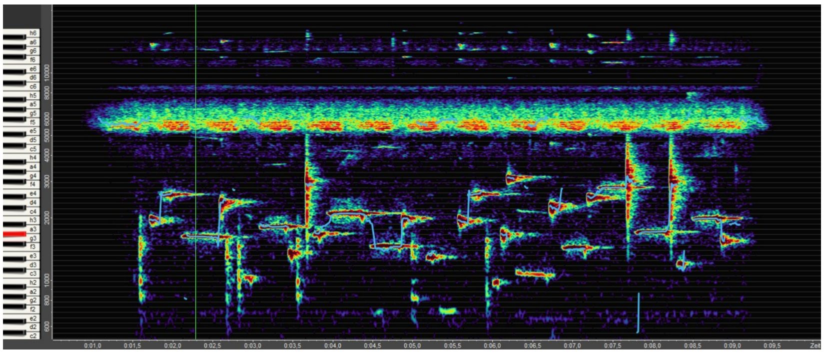
Musician Wren: Singing of 3 different birds (1 verse each)

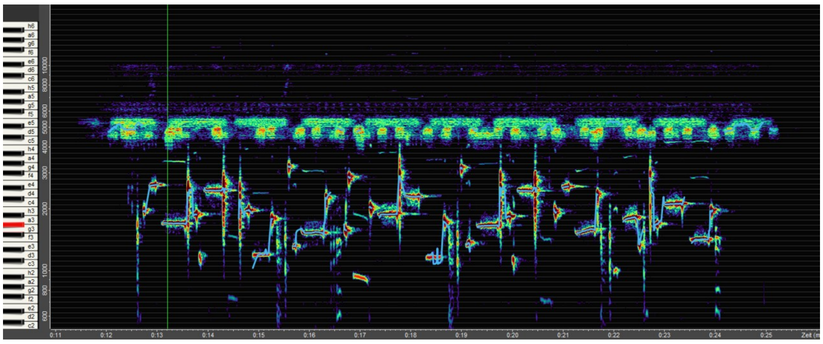
From Wikipedia:
"The Musician wren lives in South America. Its song consists of antiphonal tones of both sexes. It is a series of clear, insistent whistles that vary greatly in pitch. A phrase is repeated many times with little variation before he switches to another song. His call includes a harsh chuk sound."

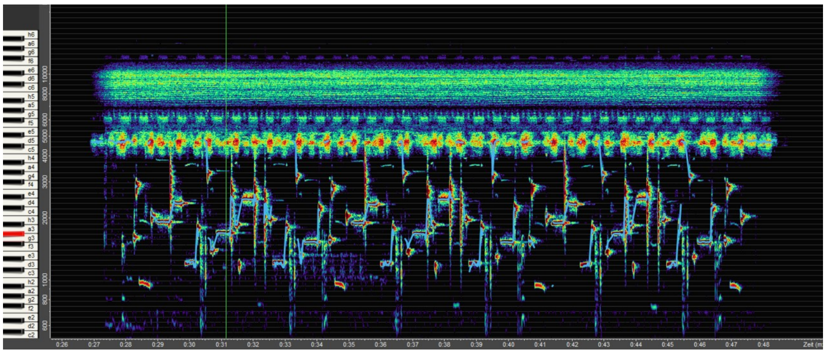
Video: Musician Wren / Orpheuszaunkönig (Cyphorhinus arada) - slowed down 0-2-4-8-16x - with notation https://youtu.be/130WudUnjj0

The vocal range is slightly more than 2 octaves, E5 (660 Hz) to G#7 (3250 Hz). The entire partial tone spectrum extends up to 96 kHz. At G#6, the partials can be read up to the 28th partial (seventh F#11 at 48 kHz).
In the spectrogram of wren 1, the frequency spectrum above the cricket sound between the 4th and 8th partial (11-16 kHz) can be clearly and calculably recognized.
The lower tones sound only faint, but the very clear and intense whistle or flute tones between 1500 and 3000 Hz lie in the frequency range in which our ears can still distinguish pitches relatively well and in which our hearing is particularly sensitive.

(In all 3 recordings different crickets can be heard, in the first at 5500 Hz, in the second at 4500 Hz and at 5300 Hz, in the third at 4500 Hz, 6000 Hz and 7.5-10.5 kHz).







