FROM THE SHEPARD TONE TO THE PERPETUAL MELODY AUDITORY ILLUSION

Pedro Patrício
Research Centre for Science and Technology of the Arts (CITAR)
Portuguese Catholic University – School of the Arts
Rua Diogo Botelho 1327, 4169-005 Porto-Portugal
ppatricio@porto.ucp.pt

ABSTRACT

This paper discusses the use of the Shepard tone as a sound source in musical composition. This tone has two musical interests. First, it underlines the difference between the tone height and the tone chroma, opening new possibilities in sound generation and musical perception. And second, considering the fact that it is (in a paradoxical way) locally directional while still globally stable and circumscribed it allows us to look differently at the instrument's range as well as at the phrasing in musical composition. Thus, this paper proposes a method of generating the Shepard tone relying upon an alternative spectral envelope, which as far as we know, has never been used before for the reproduction of the Shepard scale illusion. Using the proposed sound source, it was possible to successfully reproduce the Shepard scale illusion, even when applied to a melody. The melody was called "Perpetual Melody Auditory Illusion" because when it is heard it creates the auditory illusion that it never ends, as is the case with the Shepard scale illusion. Moreover, we composed a music titled "Perpetual Melody – contrasting moments", using exclusively the sound source as sound generator and the melody as musical content.

1. INTRODUCTION

Shepard [1] built an auditory illusion widely known as the Shepard scale illusion through digital sound synthesis. The Shepard scale illusion is a musical scale divided into 12 equal parts. When listened it could create the auditory perception that rises perpetually along a spiral, as happen in the visual plan with the oblique lines of a Barber Pole in motion [2, 3] or with the visual illusion of Escher's "staircase to heaven" [1, 4].

In general terms, Shepard conceived the auditory illusion by cancelling the vertical or rectilinear relationship of the sounds through the creation of the Shepard tone [3].

The Shepard tone is a complex sound constituted by ten components separated by octave intervals, to which a bell-shaped (i.e. Gaussian curve) spectral amplitude envelope is applied.

Using digital sound synthesis techniques, it is possible to isolate perception attributes related to the perception of the height of the sound (i.e. pitch – commonly correlated) in complex sounds (i.e. sounds constituted by more than one component, e.g. noise and Shepard tones). These attributes are the tone height and the tone chroma [3, 5, 16].

The tone height represents the vertical or rectilinear dimension of the sounds while the tone chroma represents its horizontal or circular dimension [1, 5, 6, 7, 8]. In other words, the tone height is the auditory perception attribute that allows one to order the sounds in a scale that extends from the high until the lower-pitched sounds, whereas the tone chroma is the position of the sound within the octave [7].

Thus, in Shepard tone, the tone height of the sound is eliminated while at the same time its chroma is preserved [3]. It will probably be easy to identify the musical note of the sound (i.e. the tone chroma) but it will be very difficult to identify the octave that the sound belongs to (i.e. the tone height) [4]. One could say that a Shepard tone does not belong to any octave in particular, but it belongs to all octaves at the same time, due to the fact that all octaves (ten) are simultaneously represented in the same Shepard tone.

The circular dimension of the Shepard tones allows representing them visually as equally spaced points around a circle [1] (see Figure 1 below).

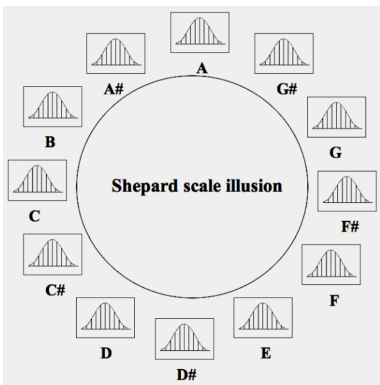

Figure. 1. The figure represents the tone chroma of the Shepard tones in the Shepard scale illusion.

In the Shepard scale illusion, each complex sound has a fixed position and always represents the same sound along any octave, i.e. a C note represents all Cs that are in a guitar. Namely, the tone chroma is always the same for all C, D, E, etc. [5, 16].

