Fixed Average Spectra of Orchestral Instrument Tones

JOSEPH PLAZAK
Ohio State University

DAVID HURON
Ohio State University

BENJAMIN WILLIAMS
Ohio State University

ABSTRACT: The fixed spectrum for an average orchestral instrument tone is presented based on spectral data from the Sandell Harmonic Archive (SHARC). This database contains non-time-variant spectral analyses for 1,338 recorded instrument tones from 23 Western instruments ranging from contrabassoon to piccolo. From these spectral analyses, a grand average was calculated, providing what might be considered an average non-time-variant harmonic spectrum. Each of these tones represents the average of all instruments in the SHARC database capable of producing that pitch. These latter tones better represent common spectral changes with respect to pitch register, and might be regarded as an "average instrument." Although several caveats apply, an average harmonic tone or instrument may prove useful in analytic and modeling studies. In addition, for perceptual experiments in which non-time-variant stimuli are needed, an average harmonic spectrum may prove to be more ecologically appropriate than common technical waveforms, such as sine tones or pulse trains. Synthesized average tones are available via the web.

Submitted 2010 January 22; accepted 2010 February 28.

KEYWORDS: *timbre, SHARC, controlled stimuli*

MOST modeling studies and experimental research in music perception involve the presentation or input of auditory stimuli. In many cases, researchers have aimed to employ highly controlled stimuli that allow other researchers to precisely replicate the procedure. For example, many experimental and modeling studies have employed standard technical waveforms, such as sine tones or pulse trains. In other cases, researchers have identified a particular spectral recipe, such as 10 equally-weighted harmonics (e.g. Plomp & Levelt, 1965; Huron & Sellmer, 1992). At the same time, researchers recognize the importance of using sounds that better approximate the sorts of sounds encountered in common listening situations. Many perceptual studies make use of commercial sound recordings, or experiment-specific recorded examples, played either on acoustical instruments or using MIDI devices. The choice of musically-pertinent stimuli is often regarded as a dichotomy between "control" and "ecological validity." With current technology it is possible to have both: stimuli can be produced that closely resemble musically appropriate sounds, yet are sufficiently well defined so as to permit replication by other researchers. As a potential tool for researchers, we present in this paper average harmonic spectra for the pitches B0-G7, and also a grand average spectrum from the sum of these pitches. For many applications, these highly replicable spectra will prove more ecologically valid than technical waveforms.

For many phenomena, the choice of stimulus materials may prove unimportant. When different stimuli converge on the same results, we may infer that the specific timbres employed are inconsequential. However, in many other cases, the choice of timbres can prove critical. The results for simple tones (such as sine waves) may differ from results using complex tones (such as recorded piano music). One possible approach to creating an ecologically useful musical stimulus is to identify an "average" musical sound. In this brief report, we describe such an effort.

The notion of an average sound raises a host of questions related to the population of sounds for which some sound purports to be the average. An average value may not be representative. For example, the average value in a bi-modal distribution will not be a “typical” value. Similarly, an “average spectrum” may not be typical. This problem is particularly salient when mixing elements from different domains. For example, in mixing strings, brass and woodwinds together into a single tone, the resulting tone may fail to be representative of any actual tone encountered by listeners. By way of illustration, consider the following facts: (1) the most common sex is female; (2) the most common nationality in the world is Chinese; (3) the most sold musical instrument is the harmonica. From these facts, it would not be appropriate to deduce that the most representative musician would be a female Chinese harmonica player. Apart from the theoretical issue of an “average” there is the practical problem of identifying the population of sounds for which some sound purports to be an average. Ideally, we would aim to use an average tone from the population of all musical sounds heard by listeners. This might involve, for example, sampling the sorts of music to which a listener is exposed and determining the frequency of occurrence for various instruments and various pitches. Moreover, we would expect such a sample to be sensitive to the cultural background of the listener as well as the listener’s stylistic preferences.

Rather than employing the above sampling method, we have elected to pursue a simpler approach. In this study we rely on the SHARC timbre database assembled by Gregory Sandell (1991). The SHARC database consists of harmonic spectra for a variety of standard Western art music instruments. These spectra were generated from recorded instrument tones available in the McGill University Master Samples collection (Opolko & Wapnick, 1987). The Master Samples collection contains a large number of Western instruments spanning roughly five centuries, and includes recordings of every individual pitch produced by such instruments as the alto shawm, krumhorn, harpsichord, piano, etc. In analyzing the spectra for these instruments, Sandell focused on the instruments of the modern classical orchestra. This included the common string, woodwind, and brass instruments. Table 1 provides a complete list. Figure 1 plots the range or compass for each of the SHARC instruments/treatments.