All 3 birds repeat their verses in their song with slight variations. The long verse of wren 3 contains three repetitions with a few small changes. However, it also sings shorter verses.
Duration: 1) 7.5 s - 2) 11.5 s - 3) 20 s

The first impression is that each wren jumps up and down in rapid alternation between pitches that are further apart. It is a rhythmically variable sequence of very short and some somewhat longer, more clearly perceptible notes. Wren 1 has a total of 28 notes with 6 stronger sounding notes (sequence: G#6-A#6-C7-F#6-B6-F6) and Wren 2 has 38 notes with 12 stronger sounds (G#6-D#7-D6-F#6-A#6-D7-G#6-D#7-F#6-A6-C7-F6).

The specific rhythm is created not only by the variable succession of several short and single longer notes as well as high and low notes, but also by the fact that very short, sharp noises are repeatedly heard between groups of 2-6 notes. When slowed down, it can be recognized that these short noises create a structure of individual connected interval sequences, e.g. B3-E4-G#3-D4 or C3-A3-F3 or D4-D3-A3 or G3-F4-A#2-B3-A#3 or D3-B3-G3.

However, although individual sounds can be distinguished relatively clearly for our ears as pitches, we cannot establish a direct connection between the tones in the rapid succession between the sometimes very large intervals. It sounds more like "new music", which is sometimes not very singable and jumps up and down in the intervals without a traceable melody.

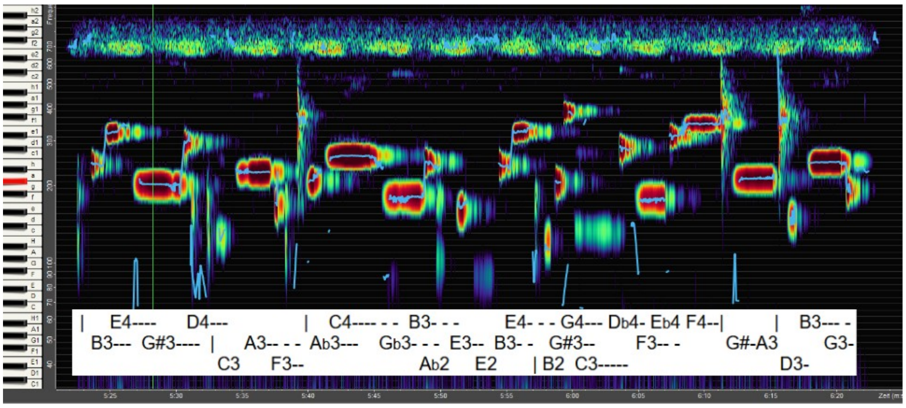
But if I play the first motif (B3-E4-G#3-D#4) as a loop, I can spontaneously whistle along the "E major triad" (B3-E4-G#3) without any problems (and without knowing what notes they are). However, the high D#4 is so short that I don't recognize it. The human whistling tones are in the same frequency range as the "whistling tones" of the wren (1945 Hz / 2589 Hz / 1631 Hz = 6:8:5 = fifth:octave:third). This means that I can whistle exactly the same notes as the wren.
If I take a closer look at the whole sequence of notes in the slowdown, I see that the bird glides from G#3 to G3 on the third note, and then the D4 is in the exact ratio of a fifth (3:2) to the G3.

The spectrogram of the other motifs in the verse also shows that it is not a random sequence of intervals, but that most of the tones are in a spectral relationship to one another. Like all somewhat more intense sounds, the first B3 continues to sound when the E4 is heard, and the echo of the E4 continues to resonate almost to the end of G#3. And even the short D4 continues to sound, although the bird makes a bump noise immediately after the D4. When he then enters the second motif with C3, this is the exact ninth (D4/C3 = 9:4) to the still sounding D4. And so it continues with the sequence C3-A3-F3 (3:5:4), a pure "F major" triad in which the third A/F sounds quasi two-part at the end. There is another similar triad as a quartal sext sound at the end of the verse D3-B3-G3 ("G major" - 151:249:199 Hz = 3:5:4).
This sounding together ("consonare") applies to all pitches and their relationships in a phrase.
It is not to be confused with our concept of "consonance". The interval G#3-D3 (a "tritone") also sounds exactly together in the ratio 7:5 and that doesn't sound dissonant at all, not even for our musically and culturally imprinted ears.