One should notice, that Jean-Claude Risset (1938) used the original Shepard scale illusion in the second movement (Fall) of his musical oeuvre titled “Computer Suite for Little Boy” (1968). The illusion could be heard between minute 1.15 and 2.20.

One of the purposes of this paper is not to directly use the original Shepard scale illusion in music compositions, but to go a step further and compose through the Shepard scale illusion applied to a melody.

Furthermore, to successfully reproduce the Shepard scale illusion, an alternative spectral envelope for the individual amplitudes that, as far as we know, has never been used before, is also being proposed in this work.

In the sections which follow, related work with spectral envelopes used in works that revisited the Shepard scale illusion or the Shepard tone and a detailed section about the construction of the sound source will be presented, followed by the sections of experimentation, evaluation and conclusion of this proposal.

2. RELATED WORK

To better structure this section, two types of spectral envelopes sets have been created according to the following criteria: One set for the authors who used fixed spectral envelopes and another set for the authors who used adjustable spectral envelopes.

In the fixed spectral envelopes, the individual amplitudes are always the same for all complex sounds, while in the adjustable spectral envelopes, the individual amplitudes vary along the complex sounds. However, when a cycle of the scale divided into 12 equal parts is completed, the individual amplitudes always return to their original values [8].

2.1 Fixed Spectral Envelopes

The following authors used fixed spectral envelopes: Pollack [9] used a triangular spectral envelope with fixed amplitudes, logarithmically weighted. The individual amplitudes are always the same for the central components (five and six), while the amplitudes of the components either below or above the central components decrease logarithmically and progressively in both directions (see Figure 2 below).

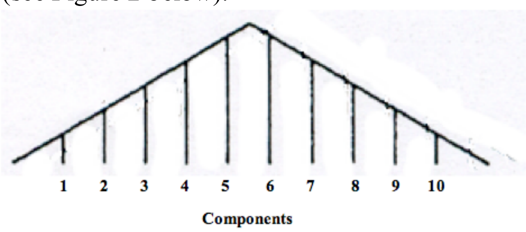

Figure 2. The figure depicts Pollack’s spectral envelope for the individual amplitudes. The vertical lines symbolise the ten components of the complex sound separated by octave intervals, while the oblique lines represent the spectral envelope. Adap. from Pollack [9].

Burns [6] presented a triangular spectral envelope very similar to the shape of the spectral envelope used by Pollack [9], in which the individual amplitudes are weighted in the following manner: minus 6 dB per octave from the central component in both directions, (see Figure 3 below).

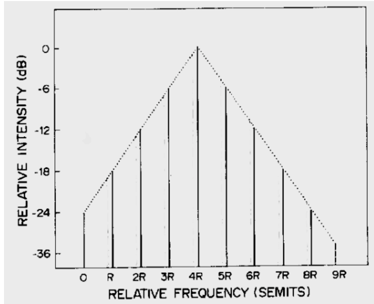

Figure 3. The figure depicts Burns' spectral envelope for the individual amplitudes of a complex sound. The dotted line represents the amplitude envelope. Adap. from Burns [6].

In Nakajima et al. [10] the four central components have the same amplitude, while toward both ends of the spectrum, the sound pressure levels are attenuated by leaps of minus 20 dB per octave (see Figure 4 below).

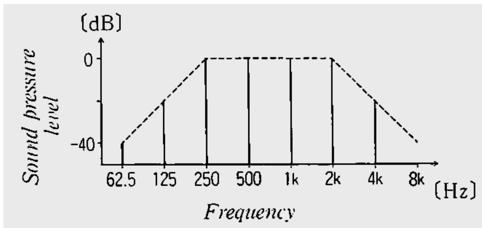

Figure 4. The figure depicts Nakajima’s spectral envelope for the individual amplitudes of a complex sound. The dashed lines represent the amplitude envelope. The image was extracted from Nakajima et al. [10].

2.2 Adjustable Spectral Envelopes

Authors such as Shepard [1], Risset [5], Ueda & Ohgushi [8] (see Figure 5) or Deustch [11] (see Figure 6) used a spectral bell-shaped envelope (Gaussian curve).

This kind of shape of spectral envelope allows a smooth transition among the components and simultaneously highlights the central components while making the components of the extremities become inaudible [9].