List of all instruments in the SHARC database

Bach trumpet	Eb clarinet	oboe	viola martelé
C trumpet	French horn	tuba	viola muted vibrato
C trumpet muted	French horn muted	bass flute vibrato	viola pizzicato
contrabass	English horn	alto flute vibrato	viola vibrato
contrabass martelé	bass trombone	flute vibrato	violin martelé
contrabass muted	trombone	piccolo	violin muted vibrato
contrabass pizzicato	trombone muted	cello martelé	violin pizzicato
contrabass clarinet	alto trombone	cello muted vibrato	violin vibrato
bass clarinet	contrabassoon	cello pizzicato	violin ensemble
Bb clarinet	bassoon	cello vibrato	

Table 1. A list of all instruments in the Sandell Harmonic Archive (SHARC) database.

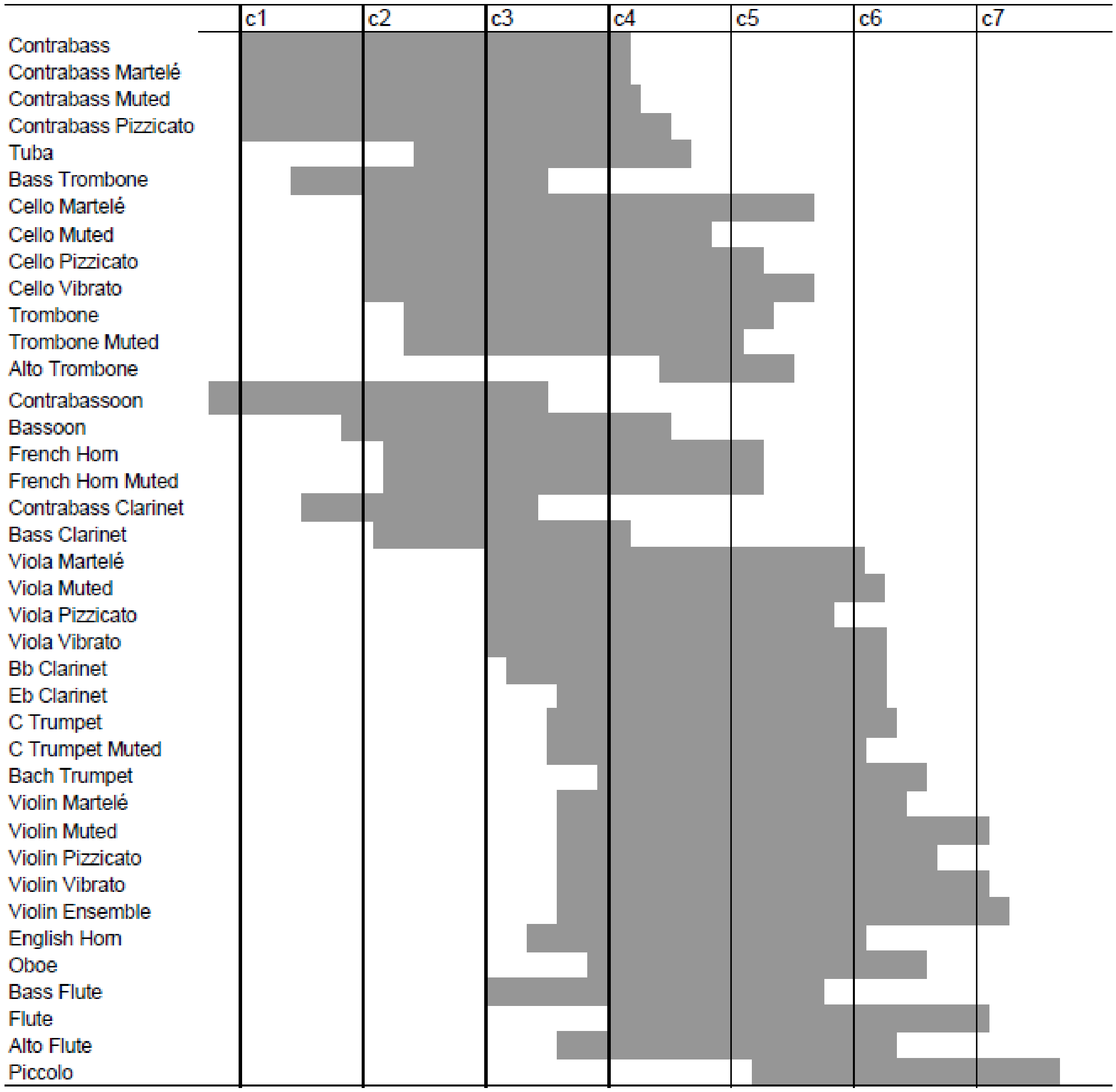

Figure 1. Instrument ranges for all instruments in the SHARC database. The horizontal axis represents pitch from A#0 to G7. Ranges are indicated by shaded bars. N.B. The database available on the web is missing information for contrabass pizzicato G1, contrabass martelé A#3, viola martelé A#3, and violin ensemble D4. In addition, there was corrupted data for the contrabassoon A#0.

METHODOLOGY

The SHARC database represents 39 sound sets produced by 23 instruments and spans the pitches from A#0 to G7. Several instruments are represented more than once in the database. For example, spectral data from a violin played with vibrato is represented separately from a violin performed with a martelé (i.e. hammered) bowstroke. The data are organized as separate note files for each recorded tone. Each note file contains amplitude and phase information for all possible harmonics below 10 kHz. For the purposes of this project, we considered only the amplitude information for each harmonic. In the SHARC database, amplitudes are expressed as decibels relative to the amplitude of the strongest harmonic for that recorded tone.

Our goal was to determine an average harmonic spectrum, which was to be found by averaging the mean harmonic spectra for each pitch in the database. The average spectrum was calculated as follows:

1. Decibels are relative measurements and so cannot be directly averaged together. Each decibel value must first be converted to a relative squared amplitude value before averaging. For each harmonic, the decibel data were converted to squared amplitude data relative to the strongest harmonic using the following formula (Rossing, Moore, & Wheeler, 2002):

$$\text{Relative Squared Amplitude} = 10^{(dB / 10)}$$

2. The mean spectrum for each pitch was determined by averaging the harmonic amplitudes for all available instruments capable of generating that pitch. For example, 12 sounds in the SHARC database are pitched at C2. Hence the average spectrum for C2 combines the data from all 12 of these recordings. By contrast, 34 sounds are pitched at C4; all 34 spectra were therefore averaged for the pitch C4. By way of illustration, Figure 2 plots the average spectra for seven pitches: pitch-class C for each of 7 octaves.