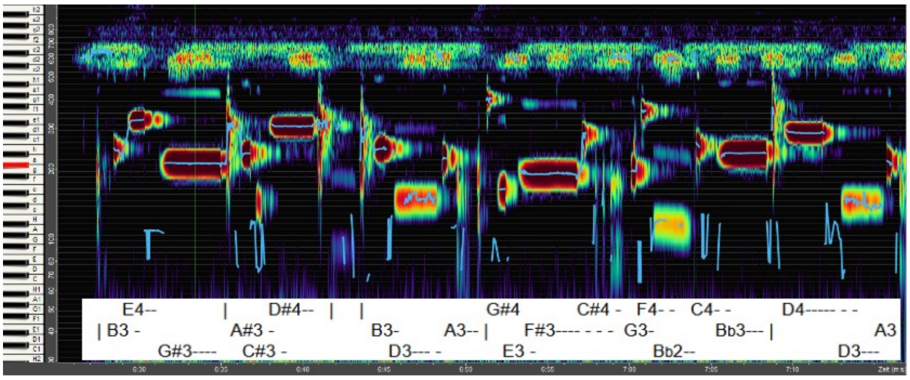
I listened to a total of 20 recordings of Musician Wren singing and found the same verse models with similar jumping interval sequences and variable rhythm in all of them.

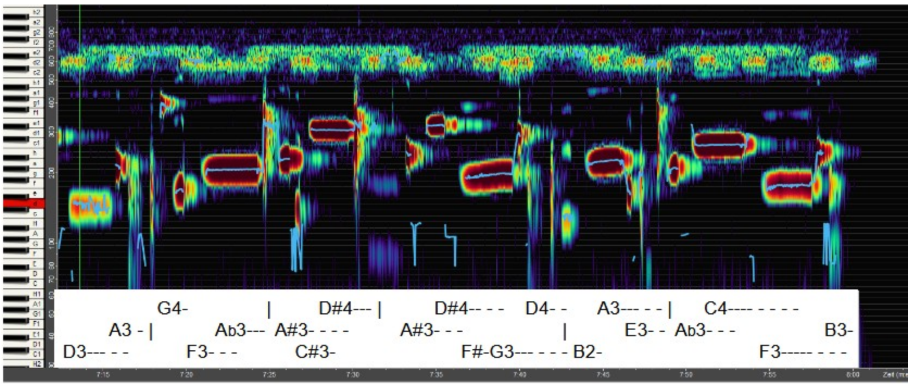
All 3 verses in the spectrogram 8-fold slowed down with notation:

