Reverse engineering the sound of the jazz saxophone

Luca Cazzanti, Vanessa Sielert, and Maya R. Gupta

Dept of EE, University of Washington
Seattle WA, 98195-2500

UW Electrical Engineering

UWEE Technical Report
Number UWEETR-2006-0011
June 2006

Department of Electrical Engineering
University of Washington
Box 352500
Seattle, Washington 98195-2500
PHN: (206) 543-2150
FAX: (206) 543-3842
URL: http://www.ee.washington.edu

Reverse engineering the sound of the jazz saxophone

Luca Cazzanti, Vanessa Sielert, and Maya R. Gupta

Dept of EE, University of Washington
Seattle WA, 98195-2500

University of Washington, Dept. of EE, UWEETR-2006-0011
June 2006

Abstract

Differences between jazz sounds and symphonic sounds are investigated via the saxophone long tone spectra. A hypothesis is presented that relates these different styles of sound to the pitch of the saxophone mouthpiece when played in isolation. Specifically, playing a low pitch on the mouthpiece is hypothesized to create a sound more typical of jazz, while playing a high pitch on the mouthpiece is hypothesized to create a sound more typical of symphonic music. Experimental data and the nearest-neighbor method using the Itakura-Saito distortion provide statistically significant evidence for the proposed hypothesis. The interaction between different mouthpieces and the mouthpiece pitch is also studied. This work has implications for scientists, musicians, and engineers.

1 Background

This article details an experimental comparison of symphonic and jazz saxophone sounds, and the production of these different styles of sound. After presenting some comparisons of the spectral differences, a constructive hypothesis is presented as to how the different sounds can be produced. An experiment to test this hypothesis is detailed, and the resulting statistically significant evidence is given, supporting the proposed hypothesis. This work has implications for understanding musical sound, synthesizing music, designing musical equipment, compressing music, recording technologies, composing, and for saxophone pedagogy.

Musicians generally agree that jazz wind instrument sounds, such as the saxophone sound, have more higher harmonic content than symphonic wind instrument sounds. In Section 2 some example spectra are given, and harmonic content is discussed. A saxophone is composed of two sections, a mouthpiece and the main bore. A musician can play a saxophone mouthpiece by itself, and a tone called the *mouthpiece pitch* is produced. On an alto saxophone mouthpiece, a musician can reach an entire octave of mouthpiece pitches. The chosen mouthpiece pitch can make the fingered note flatter or sharper, and the mouthpiece pitch also affects the overall quality of the saxophone tone. In Section 4 we discuss the effect of the mouthpiece pitch on the full saxophone tone spectra. Several saxophone pedagogues contend that a mouthpiece pitch of $A5$ produces a tone typical of symphonic playing. Some pedagogues hypothesize that a lower mouthpiece pitch will produce a sound more typical of jazz. Experimental evidence for this is given in Section 5.

2 Jazz and symphonic spectra

Symphonic tones are described by musicians as “focused, consistent, round, and warm,” while jazz tones are often described as “brasher, edgier, more colorful, or having more buzz” (1). These jazz descriptions are generally thought to correspond to stronger high-frequency energy than heard in symphonic tones. To confirm this, long tones were recorded from three professional jazz and three highly-qualified symphonic saxophone players. The magnitude of three of the six jazz reference spectra, and three of the seven symphonic reference spectra are displayed in Figure 1. Details of the recording and processing are given next, and then the spectral differences are discussed.

