microphone tests were conducted using an ASUS S300CA-BBI5T01 laptop (circa 2013) computer and a MacBook Air. White noise was played through a pair of Bose Companion 2 computer speakers. The white noise was recorded on each computer using three different microphones: the internal built-in microphone, an Audio-Technica AT2020 USB Condenser microphone, and a Zoom H2 recorder (used as a microphone). The speakers are not recording studio-quality monitors, so they introduce some imperfections. However, this is not an issue. The area interest is how each various microphone responds to the audio signals coming from the speakers.



Cursor: 10524 Hz (E9) = -98 dB Peak: 10440 Hz (E9) = -98.5 dB

Algorithm: Spectrum
Size: 512
Function: Hanning window
Axis: Linear frequency
Grids ☑

Figure A.1: ASUS built-in

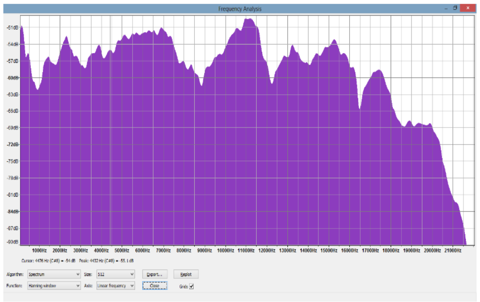


Algorithm: Spectrum
Size: 512
Function: Hanning window
Axis: Linear frequency
☑ Grids

Figure A.2: MacBook built-in

Surprisingly, the built-in microphone for the ASUS drops off after 6K Hz. A similar test using an HP Envy 6-1010US laptop (circa 2012) produced similar results. The built-in microphones on both PCs would not capture higher frequencies and could potentially lose upper harmonics. For this study, the highest pitch that was recorded is B5 which is approximately 988 Hz. H6 for B5 is 5,928 Hz and would likely be captured. However, H7 is 6,916 Hz and is beyond the range of the built-in PC microphones. Since the flute range extends up to C7, the built-in microphone would not be appropriate for the third octave of the flute.

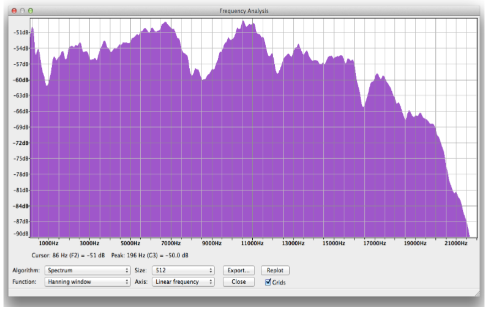
The MacBook internal microphone performed better, but there is a dip around 8K Hz. It clearly outperforms the PC around 11-12K Hz. The built-in MacBook microphone would suffice for this study. However, the uneven sensitivity of the microphone is of some concern as it could distort levels of some harmonics. 12K Hz is the range of the 7th harmonic for A6, so the built-in microphone on the MacBook would lose information for the upper end of the flute's third octave.



Cursor: 4476 Hz (C#8) = -54 dB Peak: 4432 Hz (C#8) = -55.1 dB

Algorithm: Spectrum
Size: 512
Function: Hanning window
Axis: Linear frequency
Grids ☑

Figure A.3: ASUS with Audio-technica

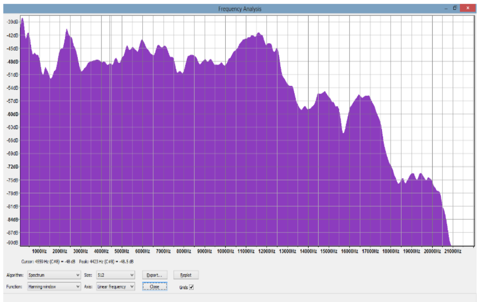


Cursor: 86 Hz (F2) = -51 dB Peak: 196 Hz (G3) = -50.0 dB

Algorithm: Spectrum
Size: 512
Function: Hanning window
Axis: Linear frequency
Grids ☑

Figure A.4: Macbook with Audio-technica

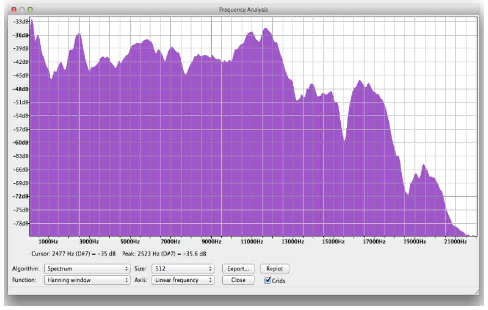
The Audio-Technica microphone performs better on both computers and drops off slowly after 16.5K Hz. The H2 zoom record starts to drop off after 12.5K Hz for both computers. Either microphone is sufficient for this study. Overall, the Audio-Technical performed better and was used for all the flute tone recording sessions.