Wren 1

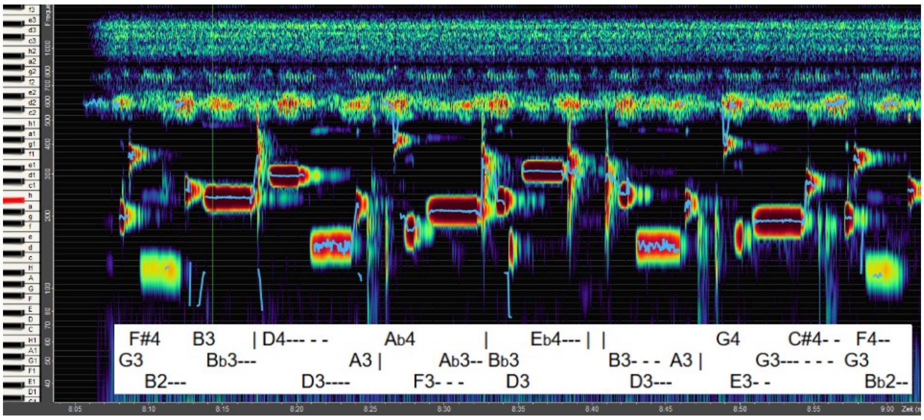


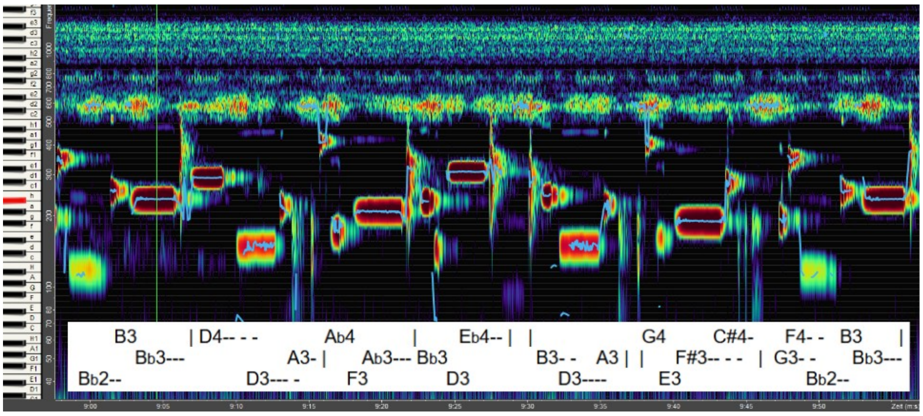
Wren 2


E4-- | D#4-- | | G#4 C#4 - F4- - C4- - D4----- - -
| B3 - A#3 - B3- A3-- | F#3---- - - - G3- Bb3--- | A3
G#3---- C#3 - D3--- - E3 - Bb2-- D3---
6:30 6:35 6:40 6:45 6:50 6:55 7:00 7:05 7:10 Zeit (m:s)


G4- | D#4--- | D#4-- - - D4- - A3--- - - | C4---- - - - -
A3 - | Ab3--- A#3- - - - A#3- - - | E3- - Ab3- - - B3-
D3--- - - F3- - - C#3- F#-G3--- - - - B2- F3----- - - -
7:15 7:20 7:25 7:30 7:35 7:40 7:45 7:50 7:55 8:00

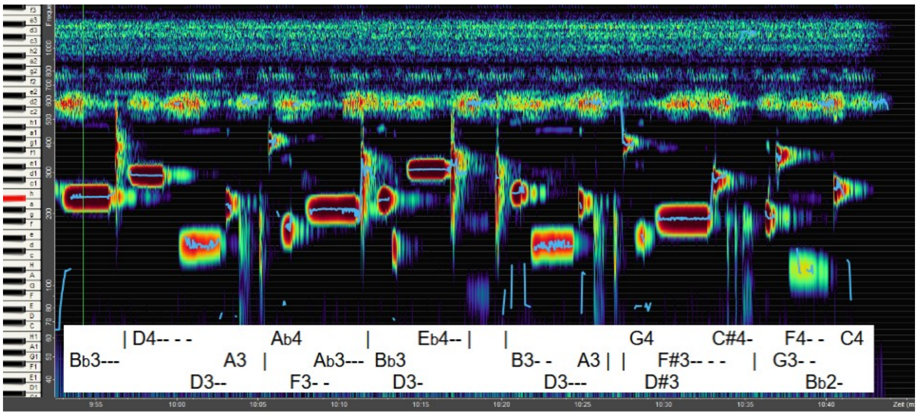
Wren 3

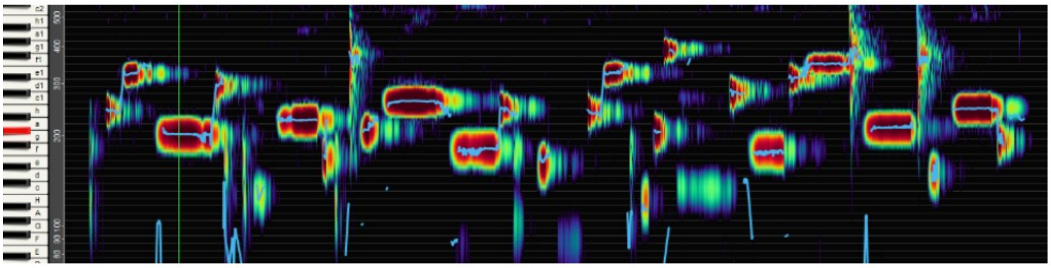

F#4 B3 | D4--- - - Ab4 | Eb4--- | | G4 C#4- - F4--
G3 Bb3--- A3 | Ab3-- Bb3 B3- - - A3 | G3--- - - - G3
B2--- D3---- F3- - - D3 D3--- E3- - Bb2--
8:10 8:15 8:20 8:25 8:30 8:35 8:40 8:45 8:50 8:55 9:00 Zeit


B3 | D4-- - - Ab4 | Eb4-- | | G4 C#4- F4- - B3 |
Bb3--- A3- | Ab3--- Bb3 B3- - A3 | | F#3-- - - | G3- - Bb3---
Bb2-- D3--- - F3 D3 D3---- E3 Bb2--
9:00 9:05 9:10 9:15 9:20 9:25 9:30 9:35 9:40 9:45 9:50 Zeit (m:s)