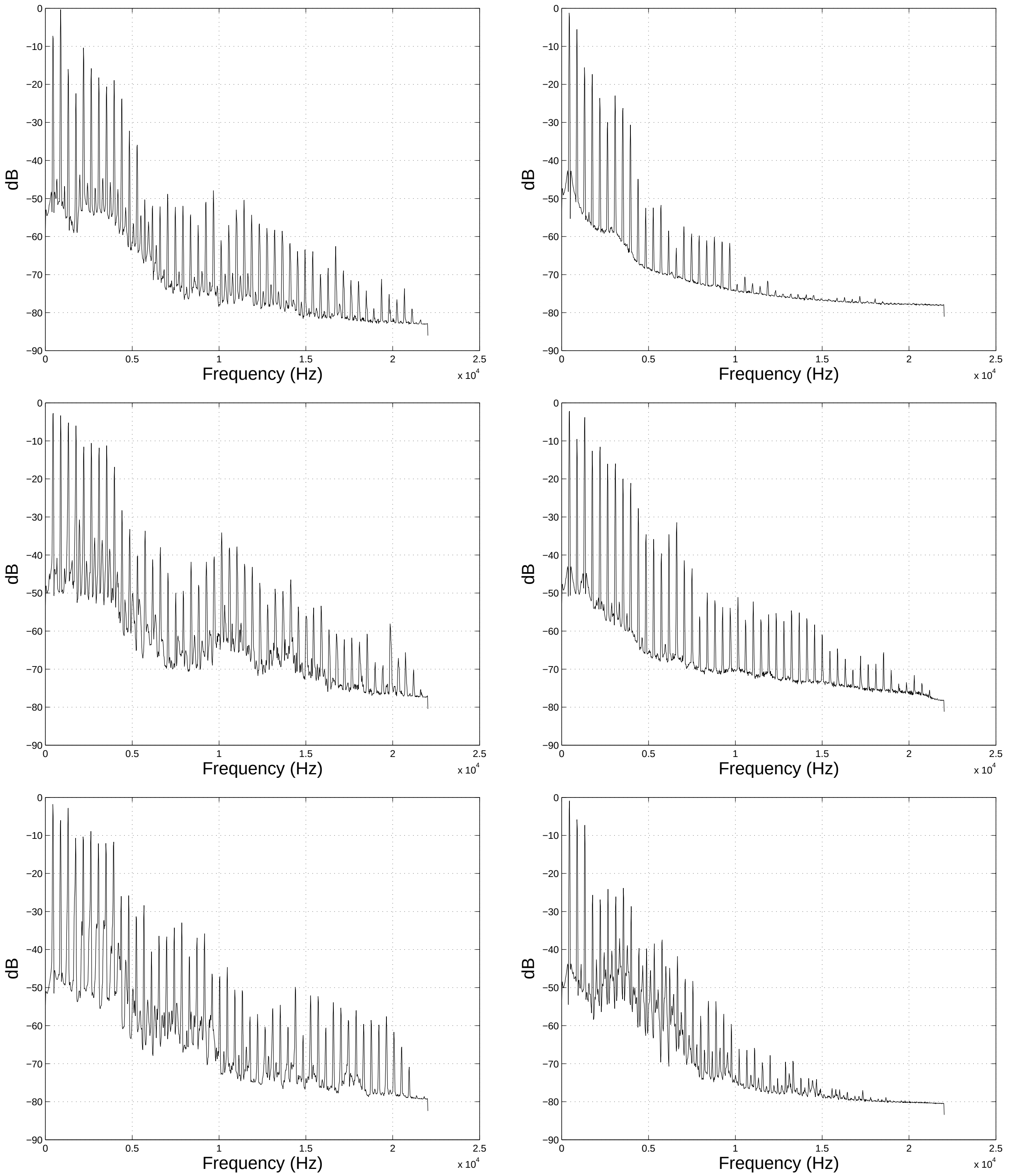

Figure 1: Three jazz reference spectra (left) and three symphonic reference spectra (right).

2.1 Recording details

Reference musicians used their choice of their personal equipment (reed, mouthpiece, and saxophone). Each musician was recorded playing an $A4$ (440 Hz) for a duration of three to five seconds, twice, with a brief pause between the tones. The attack and decay of each tone was cut from the signal before processing, as the focus of this study is the steady-state sound.

Reference tones were recorded on the same day, in the same room, and in the same manner as all the data collected in this study. Subjects were recorded on the stage of the Tryon Festival Theatre in the Krannert Center for the Performing Arts on the campus of the University of Illinois at Urbana-Champaign. The Tryon theatre is roughly rectilinear in shape, with a distance of 77 ft. from the proscenium to the back wall of the theatre, 32 ft. from the floor to ceiling, and 70 ft. between the side walls of the theatre.

The geometrical and acoustic set-up was the same for each musician in this study. The musician stood on the stage pointing forward. To minimize the influence of the room, a large Duvatine drape was eight to ten feet behind the musician. The closest wall behind the drape was forty feet away. The next closest wall was sixty feet away. The microphone was two to three feet in front of the musician. All musicians pointed the saxophone bell at the microphone for all the tones they recorded. In this way, it was assured that most of the recorded energy was direct. However, some energy may have been reflected by the floor. The floor was a sprung performance floor with painted masonite as the top layer. The influence of the floor on the spectra is a function of the musician's height (or more specifically, of the height at which the musician held the instrument). The heights of the set of classical reference musicians and the set of jazz reference musicians each varied, without extremes. For the experiment detailed in Section 5, the same musicians played the saxophone with both the $A5$ and E^b5 embouchures, and thus any differences due to height from the floor would have been controlled in the experiment.

