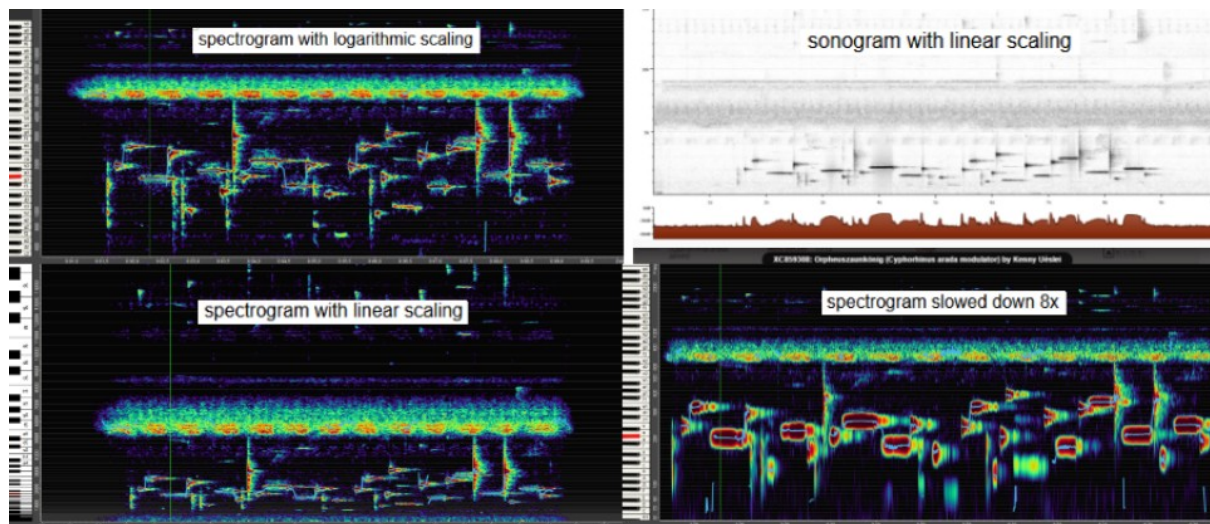# Comparison of logarithmic spectrogram and linear sonogram -

The logarithmic ear

Excerpt from the text: "Musician Wren: Singing of 3 different birds"
https://www.entfaltungderstimme.de/pdfs/Klangkosmos/Musician-Wren_Song-Analysis.pdf