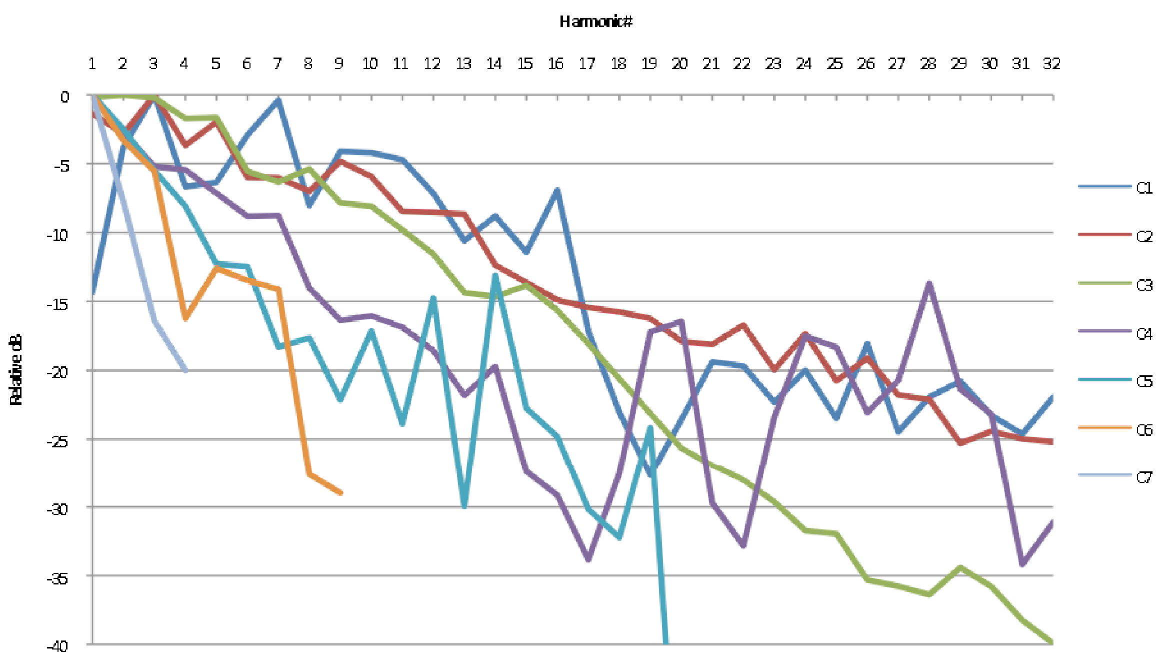

Figure 2. Average harmonic spectrum of pitch-class C for each of seven octaves (C1-C7). The horizontal axis displays the first 32 harmonics. The vertical axis plots the corresponding normalized level where the most energetic harmonic is deemed to be 0 dB.

3. Four instruments were found to have missing data. The corresponding spectra were therefore omitted in the calculation of the average spectrum for the pertinent pitches. In addition, the spectrum for A#0 (the lowest pitch in the database) was found to be corrupted. Spectral data for this pitch was available for only a single instrument (contrabassoon), therefore this pitch was omitted.

4. A mean spectrum for all pitches combined was also calculated by averaging the mean relative squared amplitudes for each pitch (B0 - G7).

5. The mean spectrum of the average harmonic spectrum was converted back to decibels using an arbitrary amplitude reference value of 1.0 = 0 dB:

$$dB = 10 \log_{10} (\text{Amplitude} / \text{Reference Amplitude})$$

6. The pitch B0 exhibited the maximum number of harmonics (324), and so the overall average musical tone was calculated to include 324 harmonics. In order to render this average spectrum consistent with the SHARC database, in which the strongest harmonic is equal to 0, we normalized the average

spectrum by adding 2.28 dB to each component of the spectrum so that the decibel level of the strongest harmonic equaled 0 dB.

7. Harmonics lower than -60 dB are almost certainly inaudible, so these harmonics were removed from the spectrum of the average musical tone. This reduced the number of spectral components to 132 harmonics. Table 2 reproduces these 132 normalized values. The resulting complex tone was synthesized using MAX/MSP (Puckette, 1991).