Cursor: 4559 Hz (C#8) = -48 dB Peak: 4423 Hz (C#8) = -48.5 dB

Algorithm: Spectrum
Size: 512
Function: Hanning window
Axis: Linear frequency
Grids ☑

Figure A.5: ASUS with H2 zoom



Cursor: 2477 Hz (D#7) = -35 dB Peak: 2523 Hz (D#7) = -35.6 dB

Algorithm: Spectrum
Size: 512
Function: Hanning window
Axis: Linear frequency
☑ Grids

Figure A.6: MacBook with H2 zoom

APPENDIX B

Recording Details

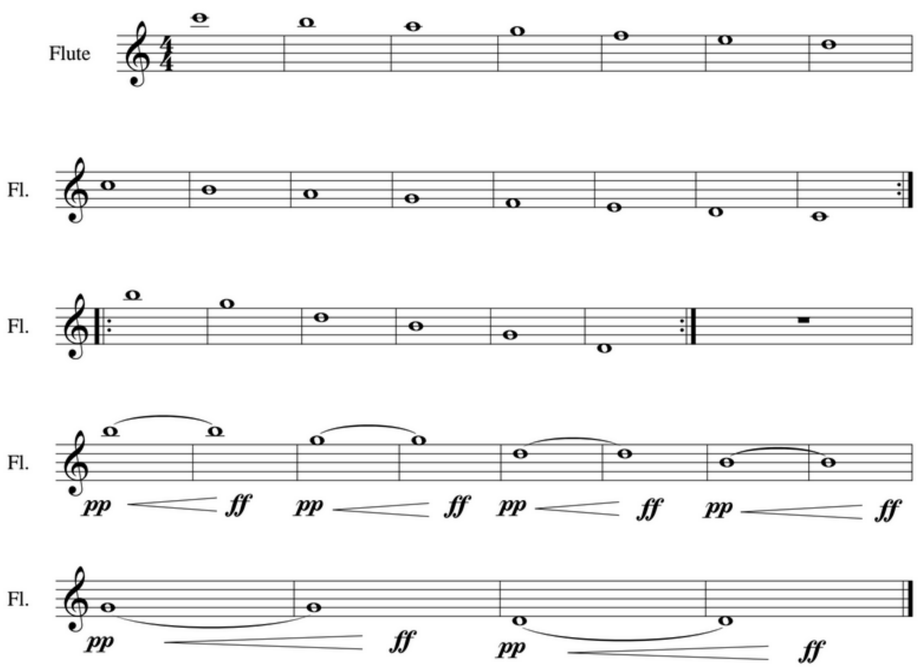


Figure B.1: Sheet music for recording

These are the steps for each recording session:

1. All subjects that were recorded were asked to read and sign an informed consent form. The text in that form is at the end of this appendix.
2. Subjects were asked to play a few notes to warm up. This allowed some time for testing the recording equipment setup.
3. Subjects were asked to play at a *mf* dynamic level using a tone that they considered appropriate for a passage marked *cantabile* or *dolce*. To re-

duce any anxiety, the subjects were reminded that their recording would be anonymous. Further, they were told if they cracked a note to play the note again to their satisfaction.

4. They were then instructed to play the notes from lines 1-2 of the sheet music. Each whole note should be 4-5 seconds long. The first time, all notes would be played with straight-tone, and then repeated with vibrato.
5. Next, they were asked to play line 3, first with straight-tone and then with vibrato. Subjects were asked to play the last two lines on the sheet music using dynamics; starting *pp* and ending *ff*. Each note should have a duration of approximately 8 seconds. They were instructed to play the last two lines first with straight-tone, and then to repeat and play with vibrato.
6. Finally, subjects were handed a "control-flute" (a Miyazawa PA-202 with a Dan Sheridan Headjoint). They were allowed to play a few minutes to acclimate to the new instrument. Then they were asked to play line 3, first with straight-tone and then with vibrato.

Most of the recording sessions followed this format. In some situations, the musicians did not play all of the notes, either because of their inability to produce a note, or a note was accidentally omitted. Some less experienced flutists were not comfortable using vibrato and only played the straight-tone notes. Four of the university students were recorded early in the project and played fewer notes. The high school students were only asked to play notes from an inverted G major chord using straight-tone.

INFORMED CONSENT TO PARTICIPATE IN A FLUTE TONE QUALITY STUDY

A research project for the analyzing and visualizing flute tone is being conducted by Ron Yorita, in the Department of Computer Science at Cal Poly, San Luis Obispo. The purpose of the study is to gain insights into flute tone and to build a computer program that will provide quantitative data about tone quality.

You will be asked to play some notes on your instrument for recording purposes. Your participation will take approximately 10 minutes. Please be aware that you are not required to participate in this research and you may discontinue your participation at any time without penalty. There are no risks associated with participation in this study.

Your confidentiality will be protected your identity will not be divulged without your expressed permission. The recordings will be used for analysis of tone and may be used to demonstrate particular tone characteristics. By participating in the study, you are agreeing to allow the researcher to include your recordings anonymously for others to hear. Potential benefits associated for the study include improvements to an application program design as well as input for a computer science masters thesis project.