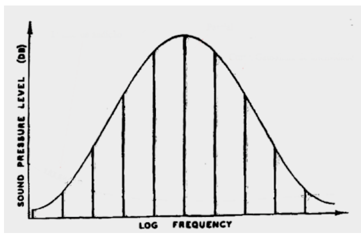

Figure 5. The figure depicts Shepard’s spectral envelope for the individual amplitudes. The vertical lines represent the ten components of the complex sound separated by octave intervals. Adap. from Shepard [1].

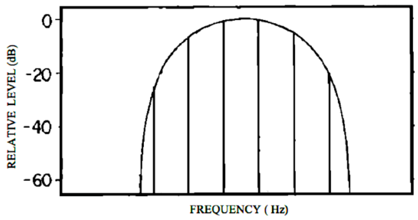

Figure 6. The figure depicts Deutsch's spectral envelope for the individual amplitudes. The vertical lines represents six components of the complex sound separated by octave intervals. Adap. from Deutsch [11].

One should notice, that in all spectral envelopes presented until now, the components of the extremities always have the lowest amplitudes, whereas the central components always have the higher ones.

The construction of the sound source will be discussed in detail in the following sections.

3. SOUND SOURCE

In this work, only digital sound synthesis was used to construct the sound source.[1]

The sound source consists of nine components separated by octave intervals, in which an alternative spectral envelope for the amplitudes is applied.

Primarily, the individual frequencies of the sound source will be automatically determined followed by the corresponding individual amplitudes.

3.1 Automatic Determination of the Sound Source's Individual Frequencies

The first Shepard's equation [1] was implemented in Pure Data (Pd) so as to automatically determine the individual frequencies of the sound source (see equation 1 below).

$$f(t,c) = f_{\min}.2^{[(c-1)t_{\max}+t-1]/t_{\max}}, \quad (1)$$

Where:

$f_{\min}$ = 32.72-Hz, represents the frequency of the first component (fundamental frequency) of the Shepard scale illusion's first complex sound;[2]

$t_{\max}$ = 12, represents the maximum number of complex sounds;

t - represents the complex sound's number;

c - represents the component's number.[3]

[1]A MacBook Intel Core 2 Duo of 2-GHz, an Operating System Mac OS X (version 10.4.11) and the Pure Data (Pd) software (version 0.39.3 – extended) were utilised in the construction of the sound source. More information about Pd at http://puredata.info/

[2]In this work, the original value of $f_{\min}$ (4.863-Hz) of the Shepard's equation [1] was replaced by 32.72-Hz with the purpose of moving the spectrum to the right and thus, obtain a more brilliant timbre and musically more interesting than the original Shepard tone.

[3]Only nine components were used in this work, because after the ninth component and from the fourth complex sound, the individual frequencies overcome the relative maximum limit of the human hearing capacity of a young person, which is around 20-kHz [2, 12, 13].

At the end of this procedure, a chromatic scale divided into 12 equal parts was obtained.[4]

The process of automatic determination of the individual amplitudes of the sound source will be described in detail in the following section.

3.2 Alternative Spectral Envelope Applied to the Sound Source

The alternative spectral envelope purposed in this article is based on a filter used in the measurement of the intensity of noise, called Filter A or A-Weighting curve (see Figure 7 below).

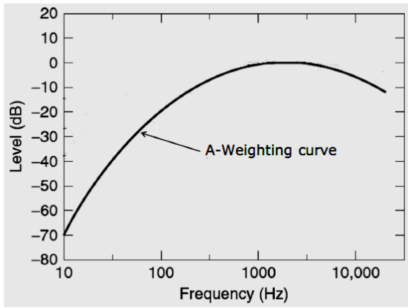

Figure 7. The figure depicts the Filter A or A-Weighting curve. One should notice, that on 1000-Hz (the reference frequency) the sound pressure level is equal to zero dB. Adap. from Wong [14].

The A-Weighting curve was extracted from the inversion of 40 Phon equal-loudness curve of Fletcher & Munson (1933) [2], and intends to simulate the reaction of the human ear throughout the curve [14]. Namely, different frequencies, whose amplitudes lie along the curve, will be listened with the same intensity of a pure tone of 1000-Hz (the reference frequency).