Harmonic	dB	Harmonic	dB	Harmonic	dB	Harmonic	dB
1	0	34	-22.8656	67	-43.401	100	-53.9705
2	-2.49211	35	-23.2739	68	-43.6745	101	-54.6125
3	-4.12197	36	-25.2515	69	-44.317	102	-54.6505
4	-6.36542	37	-25.3359	70	-44.8094	103	-55.1575
5	-6.25719	38	-26.2105	71	-45.1452	104	-55.3995
6	-7.49675	39	-27.1334	72	-45.7805	105	-55.8417
7	-7.69567	40	-27.3785	73	-46.0414	106	-56.1866
8	-9.26978	41	-29.2297	74	-46.0112	107	-56.2434
9	-8.94704	42	-29.6742	75	-46.9429	108	-56.4011
10	-9.7248	43	-30.2306	76	-47.0066	109	-57.0094
11	-9.95047	44	-30.8659	77	-47.2059	110	-56.6134
12	-10.1035	45	-31.247	78	-47.7963	111	-57.2086
13	-11.2435	46	-32.5626	79	-47.6947	112	-57.2518
14	-12.413	47	-32.8658	80	-48.5066	113	-56.9857
15	-13.1692	48	-33.1618	81	-48.3718	114	-57.8192
16	-13.6776	49	-33.9285	82	-48.9478	115	-57.6201
17	-15.1513	50	-34.8408	83	-49.3049	116	-58.4377
18	-16.0201	51	-35.5584	84	-49.4008	117	-57.876
19	-16.5112	52	-35.7774	85	-50.303	118	-58.8556
20	-16.9168	53	-36.5521	86	-50.4143	119	-58.8964
21	-17.4663	54	-36.9508	87	-50.6155	120	-58.8947
22	-18.5106	55	-37.5281	88	-50.6431	121	-58.5147
23	-18.7082	56	-38.2605	89	-51.5997	122	-59.0432
24	-18.7753	57	-39.0375	90	-51.2573	123	-59.2468
25	-19.2537	58	-39.0979	91	-51.4871	124	-59.3495
26	-19.7006	59	-40.3298	92	-51.7691	125	-58.8656
27	-20.116	60	-40.623	93	-52.0673	126	-58.9029
28	-19.9884	61	-40.9314	94	-52.4427	127	-58.8187
29	-20.6929	62	-41.336	95	-53.0827	128	-59.866
30	-20.1128	63	-41.7689	96	-52.3241	129	-59.9949
31	-21.8898	64	-42.113	97	-52.8004	130	-59.5349
32	-21.5211	65	-42.2577	98	-53.4896	131	-59.9942
33	-22.6679	66	-42.7164	99	-53.7338	132	-59.921