If you have questions regarding this study or would like to be informed of the results when the study is completed, please feel free to contact Ron Yorita at ron.yorita@gmail.com and/or the professor advising this project, Dr. Clements at clements@calpoly.edu. Thank you for your participation in this research.

Volunteer signature ____________________ Date__________

APPENDIX C

Survey Questions

These are the Flute List survey questions. Other surveys administered to the Cal Poly students contained fewer demographic questions. All of the surveys administered contained the informed consent and six comparison pairs with audio clips.

The following paragraph appeared on top of every survey screen:

Listen to each recording by clicking on the audio icon (located above and to the right of the question). Then describe each tone using the text boxes, and rate the tone quality by using the radio buttons next to each text box. This is subjective, and there are no correct or incorrect answers, so feel free to express your opinions.

Question 1: I have read the informed consent and I agree

> A research project on analyzing and visualizing flute tone is being conducted by Ron Yorita, in the Department of Computer Science at California Polytechnic State University, San Luis Obispo. The purpose of the study is to gain insights into flute tone and to build computer programs that will provide quantitative data about tone quality.
>
> You will listen to flute tones and give comments and opinions on quality. Your participation will take approximately 10-15 minutes. Please be aware that you are not required to participate in this research and you may

discontinue your participation at any time without penalty.

Your confidentiality will be protected, and your identity will not be divulged. By taking part in this survey you are agreeing to allow your comments and opinions to be anonymously included as part of the thesis.

If you have questions regarding this study, please feel free to contact Ron Yorita ron.yorita@gmail.com or the professor advising this project, Dr. John Clements clements@calpoly.edu.

Thank you for your participation in this research.

Question 2: Gender (male, female)

Question 3: If you are a musician, indicate if you are a student, professional, or hobbyist

Question 4: If you are a musician, specify what instrument(s) you play (enter voice if vocalist) and approximately how long youve played

Question 5: If you have any favorite flutist please list one or two of them

Question 6: Indicate your top 2 favorite music genres (alternative, blues, classical, country, easy listening, electronic, folk, jazz, latin, new age, pop/rock, R&B, rap & hip-hop, soul, vocal, other)

Question 7: Are you a flutist? (if not, the survey skips to question 14)

Question 8: If you had private flute lessons, how many years did you receive private lessons?

Question 9: If you are a flute instructor, how many years have you been teaching?

Question 10: What type of flute do you play (professional model, intermediate model, student model)?

Question 11: What is the make and model of your flute?

Question 12: Approximately how many hours per week do you currently play, practice, and rehearse?

Question 13: How important is good tone to you? (0-not at all 5-highest priority)

Question 14 (pair #1): Describe the tone quality (use adjectives or short phrases) for the first note and for the second note

Question 15 (pair #1): Rate the tone quality (1-poor 5-great) for the first note and for the second note

Question 16 - 25: Similar questions for pair #2 through pair #6

Question 26: Please enter any comments or suggestions

APPENDIX D

Demographics

D.1 Full Flute List

The Flute List survey had 121 participants. Nearly 63% (73 out of 121) of the participants are flute teachers. There is a fairly even distribution of years of teaching experience as shown in figure D.1. The aggregated teaching experience for these 73 teachers is 1,651 years. This is an average of 22.6 years per teacher. The amount of time of playing, practicing, and rehearsing for the 121 participants is 15.3 hours per week.

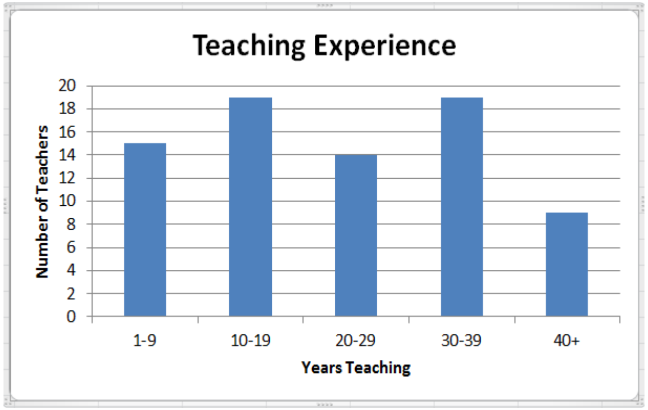


Figure D.1: Flute List - teaching experience

The survey participants were asked the importance of "good" tone quality. 66% indicated that it is the highest priority (5 on a scale of 0-5). 30% rated this as 4. This indicates 96% of the Flute List members rated tone quality a high priority.

- 66% female vs. 44% male
- 84% play professional model flutes, and 12% play intermediate model flutes
- The four most frequently mentioned flute brands were: Powell (19), Haynes (17), Brannen (15), and Muramatsu (15)

D.2 FL10s

Most of the demographics for the FL10s are in section 7.2.1. This list contains a few supplemental items:

- 76% female vs. 24% male
- 100% play professional model flutes
- The four most frequently mentioned flute brands were: Powell (8), Haynes (6), Muramatsu (6), and Brannen (5)

D.3 Favorite Flutists

The list of favorite flutists compiled during this study is shown on the next page.