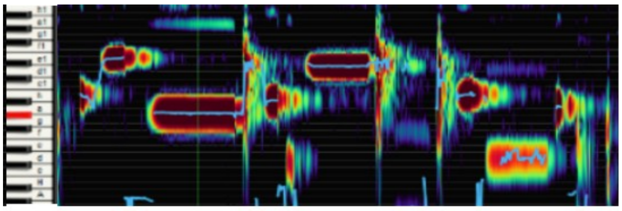
Wren 1: all intervals in their harmonical proportion

(calculated according to the exact frequency ratios in the spectrogram)

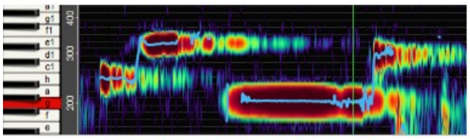


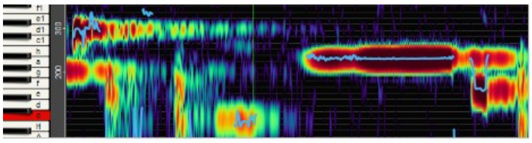


B3-E4 (6:8 Quarte) - E4-G#3 (8:5 Sexte) - G3-D4 (2:3 Quinte) - D4-C3 (9:4 None) - C3-A3 (3:5 Sexte) - A3-F3 (5:4 Terz) - Gb4-Ab3 (7:4 Septime) - Ab3-4 (4:5 Terz) - C4- Gb3 (10:7 "Tritonus") - Gb3/F#3-B3 (6:8 Quarte) - B3-E3 (3:2 Quinte) - E3-E2 (2:1 Oktave) - E2-B3 (1:3 Quinte) - B3-E4 (3:4 Quarte) - E4-B2 (8:3 Quarte) - B2-G#3 (3:5 Sexte) - G#3-G4 (7:13 Große Septime) - G4-C3 (3:1 Quinte) - Db4-F3 (8:5 Sexte) - F3-Eb4 (4:7 Septime) - Eb4-F4 (8:9 Sekunde) - F4- Ab3 (10:6 Sexte) - Ab3/A3-G4 (4:7 Septime) - G4-D3 (8:3 Quarte) - D3-B3 (3:5 Sexte) - B3-G3 (5:4 Terz)



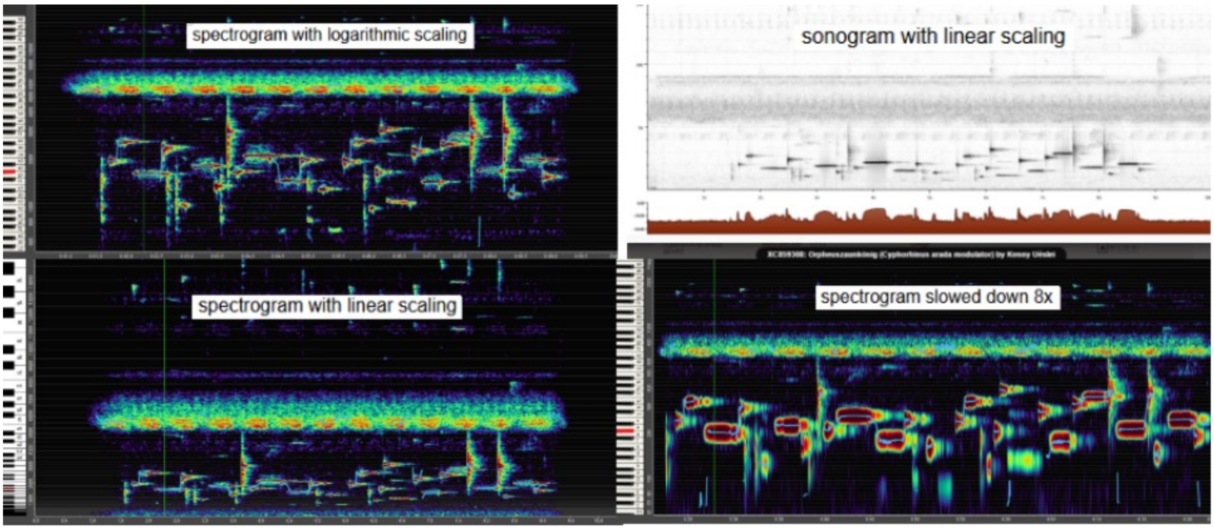
Wren 2 begins its verse in the same way as Wren 1, but it remains constant on G#3 and then moves to the fifth D#4. While the short D#4 continues to sound until the next long D#4, it moves a fourth lower to A#3 (4:3), which also continues to sound, and from there a sixth (5:3) lower to C#3 and then returns in a ninth (4:9) to D#4. The next phrase begins a third lower with B3, to which the sixth (5:3) D3 sounds with the following fourth (2:3) G3. It is a sounding G major triad.