To determine the alternative spectral envelope of the sound source, the following Wong's mathematical equation [14] was used (see equation 2 below).

$$W_A(f) = 20\log\left[\frac{f_4^2 f^4}{(f^2+f_1^2)(f^2+f_2^2)^{\frac{1}{2}}(f^2+f_3^2)^{\frac{1}{2}}(f^2+f_4^2)}\right] - W_{A1000} \quad (2)$$

Where:

W_{A1000} = -2 dB, represents the constant of normalisation;

f = represents the frequency in Hz;

f_1 = 20.6-Hz;

f_2 = 107.7-Hz;

f_3 = 737.9-Hz;

f_4 = 12194-Hz.

The equation allows to reconstruct the A-Weighting curve and to determine the same sonic intensity of any frequency introduced in its variable (f). The remaining

[4]The table that contains the individual frequencies of the sound source is in the Appendix section (see Table 3).

elements of the equation (e.g. W_{A1000}, f_1, f_2, f_3 and f_4) are constants and their values are provided by the author [14].

To automatically compute the individual amplitudes of the sound source, the equation 2 was implemented in Pd.[5]

Figure 8 depicts the image of the alternative spectral envelope obtained for the sound source.

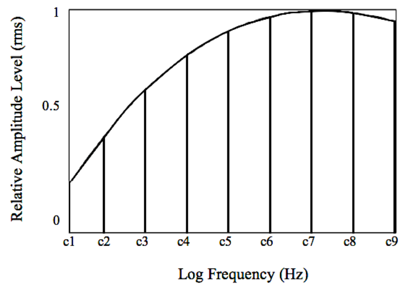

Figure 8. In the figure the vertical lines represent the nine components separated by octave intervals. For musical purposes, the initial values obtained, which ranged between zero and two were converted in Root Mean Squares (RMS) values, namely into the range between zero and one. Furthermore, one should notice, that the rightmost components of the spectrum always have the higher amplitudes.

The following section will be presented so as to test if the sound source is able to successfully reproduce the Shepard scale illusion, even when applied to a melody.

4. EXPERIMENT

In this section, the sound source was applied to two musical exercises, with the purpose to find firstly, if the Shepard scale illusion is successfully reproduced and secondly if the sonic effect of the Shepard scale illusion (sensation of sonic perpetuity) is maintained during the repetition of a melody. For a better view of the melody, its pentagram is presented in the Figure 9.

Figure 9. The melody is constituted by ten musical notes spread along two measures. It goes up one semitone every two measures and is repeated 12 times.

4.1 Method

Twenty students of music performance and composition at ESMAE-IPP, Porto - Portugal served as subjects in this experiment. The group consisted of 14 musicians, 6 composers, 12 male, 8 female, with an average age of 20 years (range 18 – 30 years), and with of an average of 14 years of musical training (range 6 – 20 years).

As equipment, a laptop computer, an application programmed in Pd for playing the musical exercises[6] and a Roland AC-60 loudspeaker for diffusing the sound were used. The sound was diffused in a rehearsal orchestra room, at moderate loudness to all subjects at once.

A questionnaire was delivered to each subject. Before filling it in, the subjects had the opportunity to read information about the Shepard scale illusion, in addition to an oral explanation about the theme in question, and the audition of the original sound of the Shepard scale illusion (the reference sound).[7] Subsequently, the subjects were asked to respond YES or NO to the following questions:

- Was the Shepard scale illusion successfully reproduced (exercise number one);
- Was the sensation of sonic perpetuity maintained during the repetition of the melody (exercise number two).

4.2 Results

After processing the data the results of this experiment can be seen in Tables 1 and 2.

Exercise One		
Answers	Subjects	Percentages
YES	17	85%
NO	3	15%

Table 1. This table presents the results obtained in the first exercise.

Seventeen subjects answered YES, whereas three answered NO. This means that 85 percent of the inquired subjects considered that the sound source reproduced successfully the Shepard scale illusion, whilst 15 percent replied negatively.

Exercise Two		
Answers	Subjects	Percentages
YES	15	75%
NO	5	25%

Table 2. This table presents the results obtained in the second exercise.