Count	Lastname	Firstname
37	Pahud	Emmanuel
22	Rampal	Jean-Pierre
18	Galway	James
11	Baker	Julius
11	Bennett	William
10	Bouriakov	Denis
6	Piccinnini	Marina
5	Beynon	Emily
5	McGhee	Lorna
5	Robison	Paula
5	Still	Alexa
4	Jennings	Christina
4	Platillo	Greg
4	Stallman	Robert
4	Wincenc	Carol
4	Zoon	Jacques
3	Anderson	Ian
3	Baxtresser	Jeanne
3	Debost	Michel
3	Dick	Robert
3	Dufour	Mathieu
3	Gedigian	Marianne
3	Walker	Jim
2	Clark	Ian
2	Cox	Michael
2	DuFour	Matthew
2	Gaubert	Philippe
2	Graf	Peter Lucas
2	Moyse	Marcel
2	Porter	Amy
2	Ryerson	Ali
1	Aitken	Robert
1	Baeten	Aldo
1	Barth	Molly
1	Beaudiment	Julien
1	Boyd	Bonita
1	Boyd	Bonita
1	Brown	Rachael
1	Buyse	Louise
1	Caroli	Mario
1	Chapuis	Isabelle
1	Choi	Jasmine
1	Dolphy	Eric
1	Felber	Jill
1	Ferrandis	Jean
1	Formisano	Davide
1	Gallois	Patrick
1	Gleghorn	Authur
1	Goldberg	Bernard
1	Greenbaum	Adrianne
1	Guzman	Viviana
1	Harris	Tracy
1	Hoepner	Susan
1	In Sterio	
1	Kellerman	Wouter
1	Khaner	Jeffrey
1	Kincaid	William
1	Kirk	Roland
1	Koefler	Michael
1	La Berge	Anne
1	Langevin	Robert
1	Larieau	Maxence
1	Larson	Rhonda
1	Laws	Hubert
1	Litz	Han
1	Mann	Herbie
1	Marion	Alain
1	McBirnie	Bill
1	Monroe	Ervin
1	Morris	Gareth
1	Newton	James
1	Nicolet	Aurele
1	Norman	Chris
1	Nyfenger	Thomas
1	Panitz	Murray
1	Praful	
1	Rangell	Nelson
1	Rees	Carla
1	Robertello	Tom
1	Schneemann	Marieke
1	Smith	Joshua
1	Sparks	Mark
1	Stone	Dorothy
1	Thaves	Darrin
1	Toote	Linda
1	Torres	Nestor
1	Valle	Orlando "Maraca"
1	Weisberg	Tim
1	Wilson	Ransom
1	Zuckerman	Eugenia