The recording was done with a Neumann U87 microphone and a True Systems Precision 8 microphone amplifier. All subjects were recorded to digital audiotape (DAT) at a sampling frequency of 44.1 kHz. Spectra were computed up to 22.05 kHz, in accordance with Shannon's sampling theorem. All of the processing was done with the Fast Fourier Transform, with Fourier coefficients every 21.5 Hz.

2.2 Spectral differences

Visually, the reference spectra in Figure 1 confirm the widely-held belief that there is more high frequency content in jazz long tones than in symphonic long tones. For all seven of the symphonic reference spectra (three of which are shown), the harmonics fall off roughly linearly in dB, while for the six jazz reference spectra, the harmonics below 5 kHz are much stronger, then fall off somewhat abruptly around 5 kHz. Continuing out to the extent of human hearing, there is more energy in the higher jazz harmonics, and a greater amount of inharmonic energy as well. Further experiments would be needed to determine if the characteristic jazz sound was due predominantly to the greater energy in the harmonics, or to the greater energy in inharmonic higher frequencies, or to the combination thereof. In the next section, objective measures are discussed.

3 Data Analysis

A standard technique in science is to fit a model to experimental data. In this case, one might fit an acoustic model of the spectral decay, or of the spectral entropy. Such models have been investigated for the spectra of saxophone and clarinet (2; 3). However, for the analysis needed for this experiment, information that discriminates jazz and symphonic saxophone sounds is needed. To that end, the jazz and symphonic sounds of the saxophone are *directly represented* by examples (prototypes) of these sounds: the thirteen reference spectra. This allows us to focus on questions about how these different spectra are created, without concerns about the validity of a model for these sounds or the fit of an appropriate model.

Given a new, test spectrum of a 440 Hz long tone, we classify it as a jazz tone or symphonic tone based on the label of the reference spectrum it most closely resembles. This paradigm of classifying a test sample to be the same as its nearest-neighbor from a set of references is a common approach to pattern recognition and statistical learning (4). Using nearest-neighbor classification avoids the need to specify a model of jazz sounds, but does require an appropriate distance measure to compare test sounds to reference sounds.

Measuring distance exactly as humans hear differences between sounds is a difficult and possibly ill-defined problem. In this study the Itakura-Saito (I-S) distortion was used. I-S distortion has been shown to be a useful measure of audio distortion for other human-audio processing tasks (5; 6; 7). Its close relative, the Itakura distortion, has been shown successful in classifying musical signals (8) in conjunction with the nearest-neighbor paradigm. Given a reference spectrum with frequency magnitudes $r \in \mathcal{R}^{1024}$ and a test spectrum with frequency magnitudes $x \in \mathcal{R}^{1024}$, the I-S distortion is:

$$\text{I-S}(x,r) = \sum_{i=1}^{1024} \left(\frac{x_i}{r_i} - \log\left(\frac{x_i}{r_i}\right) - 1 \right) \tag{1}$$

Like the decibel scale and cepstral coefficients, which have been shown useful in recognizing musical signals (9), the I-S distortion is based on the log of the frequencies. The I-S distortion is a measure of all the spectral information, and not only a function of spectral peaks. Because it may be that jazz sounds are differentiated in part by their inharmonic energy content, it was important to employ a measure that takes into account the entire spectra.

It was confirmed numerically that the I-S distortion preserves the following expected relationships between spectra: I-S distortions between spectra of one reference player are smallest; I-S distortions between jazz reference spectra show more variability than between symphonic reference spectra; and I-S distortions between jazz and symphonic spectra are relatively large compared to I-S distortions between two symphonic spectra or two jazz spectra. Psychoacoustics information, such as a measured modulation transfer function for the human auditory system or psychoacoustic masking, was not taken into account by using the I-S distortions. While using a psychoacoustic model might improve the distortion measure, it is not clear the experiments used to derive such psychoacoustic information are relevant for this application. Complete data of the I-S distortions between reference spectra and test spectra are available online (1).