Table 2. The normalized values of the remaining 132 harmonics after the inaudible harmonics were removed.

RESULTS

Figure 3 shows a grand spectral average for all tones in the SHARC database. Only those harmonics with decibel levels greater than -60 dB are plotted. As would be expected, the graph shows a smooth roll-off of energies with increasing harmonic number. Figure 4 provides more detail, showing the amplitudes for the first 32 harmonics. This figure better illustrates that much of the energy is present in the first few partials. The data shown in Figures 3 and 4 could be fitted using an exponential curve. This would be consistent with the average spectrum in speech which has already been shown to exhibit a roughly exponential decay (Cornelisse, Gagne, & Seewald, 1991). Since speech involves tube-like acoustic production, similar to wind instruments, it should not be surprising that there might be similarities in the spectral content.

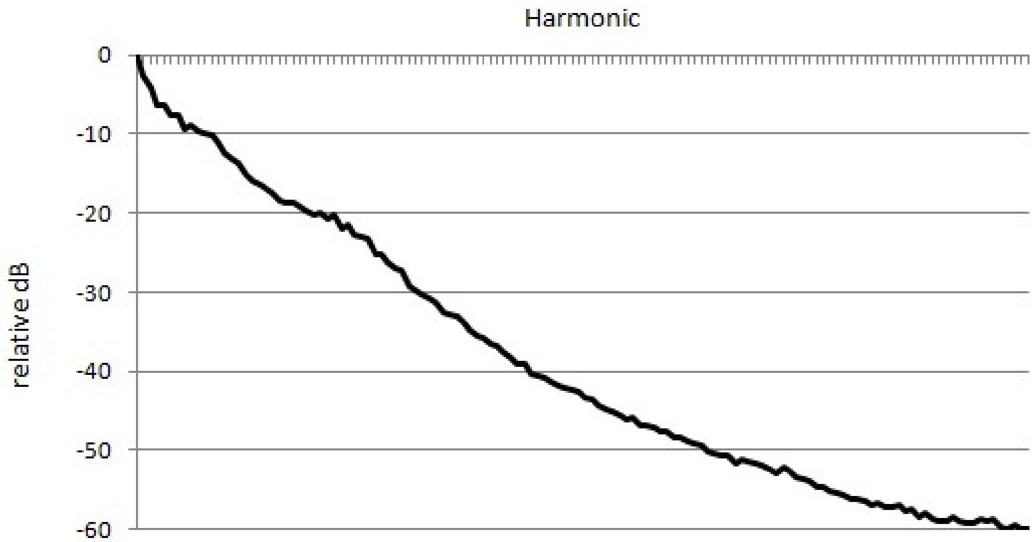

Figure 3. Average harmonic spectrum of all 1,338 tones in the SHARC database. The horizontal axis represents harmonic number ranging from 1 to 132. The vertical axis plots the corresponding normalized level where the most energetic harmonic is deemed to be 0 dB. Only those harmonics with levels greater than -60 dB are plotted.