APPENDIX E

Spectra

E.1 Flute Spectra - Repeated Notes

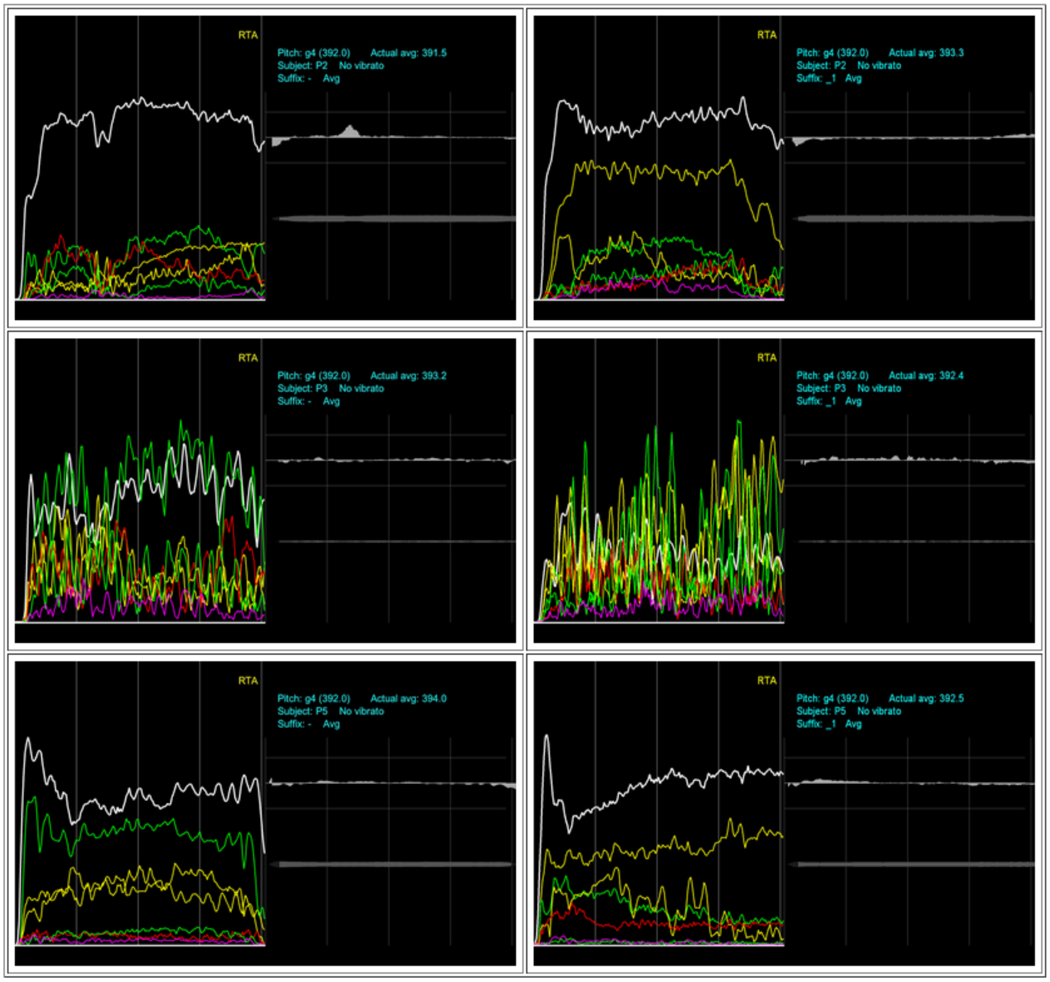


Figure E.1: G4 repeated note (part 1 of 2)