Fifteen subjects answered YES, whereas five answered NO. This means that 75 percent of the inquired subjects considered that the sensation of sonic perpetuity was

[5]The table obtained from the individual amplitudes is annexed in the Appendix section (see Table 4).

[6]The application can be downloaded or the audio examples heard at http://sites.google.com/site/pp2007pt/uk/auditory-illusion

[7]The audio file was extracted at http://acousticalsociety.org/about_acoustics/listen_to_sounds/08_03_10_demo27

maintained during the repetition of the melody, whilst the remaining 25 percent responded negatively.

Concerning to overall findings, we obtained an average of 80 percent of the answer YES against an average of 20 percent of the answer NO. Namely, the ratio of answers YES to NO is 4:1. This significant difference of proportion can be observed in the Figure 10.

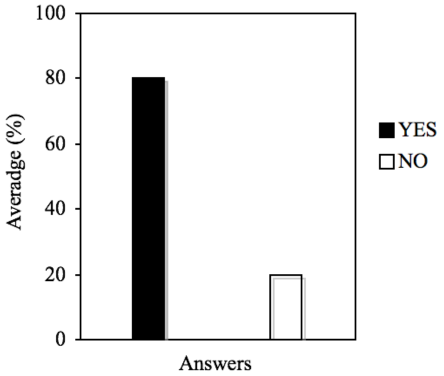

Figure 10. Average of answers YES and NO.

The overall result obtained can be statistically interpreted as follows: In a population of 20 inquired subjects, 4/5 (i.e. 16 subjects) considered that the proposed sound source reproduced successfully the Shepard scale illusion in the first exercise and the sensation of sonic perpetuity was maintained in the second one, whereas 1/5 (i.e. 4 subjects) did not considered that.

4.3 Discussion

The results obtained in the present study show that the proposed sound source method was able to successfully reproduce the Shepard scale illusion, even when applied to the melody. Furthermore, through the alternative spectral envelope used in this work, we also have shown that the bell-shaped spectral envelope used by Shepard [1] it is not an essential factor to produce the auditory illusion, as was already stated by Pollack [9].

A decrement of ten percent in the answer YES between the first and the second exercise was observed. This fact can be explained in the following way: Some subjects (25 percent, Table 2) considered that the constant change of tonality of the melody could weaken the sensation of sonic perpetuity. For these listeners the perpetuity was stronger in the first exercise, because the cycle is simpler, shorter, and consequently easier to follow and predict the moment of the return to the scale's tonic. This argument is contradictory, because when one detects the return to the scale's tonic (i.e. when one perceives a descendant jump in pitch from the last note of the scale to the tonic), meaning that the cycle is broken and the sensation of sonic perpetuity is lost.

Accordingly the majority of the subjects (75 percent, Table 2) the sensation of sonic perpetuity was stronger in the second exercise, due to the fact this be more complex i.e. as it has more musical notes, the cycle becomes larger and more difficult to follow auditorily. Moreover, as the return to the first note of the first measure is imperceptible, it creates the sensation that the melody lasts forever.

5. MUSICAL COMPOSITION

For testing the sonic plasticity and the musicality of the sound source a music titled "Perpetual Melody - contrasting moments" was composed.[8] It was integrally composed and sequenced in Pd.

In the composition only digital sound synthesis, the sound source (as sound generator) and the Shepard scale illusion (applied to the melody as musical content) were used.

To further explore the timbre of the musical work and to produce sonic segregation between the musical voices, we used vibrato generated through a basic circuit of Frequency Modulation (FM) [15].

As musical gestures, transpositions applied to the perception of the height of sound (pitch), envelopes applied to the perception of the intensity of sound (loudness), envelopes of tempo and envelopes of amplitude applied specifically to the attacks and decays were mainly used. Additionally, envelopes applied to the main FM parameters as the carrier frequency, the modulator frequency and the modulation index were also used.

6. CONCLUSIONS

The initial challenges of applying an alternative spectral envelope to the Shepard tone which as far as we know, had never been used before for reproducing the Shepard scale illusion and compose music using exclusively Shepard tone as sound source and the Shepard scale illusion applied to a melody as musical content were widely overcome.