4 Effect of mouthpiece pitch on sound spectra

A range of approximately one octave can be played on the alto saxophone mouthpiece alone, from $C5$ (523.3 Hz) to $C6$ (1047 Hz). Classical saxophone pedagogue Rousseau teaches that a characteristic alto saxophone tone quality will be produced if the player produces a tone of $A5$ (880 Hz) when playing the mouthpiece alone (10). There is no similar standard recommendation for producing a jazz tone. Instead, a wide variety of advice on how to achieve a jazz tone is available (11). Much of the advice is based in an oral tradition of "listening" to established jazz artists. Some jazz saxophone professionals hypothesize that a lower mouthpiece pitch than the $A5$ recommended for symphonic playing will produce a sound more typical of jazz. To test this hypothesis, test spectra were produced, using either an $A5$ or E^b5 mouthpiece pitch, and each was classified as per the closest reference symphonic or jazz spectrum. The E^b5 tone was chosen to be significantly lower than the $A5$, while still in a comfortable range on the mouthpiece. The choice was also based on the positive experiences of professional saxophonist Vanessa Hasbrook in playing and teaching jazz saxophone; in particular these are mouthpiece tones that most college-level saxophonists can stably produce. The null hypothesis is that the mouthpiece pitch has no correlation to the class label of the nearest spectra. The alternative hypothesis is that there is a positive correlation.

The mouthpiece pitch is a function of the player's embouchure (lip placement and lip tension) and oral cavity. Figure 2 shows the relative position of the tongue and lower lip to produce a mouthpiece pitch of $A5$ (left) and E^b5 (right). Mouthpiece pitch is partially a function of lower lip pressure on the reed, with the E^b5 less damped than the $A5$. However, lip pressure is not enough to reach the range of mouthpiece pitch from $A5$ to E^b5; the tongue position must also change. The tongue position does not directly affect the reed, but does change the shape and size of the cavity in which the reed vibrates. The role of the oral cavity in controlling musical notes is well-known to musicians (12; 13). Vibrato can also be produced through changes in the oral cavity (14).

5 Experiment

All ten subjects were saxophone students (music majors) at the undergraduate or graduate level at the University of Illinois at Urbana-Champaign. The range of prior jazz experience was from very little to professional jazz players. Experimental details are the same as in Section 2, except all subjects played the same saxophone, a Selmer Super Action Series II, serial # 580724. Each subject warmed up on the instrument using a Selmer C^* mouthpiece, and a new Vandoren reed. They were then asked to produce an $A5$ pitch on the mouthpiece alone. They were allowed as much time as needed to become comfortable reproducing the $A5$ pitch on the mouthpiece alone. Then they were asked to re-attach the mouthpiece to the saxophone and blow the $A5$ mouthpiece pitch with the mouthpiece while playing the note $A4$ (the pitch played on a saxophone is a function of the keys, and not a function of the mouthpiece pitch). The entire process was repeated with three different mouthpieces, a Selmer S90, a Meyer 6M, and a Claude Lakey 4^*4. At least two blows of three to five seconds were recorded for each mouthpiece. Then the entire process was repeated with a mouthpiece pitch of E^b5; each player played a E^b5 mouthpiece pitch but fingered and played the note $A4$ on the re-assembled saxophone. The end result was at least two recorded long tones per mouthpiece pitch per player.

The resulting frequency spectra were calculated as in Section 2. For each test spectrum, the closest reference symphonic or jazz spectrum was found, and the test spectrum was given the same label. A total of 12 reference spectra were used, 6 jazz and 6 symphonic. Of the 87 $A5$ mouthpiece pitch test spectra, 58 most closely matched a reference symphonic spectra, or 66.67%. Of the 101 E^b5 mouthpiece pitch test spectra, 80 most closely matched a reference jazz spectra, or 79.21%. For both cases, there is a statistically significant correlation between the mouthpiece pitch and reference spectra type; from a Neyman-Pearson test for statistical significance the p-values are 6×10^{-4} and 4×10^{-10} respectively. When the type of mouthpiece used is taken into account, the results are even more significant, as discussed in the next section.

The test spectra that did not fit the mouthpiece pitch hypothesis were not independent errors. In particular, all eight of Test Subject 5's sounds using the mouthpiece pitch $A5$ were closest to jazz spectra, not closest to symphonic spectra as proposed by the hypothesis under test. Test Subject 5 self-identifies as a jazz musician, and may not have been able to retain the $A5$ mouthpiece pitch when playing the mouthpiece connected to the horn. If that were true, then the data from Test Subject 5 was corrupt in that the $A5$ sounds were not actually $A5$ test sounds. Removing Test Subject 5's data from the experiment would strengthen the support of the hypothesis. This was not done, as it is only speculation that Test Subject 5's $A5$ data was corrupt.