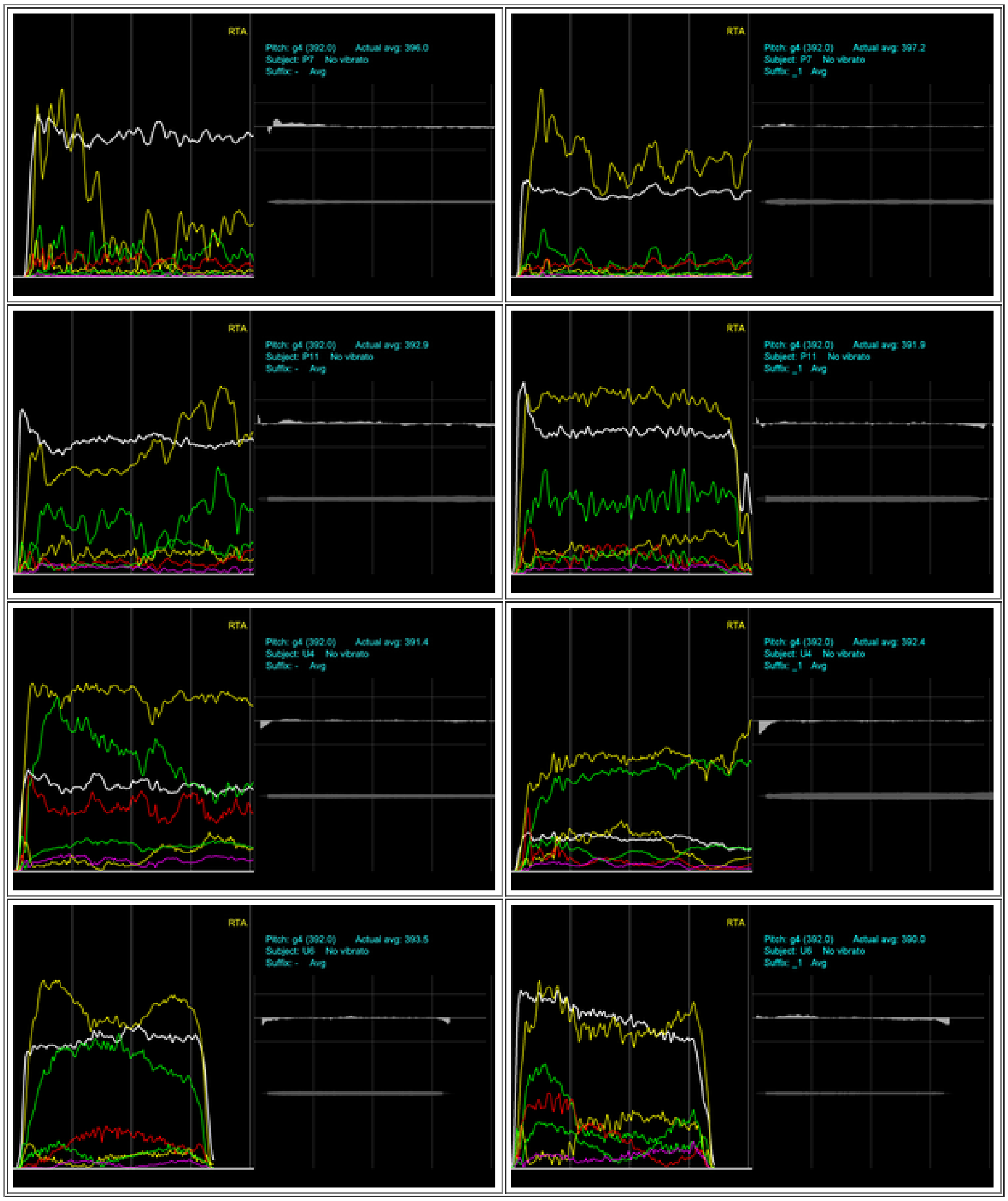
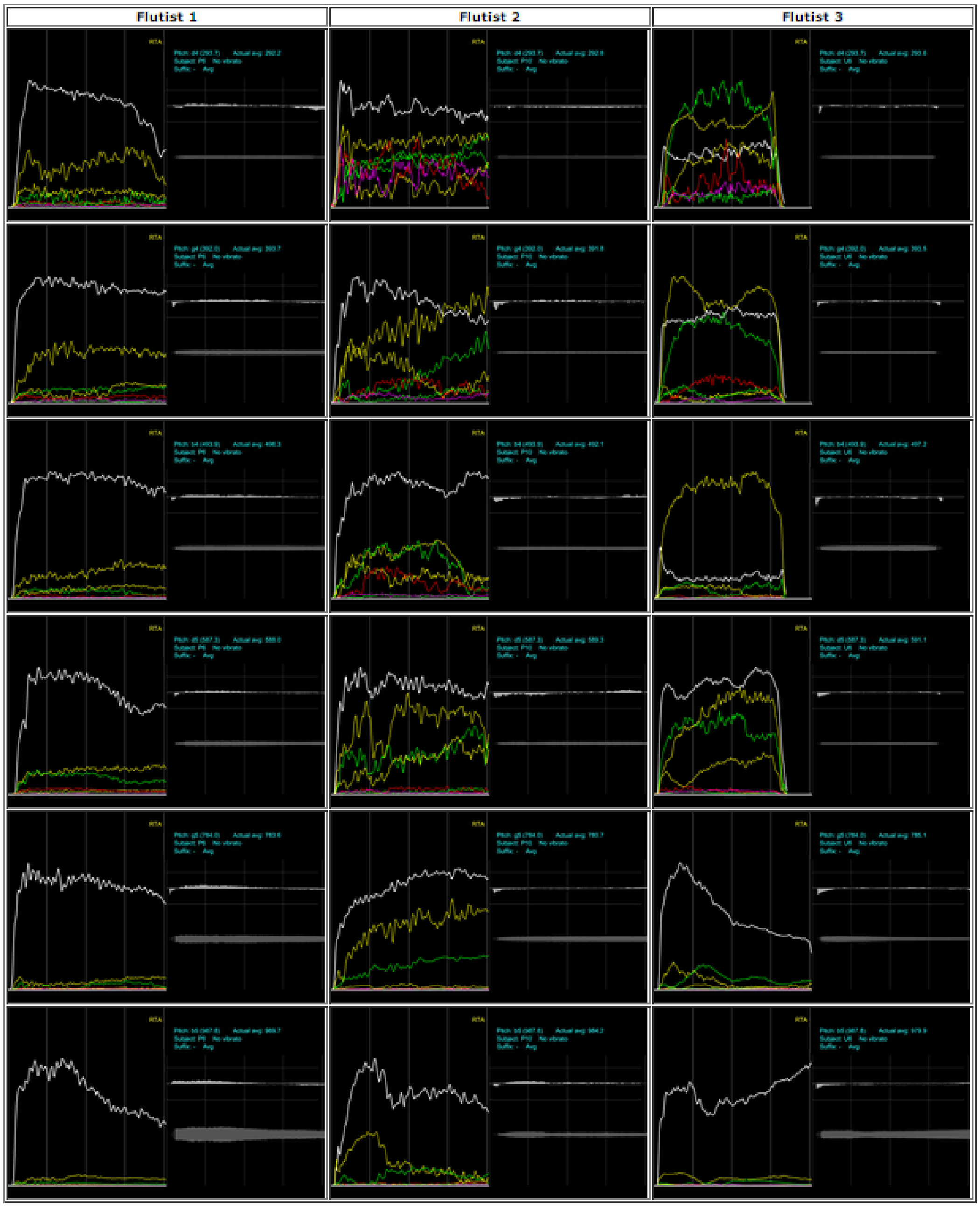


Figure E.2: G4 repeated note (part 2 of 2)