Wren 1: In this enlargement of the first two phrases, it is clearly recognizable that all notes not only sound one after the other, but that the following intervals form a sounding together. At the beginning there is an E major triad (B-E-G#), then a fifth, towards the second phrase a ninth and then an F major triad (C-A-F). At the end the third A/F continues sounding.

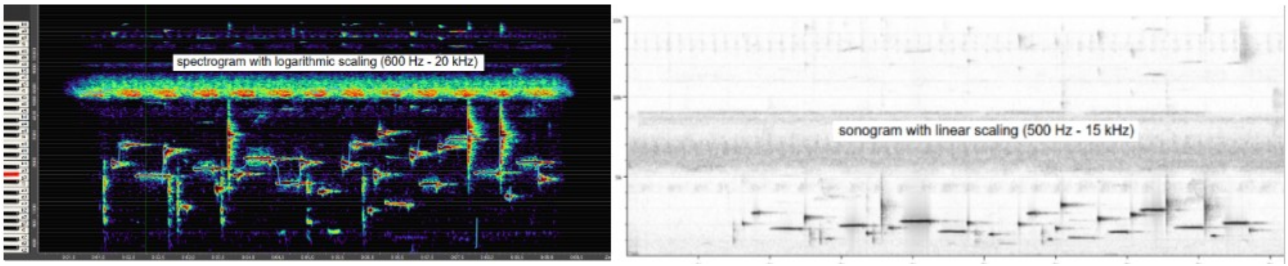
From the appendix of the video:
Comparison of different types of spectrograms



My analyses are carried out using the *"Overtone-Analyzer"*, with which the spectrogram of an audio file can be displayed in progress. The pitch and volume of each partial tone or each frequency of a sound spectrum can be read, the frequency ± 50 cents and the volume up to tenths of a decibel.
In the original position, the complete spectrum up to 22 kHz can be read. The relative intensity of each frequency in the spectrum is indicated by the color, the dark red partials are the loudest. The amplitude of the oscillation can be recognized by the width of the respective frequencies. As can be seen in the 8-fold deceleration, it can, for example, comprise a fourth at G#6 (1640 Hz) (F6-Bb6 - 1376-1896 Hz).

At 2000 Hz (B6 for the musician wren), a semitone is about 260 Hz (about 25 Hz for A3).
In the range between 1500 and 8000 Hz, all the songbirds I analyzed sing octaves (1:2), fifths (2:3), fourths (3:4), thirds (4:5 or 5:6) and also sevenths (4:7) with a difference of sometimes only 5 Hz, unless the proportion is even frequency-accurate, which is usually the case.
For the analysis it is helpful that on the Overtone-Analyzer you can zoom into a sound in the vertical of the spectrum and in the horizontal time axis like with a sound microscope. A pitch marker shows the loudest sound in a spectrum, which is not always the fundamental tone.
Compared to the song of the blackbird or the robin, the song of the musician wren is very, very simple and can easily be analyzed with frequency accuracy.

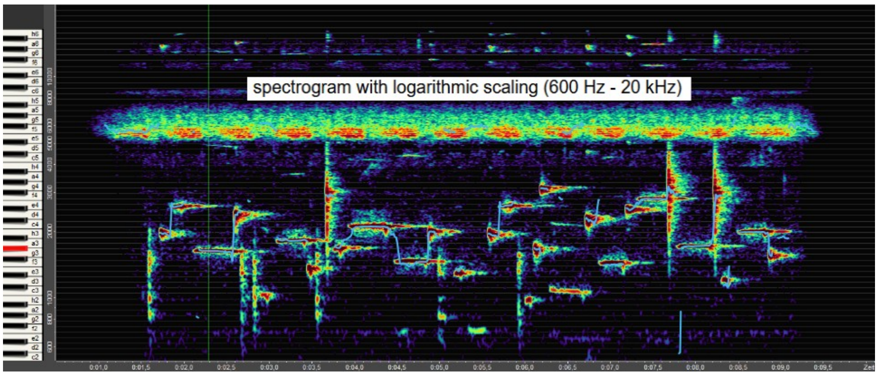
Comparison of logarithmic spectrogram and linear sonogram