6 Effect of mouthpiece

Of the four different mouthpieces used in the experiment, the Selmer 80 and Selmer 90 are sold as symphonic mouthpieces, and the Meyer 6M and Claude Lakey are sold as jazz mouthpieces. There are five main parameters to mouthpiece design: material (rubber, metal, etc.), mouthpiece chamber shape and size, the mouthpiece baffle (shape and size of the cavity next to the reed), tip opening (the vertical distance between tip of reed and mouthpiece), and the lay (the rate at which mouthpiece slopes back from the tip to touch the reed). The four mouthpieces used in this study were all made of the same material (rubber), and the differences between them is best characterized by the tip opening according to the professional saxophonist: the Selmers have a tip opening of .067", the Meyer has a tip opening of .075", and the Lakey has a tip opening of .080". It is commonly known among saxophonists that mouthpieces with larger tip

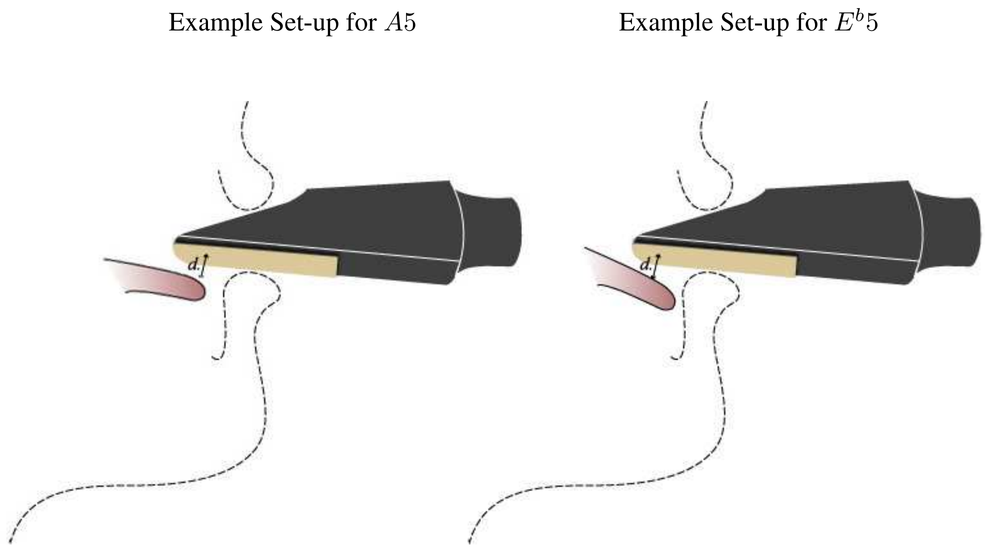

Figure 2: Tongue position rotates between playing an $A5$ mouthpiece pitch (left diagram) and an E^b5 mouthpiece pitch (right diagram). The tongue is the same distance d from the reed tip in both set-ups. The bottom lip pressure is slightly looser on the right than on the left.

openings provide more flexibility to bend pitch, but are harder for the musician to control. Smaller tip openings are easier to control, and are thus favored for symphonic playing.

The closest matches for the E^b5 and $A5$ tones are shown separated by mouthpiece in Table 1. Disregarding mouthpiece pitches, the symphonic mouthpieces S80 and S90 have 43 matches to reference symphonic spectra and 41 matches to reference jazz spectra. That is not statistically significant evidence for the hypothesis that the symphonic mouthpiece (alone) correlates to a symphonic tone; the p-value is $.3718$. However, a mouthpiece pitch of $A5$ played on a symphonic mouthpiece has a strong correlation to producing a symphonic spectra; the p-value is 2.2×10^{-5}. The Meyer jazz mouthpiece shows the strongest effect when combined with the E^b5 mouthpiece pitch to produce a jazz-like spectra; only two of those twenty-seven trials were closest to symphonic spectra. The p-value of that result is 2.1×10^{-7}, showing very significant statistical support that combining the Meyer jazz mouthpiece with the E^b5 mouthpiece pitch correlates with a sound closer to the jazz reference spectra. On the other hand, playing the Meyer jazz mouthpiece with the $A5$ mouthpiece pitch did not correlate strongly with either a jazz sound or a symphonic sound.

Table 1: Experimental results showing the correspondence of different mouthpiece pitches and mouthpieces to the type of the closest reference spectrum.