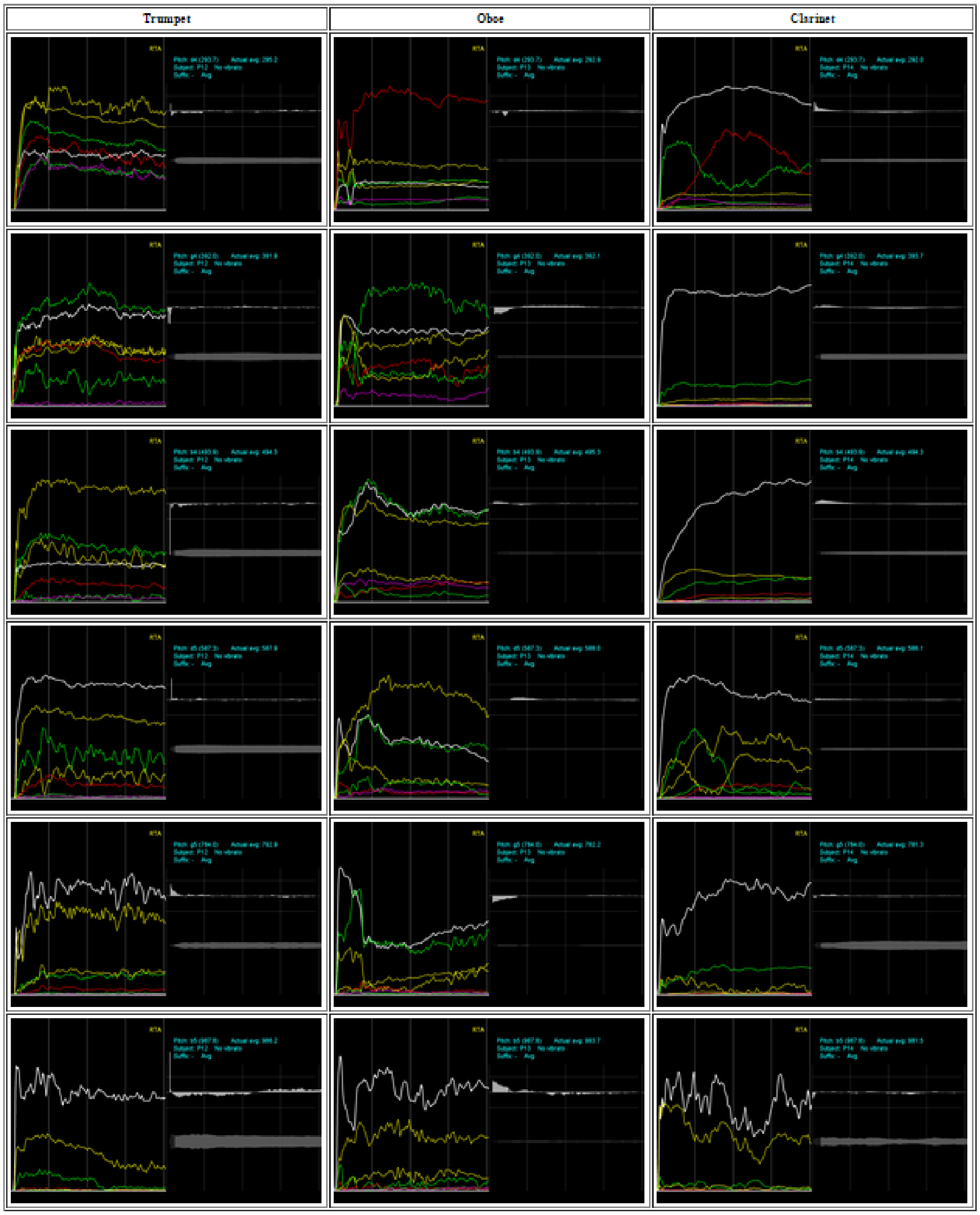
E.2 Flute Spectra - Range Of Pitches

Each column is a particular flutist. Each row is a different pitch (D4, G4, B4, D5, G5, B5).



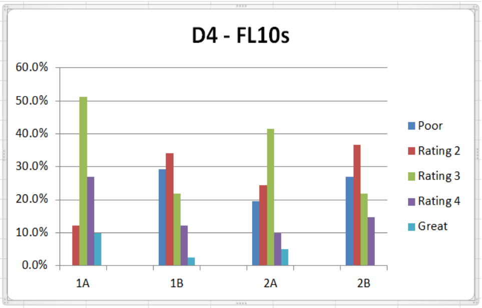
E.3 Oboe, Trumpet, and Clarinet Spectra

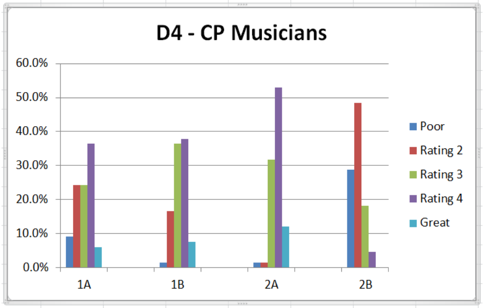
Each column is a particular flutist. Each row is a different pitch (D4, G4, B4, D5, G5, B5).