## 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 A4).
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.
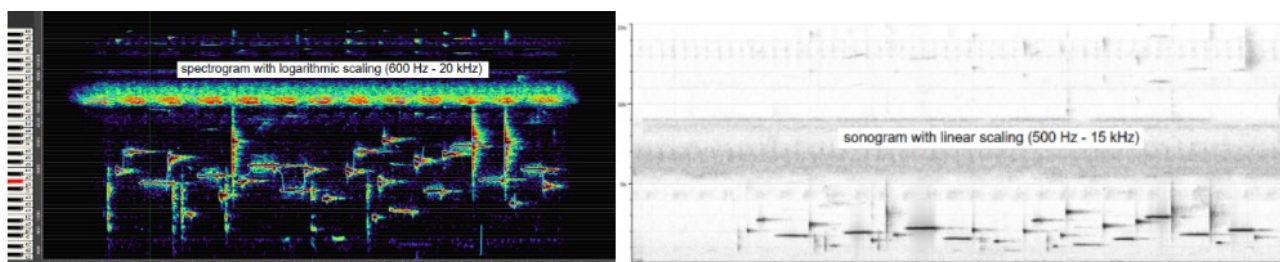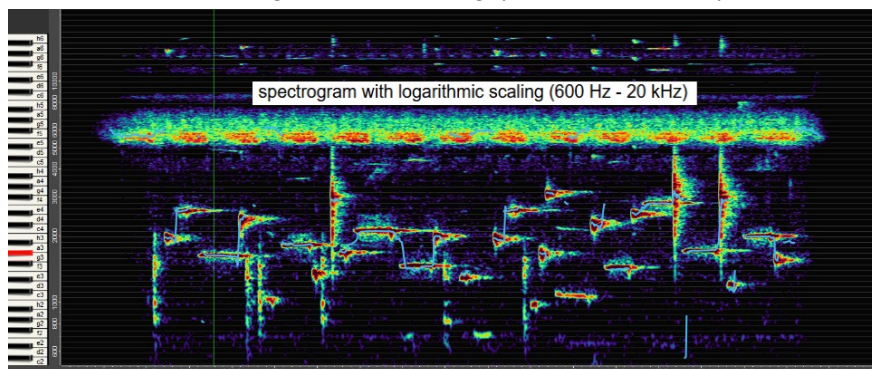
The *logarithmic spectrogram* shows at a glance the relative interval ratios of octaves, fifths and thirds in the range from 700 to 3000 Hz, as well as the condensed noise-like spectrum of cricket song at 5-6000 Hz and the high partial tone spectrum of the wren between 11 and 16 kHz (3rd-8th partial tone), in which all partial tones can be precisely determined.
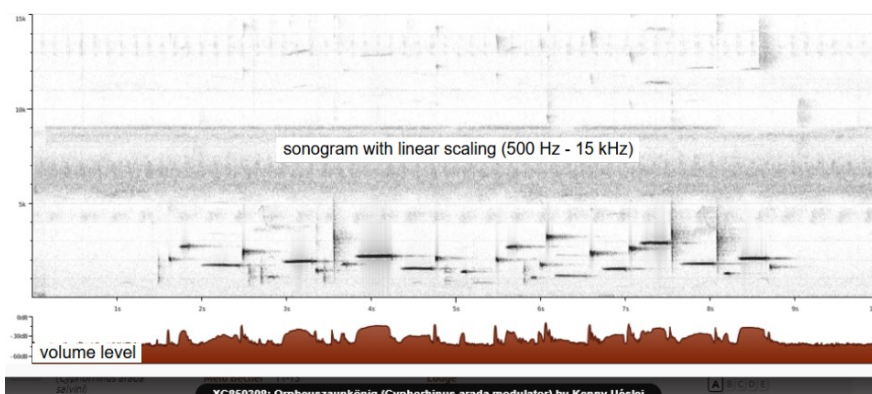
I can use the cursor to determine the frequency and relative volume for each pitch, regardless of whether it is the fundamental or a partial. However, this is even more precise in the 8-fold slowdown.

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)

spectrogram with logarithmic scaling (600 Hz - 20 kHz)

sonogram with linear scaling (500 Hz - 15 kHz)

sonogram with linear scaling (500 Hz - 15 kHz)

volume level

XC859308: Orpheuszaunkönig (Cyphorhinus arada modulator) by Kenny Uêslei

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/B8 (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".
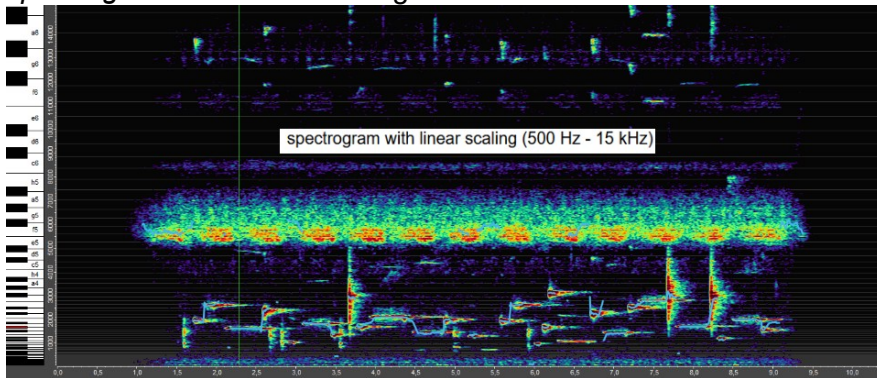This "fifth" can represent the simultaneous and successive sounding of 2 tones (C and G) with their own spectrum and the ratio of two partials to each other in a spectrum (2:3, 4:6:, 6:9, 8:12, 12:18 etc.).
And this applies to a fifth of the wren at F6 and C7 (1376 : 2064 Hz = 2:3) in the same way as 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.

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 G#6 to G6. In the 8-fold slowdown and in the logarithmic spectrogram, I can read the semitone change from the keyboard alone.

spectrogram with linear scaling

spectrogram with linear scaling (500 Hz - 15 kHz)

spectrogram slowed down 8x (60 - 2700 Hz)

spectrogram slowed down 8x (60-2700 Hz)