The sound source has successfully reproduced the Shepard scale illusion, even when applied to the melody.

The melody will be called "Perpetual Melody Auditory Illusion" because when it is heard it can create the auditory illusion that it never ends, as is the case with the Shepard scale illusion.

Through the music "Perpetual Melody - contrasting moments", the sound source has shown that it could be a sound generator with an interesting timbre and simultaneously be musically flexible: It works either with slow or fast rhythms, as well as with short or long attacks and decays in amplitude envelopes, even for high, middle and low pitches. For these reasons, the sound source designed and presented in this work could be a fruitfully sound generator for composing music.

The sound source could offer other sonic and musical possibilities of listening due to the particularity of being an ambiguous sound in pitch as happen with the Shepard tone [3]. One could probably say what sound source musical note is, but it would be very difficult to determine which octave it belongs to.

Applying the sound source to other digital sound synthesis techniques, such as Ring Modulation, with the purpose of finding new timbres and gathering a new sound

[8] The Pd patch for playing the "Perpetual Melody - contrasting moments" can be downloaded or an audiovisual file heard at http://sites.google.com/site/pp2007pt/uk/auditory-illusion

bank for composing new music will be the future direction of this work.

Acknowledgments

This work is supported by the Portuguese Foundation for Science and Technology (scholarship SFRH/BD/61599/2009), supervised by António de Sousa Dias. We would also like to thank Eric Oña, António Ferreira, Helena Mena Matos, Sofia Lourenço and especially to Luis Gustavo Martins for his suggestions and critical comments.

7. REFERENCES

[1] R. N. Shepard, “Circularity in Judgments of Relative Pitch,” in J. of the Acoustical Society of America, 36 (12), 1964, pp. 2346-2353.

[2] G. Loy, “Musimathics1: The Mathematical Foundations of Music,” Cambridge University Press, Cambridge, UK, 2006.

[3] R. N. Shepard, “Pitch Perception and Measurement,” Cognition, and Computerized Sound: An Introduction to Psychoacoustics,” In P. R., Cook (eds.), Cambridge University Press, Cambridge, UK, 1999, pp. 261–275.

[4] B. Fugiel, “Waveform Circularity From Added Sawtooth And Square Wave Acoustical Signals,” in Music Perception 28 (4), 2011, 415-423.

[5] J. C. Risset, “Paradoxes de Hauteur,” Communication au Symposium IRCAM sur la psychoacoustique musicale. University Lyon II, France, 1977, pp. 1–6.

[6] M. E. Burns, “Circularity in relative pitch judgments for inharmonic complex tones: The Shepard demonstration revisited, agai,” in Perception & Psychophysics 30 (5), 1981, pp. 467–472.

[7] D. Deutsch, K. Dooley, T. Henthorn, “Pitch circularity from tones comprising full harmonic series,” in J. of the Acoustical Society of America 124 (1), 2008, pp. 589-597.

[8] K. Ueda and K. Ohgushi, “Perceptual components of pitch: Spacial representation using a multimensional scaling technique,” in J. of the Acoustical Society of America 82, 1987, pp. 1193-1200.

[9] I. Pollack, “Decoupling of Auditory Pitch and Stimulus Frequency: The Shepard Demonstration Revisited,” in J. of the Acoustical Society of America 63, 1978, pp. 202-206.

[10] Y. Nakajima, T. Tsumura, S. Matsura, H. Minami and R. Teranishi, “Dynamic Pitch Perception for Complex Tones Derived from Major Triads,” Music Perception 6 (1), 1988, pp. 1-20.

[11] D. Deutsch, F. R. Moore and M. Dolson, “The perceived height of octave-related complexes,” in J. of the Acoustical Society of America 80 (50), 1986, pp. 1346-1353.

[12] D. Benson, “Music: A Mathematical Offering,” Web version 2008. Retrieved 2 November 2010 from http://www.maths.abdn.ac.uk/~bensondj/html/music.pdf

[13] M. Mathews, “What is Loudness?,” Cognition, and Computerized Sound: An Introduction to Psychoacoustics. In P. R., Cook (eds.), Cambridge University Press, Cambridge, UK, 1999, pp. 261–275.