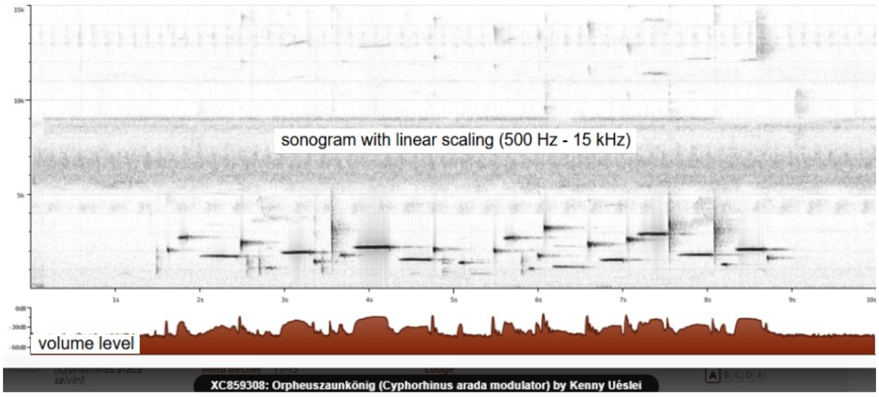
The *logarithmic spectrogram* shows at a glance the relative interval ratios of octaves, fifths and thirds in the range from 700 to 3000 Hz, as well as the condensed noise-like spectrum of cricket song at 5-6000 Hz and the high partial tone spectrum of the wren between 11 and 16 kHz (3rd-8th partial tone), in which all partial tones can be precisely determined.
I can use the cursor to determine the frequency and relative volume for each pitch, regardless of whether it is the fundamental or a partial. However, this is even more precise in the 8-fold slowdown.

In the *linear sonogram* (5-10-15 kHz), the intervals of the song lie evenly close together in diffuse gray stripes in the range of 500-5000 Hz. The cricket song occupies the same spectrum range as the song of the wren and in the range of 12-15 kHz hardly any of the partial tone spectrum can be recognized.

spectrogram with logarithmic scaling (600 Hz - 20 kHz)



sonogram with linear scaling (500 Hz - 15 kHz)



In bird song research, *sonograms* are mainly used, which, compared to the Overtone-Analyzer only reproduce the sound spectrum very diffusely and imprecisely and with which one cannot look *into* the spectrum and hear the sound and the complete frequency spectrum with its "overtones" (partials). One sees only gray bands with no frequency-accurate pitch information, and the exact frequency spectrum of the high vibrations in a sound is hard to discern, even though these high frequencies are crucial for analyzing and identifying the sounds.
In addition, the pitch scale in the sonagrams is linear, which does not correspond to the actual, i.e. integer logarithmic ratios in the frequency spectrum (1:2:3:4:5...), which means that an incorrect "sound image" ("sonagram") is displayed.

In a **linear scaling** (1000 Hz - 2000 Hz - 3000 Hz ...), the elementary structure of the integer frequency ratios cannot be captured. All partials of a sound then have the same linear distance to each other. The ratio 1:2 applies to the octave 1000/2000 Hz (B5/B6) as well as to the octave 2000/4000 Hz (B6/B7). The octave B5/B6 (1:2 - 1000/2000 Hz) at 1000 Hz is then the same size as the fifth B6/F#7 (2:3 - 2000/3000 Hz) or the fourth F#7/B7 (3: 4 - 3/4 kHz) or the major third B7/D#8 (4:5 - 4/5 kHz) or the minor third D#8/F#8 (5:6 - 5/6 kHz) or the "third" F#8/A8 (6:7 - 6/7 kHz) or the "second" A8/B9 (7:8 - 7000/8000 Hz). None of this corresponds to the actual physical proportions *within* a sound. A fifth remains a fifth, regardless of whether it comprises 1000 Hz between the 2nd and 3rd partial (B6/F#7) or 2000 Hz between the 4th and 6th partial (B7/F#8). The ear in birds and humans hears both logarithmically as a fifth ratio. It transforms the ratio 2:3 kHz and the ratio 4:6 kHz into the sound-gestalt of a fifth, a specific spectrum pattern.

The noise-like quality in the chirping sound of birds is created precisely by the compression of the frequency spectrum in these high registers, which prevents our ears from distinguishing pitches. Added to this is the speed in the sequence of tones, which we cannot realize at this speed through our hearing and which on the time scale is of course displayed linearly.