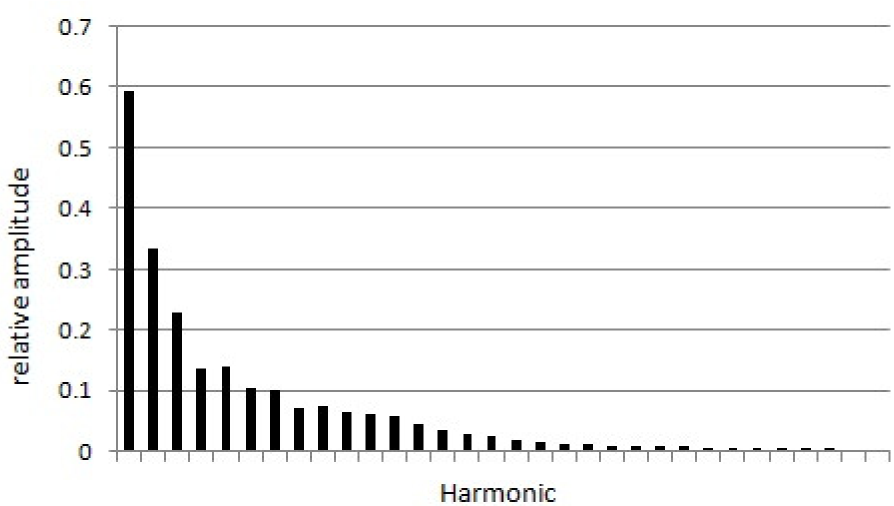

Fig. 4. Average harmonic spectrum of all 1,338 tones in the SHARC database. The horizontal axis represents harmonic number and spans a five octave range (harmonics 1 to 32). The vertical axis plots the corresponding amplitude in linear arbitrary units. Most of the sonic energy is concentrated in the first few harmonics.

LIMITATIONS AND CAVEATS

In presenting these average spectra, it is important for researchers to be fully aware of the underlying assumptions and limitations. Eight caveats bear emphasis:

1. All of the instruments are drawn from Western culture. No non-Western instruments were sampled.
2. The tones analyzed are biased towards orchestral instruments. The two most common musical instruments in Western music (namely the piano and the guitar) are absent.
3. The human voice is entirely absent.
4. The analyses are biased towards classical (art) music, rather than folk or other genres.
5. The sampled orchestral instruments are biased toward harmonic tones rather than inharmonic tones. Pitched percussive instruments such as bells, xylophone, glockenspiel, and timpani are absent.
6. In addition, there is a bias toward pitched rather than unpitched sounds. Hence, the absence of such instruments as cymbals, snare drum, bass drum, wood block, maracas, etc.
7. The spectral averages represent static amplitudes, whereas it is known that musical instruments commonly display dynamic spectra that evolve over the course of the tone (Saldana & Corso, 1964).
8. Although differences in phase are mostly inaudible (von Helmholtz, 1877) no account was taken of the phase information.

These caveats notwithstanding, the average spectra presented in this paper may prove useful in various research applications, such as harmonic modeling (e.g. Parncutt, 1989). Both the spectral recipes and synthesized versions of the average harmonic spectra are permanently archived at the Knowledge Bank website (https://kb.osu.edu/), and are also available as a MAX/MSP virtual instrument.

REFERENCES

Cornelisse, L., Gagne, J., & Seewald, R. (1991). Ear level recordings of the long-term average spectrum of speech. *Ear & Hearing*, *12*(1), 47-54.

von Helmholtz, H. (1877/1954). *Die Lehre von den Tonempfindungen*. Translated by A. Ellis as: *On the sensations of tone as a physiological basis for the theory of music*. New York: Dover Publications.

Huron, D. & Sellmer, P. (1992). Critical bands and the spelling of vertical sonorities. *Music Perception*, *10*(2), 129-149.

Opolko, F. & Wapnick, J. (1987). *McGill University Master Samples*. (Compact discs OWCD 1-87, 2-87, 3-87). Montréal, Quebec: McGill University.

Plomp, R. & Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, *38*, 548-560.

Puckette, M. (1991). Combining event and signal processing in the MAX graphical programming environment. *Computer Music Journal*, *15*(3), 68-77.

Parncutt, R. (1989). *Harmony: A psychoacoustical approach.* New York: Springer.

Rossing, T.D., Moore, F.R., Wheeler, P.A. (2002). *The science of sound*. 3rd ed. San Francisco: Addison Wesley.

Saldana, E.L. & Corso, J.F. (1964). Timbre cues and the identification of musical instruments. *Journal of the Acoustic Society of America*, *36*, 2021-2026.

Sandell, G. J. (1991). A library of orchestra instrument spectra. *Proceedings of the 1991 International Computer Music Conference*. San Francisco, CA: International Computer Music Association, pp. 98-101.