[14] G. Wong, “Sound Level Meters,” The Handbook of Noise and Vibration Control. In M. J., Crocker (eds.) Chapter 38, New York, 2007, pp. 455-464.

[15] J. Chowning, “Perceptual Fusion and Auditory Perspective,” Cognition, and Computerized Sound: An Introduction to Psychoacoustics. In P. R., Cook (eds.), Cambridge University Press, Cambridge, UK, 1999, pp. 261–275.

[16] J. C. Risset, “Paradoxes de Hauteur: Le concept de hauteur sonore n’est pas le meme pour tout le monde,” Seventh International Congress on Acoustics, 20 S 10, Budapest, 1971, pp. 613–616.

Appendix

Number of complex sounds (s)	Number of components (c) and individual frequencies (Hz) of the sound source								
	c1	c2	c3	c4	c5	c6	c7	c8	c9
s1	32.72	65.44	130.88	261.76	523.52	1047.04	2094.08	4188.16	8376.32
s2	34.6656	69.3313	138.663	277.325	554.65	1109.3	2218.6	4437.2	8874.4
s3	36.727	73.4539	146.908	293.816	587.631	1175.26	2350.53	4701.05	9402.1
s4	38.9109	77.8217	155.643	311.287	622.574	1245.15	2490.29	4980.59	9961.18
s5	41.2246	82.4492	164.898	329.797	659.594	1319.19	2638.38	5276.75	10553.5
s6	43.676	87.3519	174.704	349.408	698.815	1397.63	2795.26	5590.52	11181
s7	46.2731	92.5461	185.092	370.185	740.369	1480.74	2961.48	5922.95	11845.9
s8	49.0246	98.0492	196.098	392.197	784.394	1568.79	3137.58	6275.15	12550.3
s9	51.9398	103.88	207.759	415.518	831.036	1662.07	3324.14	6648.29	13296.6
s10	55.0283	110.057	220.113	440.226	880.452	1760.90	3521.81	7043.62	14087.2
s11	58.3004	116.601	233.202	466.403	932.807	1865.61	3731.23	7462.45	14924.9
s12	61.7671	123.534	247.069	494.137	988.274	1976.55	3953.10	7906.19	15812.4

Table 3. Values of the individual frequencies obtained through the equation 1. These values correspond to the scale of C divided into 12 equal parts.

Number of complex sounds (s)	Number of components (c) and individual amplitudes (rms) of the sound source								
	c1	c2	c3	c4	c5	c6	c7	c8	c9
s1	0.22095	0.476516	0.670893	0.814383	0.91752	0.97788	0.999143	0.992686	0.94811
s2	0.244939	0.494914	0.684613	0.824462	0.924262	0.980953	0.999602	0.990797	0.941664
s3	0.268394	0.512899	0.697982	0.834276	0.930697	0.983757	0.999888	0.988645	0.934699
s4	0.291329	0.530476	0.71101	0.843826	0.936821	0.986301	1	0.986212	0.927199
s5	0.313759	0.547648	0.723706	0.853113	0.94263	0.988595	0.999939	0.983477	0.919151
s6	0.335699	0.564417	0.73608	0.862133	0.948122	0.990651	0.999703	0.98042	0.910545
s7	0.357161	0.580788	0.74814	0.870886	0.953299	0.992478	0.99929	0.977019	0.901372
s8	0.378158	0.596765	0.759896	0.879366	0.95816	0.994087	0.998694	0.973252	0.891632
s9	0.398702	0.612351	0.771357	0.88757	0.962708	0.995486	0.997909	0.969095	0.881323
s10	0.418802	0.627551	0.78253	0.895493	0.966949	0.996684	0.996926	0.964525	0.870452
s11	0.438466	0.642369	0.793421	0.90313	0.970886	0.997688	0.995736	0.95952	0.859023
s12	0.457702	0.656814	0.804038	0.910474	0.974527	0.998506	0.994327	0.954055	0.84705

Table 4. Values of the individual amplitudes obtained through the equation 2. For all complex sounds, the amplitudes increase gradually from the first to the seventh component and decrease from the seventh to the ninth component.