Since these sonograms only indicate the dynamic level for the overall sound and do not give any precise information about the volume of each partial tone, it is not possible to make any statement about the character and structure of a sound (2-part ? - spectral sound ? - timbre ? - trills ? - vibrato ? - pulsation ? etc.).

Above all, the **octaving slowdown** and its display on the Overtone-Analyzer offers a previously unknown and practiced method in bird song research for precisely analyzing and comprehensively researching bird song in all its elements. This applies not only to the components of the singing, which are to some extent recognizable as sequences of tones, but also to the overwhelming majority, which in the original is only perceived as a diffuse noise and cannot be represented concretely and differentiated in the sonagram at all. With conventional methods, this proportion can only be described by the sound/noise surface or by the structural elements in the course of the singing, which are then statistically examined in scientific studies without knowledge of the *actual sounds* (!).
This means that one can only make sufficient statements about *what* and *how* the birds *sing* and *hear* using real spectrograms, such as those provided by the Overtone-Analyzer.

The logarithmic ear

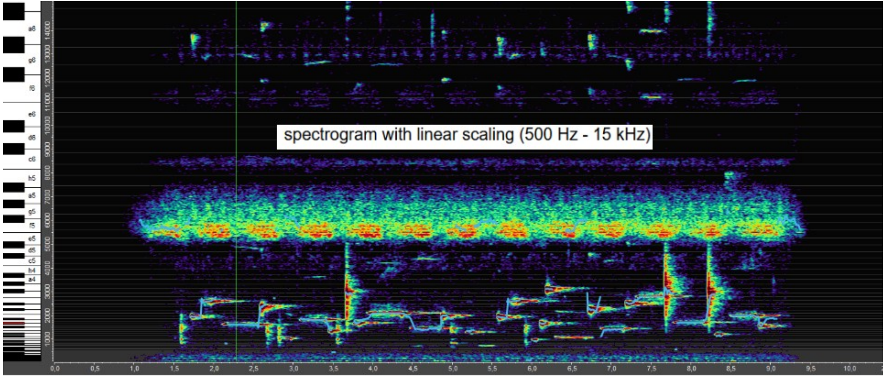
Humans and songbirds such as the blackbird or the musician wren can hear and sing a fifth because there is a corresponding circuit in their auditory/vocalization system, an innate conception of gestalt for sound structures such as the gestalt of a fifth. This innate circuit is rooted in the physiology and function of the auditory system, which has developed over the course of evolution to perceive and transform sound events in accordance with the physical order of the natural laws of sound and vibration. It is the "non-conscious ratiomorphic apparatus" (Konrad Lorenz) that calculates the vibration ratio 2:3 from the sound - the pattern or gestalt "fifth".
This "fifth" can represent the simultaneous and successive sounding of 2 tones (C and G) with their own spectrum and the ratio of two partials to each other in a spectrum (2:3, 4:6:, 6:9, 8:12, 12:18 etc.).
And this applies to a fifth of the wren at F6 and C7 (1376 : 2064 Hz = 2:3) in the same way as to a fifth that I sing from F3 to C4 (172:258 Hz = 2:3), and to the triad C6-A6-F6 (3:5:4 - 688:1720:1376 Hz), which I can whistle exactly at these frequencies by ear.

The hearing of songbirds and the hearing of humans calculates **logarithmically** (base number 2). This is why the scaling of frequencies ("pitches") in the spectrogram is not linear, but logarithmic. Humans and songbirds hear and intonate "pitches" or intervals logarithmically, just as the spectrum of each tone/sound is logarithmically constructed and structured, from the fundamental tone or 1st partial up to the highest partials (1:2:3:4:5:6:7 ... 14:15:16:17:18 ... 22:23:24:25 ....).
Spectral structures in which there is friction and no clear proportions have less sound energy and cannot be received as clearly.

next page: Spectrogram with *linear scaling and with logarithmic scaling*

For comparison, here is a spectrogram with *linear scaling* in the original position and in the 8-fold slowdown. The keyboard shows how close the pitches of the vocals are to each other. I could use the pitch marker to define the pitch, but since a semitone in this range is 260 Hz, it would be extremely imprecise. It would be impossible to recognize the glide from G6 to G6. In the 8-fold deceleration and in the logarithmic spectrogram, I can read the semitone change from the keyboard alone.

spectrogram with linear scaling



spectrogram slowed down 8x (60 - 2700 Hz)