Mouthpiece Pitch	Mouthpiece	Trials	Symphonic Matches	Jazz Matches
$A5$	S80 (Symphonic)	20	16 (80%)	4 (20%)
$A5$	S90 (Symphonic)	20	16 (80%)	4 (20%)
$A5$	Meyer (Jazz)	21	10 (48%)	11 (52%)
$A5$	Lakey (Jazz)	26	16 (62%)	10 (38%)
E^b5	S80 (Symphonic)	22	5 (23%)	17 (77%)
E^b5	S90 (Symphonic)	22	6 (27%)	16 (73%)
E^b5	Meyer(Jazz)	27	2 (7%)	25 (93%)
E^b5	Lakey(Jazz)	30	8 (27%)	22 (73%)

The Lakey mouthpiece shows a lesser interaction with mouthpiece pitch. The relatively large tip opening on the Lakey mouthpiece does make it more difficult for a musician to control the mouthpiece pitch. Because of this, we hypothesize that there may have been significant deviation from the desired mouthpiece pitch when subjects used the Lakey, causing the less significant results.

7 Discussion

Based on the experience of professional musicians, a testable hypothesis was formed for how jazz and symphonic saxophone sounds could be created by a musician by using different mouthpiece pitches. Statistically significant evidence was strongly in favor of the presented hypothesis that lower mouthpiece pitch creates steady-state spectra that are closer to that of reference jazz steady-state spectra than to reference symphonic steady-state spectra. The mouthpiece pitch is a function of the oral cavity and embouchure pressure. This experimental study leaves open the acoustic explanation for the mechanisms creating the spectral differences. A speculative partial hypothesis is that the lower mouthpiece pitch generates more low frequencies that feed into the nonlinear processes associated with the reed and feedforward and feedback processes between the front oral cavity, the reed, and the bore resonances. Such greater low frequency energy feeding into the nonlinear interactions could create more diverse harmonic content, including the greater higher harmonic energy that characterized the difference between the jazz reference sounds and the symphonic reference sounds.

Acknowledgements

The authors would like to thank Debra Richtmeyer, Christopher Plummer, Michael Jacobson, and Robert Miller for helpful discussions.

References

[1] V. Hasbrook, “Alto saxophone mouthpiece pitch and its relation to jazz and classical tone qualities,” 2005, d.M.A. thesis at the University of Illinois Urbana-Champaign. Available online at: idl.ee.washington.edu/projects.php.

[2] A. H. Benade and S. J. Lutgen, “The saxophone spectrum,” *J. Acoust. Soc. Am*, vol. 83, no. 5, pp. 1900–1907, 1998.

[3] A. H. Benade and S. N. Kouzoupiz, “The clarinet spectrum: Theory and experiment,” *J. Acoust. Soc. Am*, vol. 83, pp. 292–304, 1998.

[4] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. New York: Springer-Verlag, 2001.

[5] A. Gray and J. Markel, “Distance measures for speech processing,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 24, pp. 380–391, 1976.

[6] R. Gray, A. Buzo, A. Gray, and Y. Matsuyama, “Distortion measures for speech processing,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 28, pp. 367–381, 1980.

[7] N. Nocerino, F. K. Soong, L. Rabiner, and D. Klatt, “Comparative study of several distortion measures for speech recognition,” *Speech Communication*, vol. 4, pp. 317–331, 1985.

[8] V. Venkatachalam, L. Cazzanti, N. Dhillon, and M. Wells, “Automatic identification of sound recordings,” *IEEE Sig. Proc. Magazine*, March 2004.

[9] J. Brown, “Computer identification of musical instruments using pattern recognition with cepstral coefficients as features,” *J. Acoust. Soc. Am*, vol. 105, no. 3, pp. 1933–1941, 1999.

[10] E. Rousseau, *Saxophone High Tones*. Saint Louis: MMB Music, Inc., 1978.

[11] D. Roach, *The Cambridge Companion to the Saxophone*. Melbourne: Cambridge University Press, 1998, ch. Jazz and Rock Techniques, pp. 788–93.

[12] D. Liebman, *Developing a Personal Saxophone Sound*. Medfield: Dorn Publications Inc., 1994.

[13] K. Horch, *The Cambridge Companion to the Saxophone*. Melbourne: Cambridge University Press, 1998, ch. The Mechanics of Playing the Saxophone, pp. 75–87.

[14] J. Gilbert, L. Simon, and J. Terroir, “Vibrato of saxophones,” *J. Acoust. Soc. Am*, vol. 118, no. 4, pp. 2649–2655, 2005.