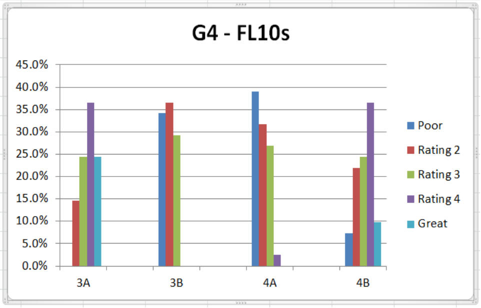


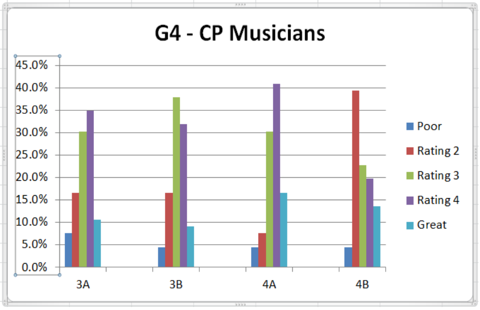
APPENDIX F

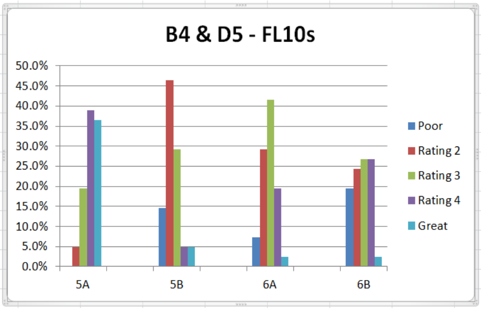
Rating Distributions

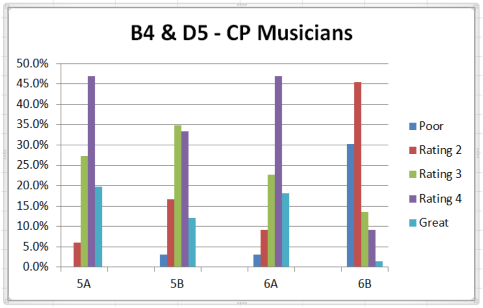












APPENDIX G

Flute List Descriptor Correlation

Using the descriptors from the full Flute List survey, terms were extracted and correlated. The following process was used:

- manually pre-process the survey text
- identify descriptors that are used five or more times for a given sample
- programmatically calculate correlations

Correlation values range from -1 to +1. Values close to +1 have a high correlation, and those with values close to -1 have low correlation. The next page contains a table of descriptors with high correlation. The subsequent page contains descripts with low correlation.

Descriptor 1	**Descriptor 2**	**Correlation**
focused	clear	0.9658
breathy	fuzzy	0.9451
focused	strong	0.9451
supported	resonant	0.9418
sharp	overblown	0.9393
edgy	hard	0.9349
overblown	loud	0.9339
metallic	hard	0.9300
unsupported	weak	0.9246
airy	breathy	0.9173
clear	strong	0.9169
dull	muffled	0.9165
trumpet	metallic	0.9138
unfocused	breathy	0.9109
nasal	trumpet	0.9022
airy	fuzzy	0.9007
resonant	nice	0.8892
forced	overblown	0.8884
fuzzy	diffuse	0.8787
unfocused	diffuse	0.8766
trumpet	pinched	0.8765
sharp	loud	0.8750
hollow	mellow	0.8726
full	rich	0.8714
unfocused	fuzzy	0.8680
brassy	hard	0.8671
airy	unfocused	0.8630
brassy	pinched	0.8587
brassy	metallic	0.8525
forced	loud	0.8519
airy	diffuse	0.8509
nasal	buzzy	0.8488
focused	supported	0.8475
trumpet	hard	0.8452
nasal	metallic	0.8429
clear	full	0.8359
hollow	uncentered	0.8336
forced	sharp	0.8294
open	soft	0.8224
metallic	pinched	0.8212

Descriptor 1	**Descriptor 2**	**Correlation**
unfocused	edgy	-0.8125
edgy	dull	-0.8070
edgy	soft	-0.8056
edgy	weak	-0.7871
airy	edgy	-0.7813
strong	diffuse	-0.7809
edgy	unsupported	-0.7651
weak	strong	-0.7605
edgy	diffuse	-0.7602
edgy	breathy	-0.7522
weak	bright	-0.7458
unfocused	strong	-0.7429
bright	diffuse	-0.7347
edgy	open	-0.7343
bright	soft	-0.7258
airy	strong	-0.7244
focused	weak	-0.7227
bright	dull	-0.7223
unfocused	hard	-0.7211
edgy	mellow	-0.7194
focused	unfocused	-0.7185
soft	hard	-0.7131
airy	bright	-0.7087
clear	diffuse	-0.7082
focused	diffuse	-0.7076
harsh	soft	-0.7065
edgy	hollow	-0.6994
unsupported	strong	-0.6980
breathy	strong	-0.6885
open	hard	-0.6883
airy	focused	-0.6781
unfocused	rich	-0.6775
strong	dull	-0.6769
unfocused	bright	-0.6765
edgy	fuzzy	-0.6752
forced	round	-0.6712
breathy	hard	-0.6697
unsupported	bright	-0.6691
breathy	bright	-0.6684
hollow	strong	-0.6680