

# COMPARING LION ROARS AND HUMAN CRYING SOUND

HAZIM FITRI BIN AHMAD FAUDZI

P152419

UNIVERSITI KEBANGSAAN MALAYSIA

## ABSTRACT

This study presents a comparative bioacoustics analysis of two vocal signals which is an adult male's crying and a lion's roar. Audio recordings of each vocalization were analyzed across three key domains which are time-domain, frequency-domain, and time-frequency domain. We first used timer plots to analyze the signals in the time domain, observing their amplitude patterns over time. Amplitude-threshold segmentation was applied to isolate individual vocal events. Next, we performed a frequency-domain analysis using the Fast Fourier Transform (FFT) to examine the spectral content and identify dominant frequencies and harmonics. Finally, we conducted a time-frequency analysis using spectrograms that illustrate how frequency components evolve over time. The results show that lion roars exhibit strong low-frequency components with clearly visible harmonics, while the human cry features a higher-pitched fundamental frequency and broader spectral energy. The lion's spectrogram displays regular bursts of roaring, with strong sounds in the low pitch range. This fits with the lion have long vocal cords and a stretch-out throat. In contrast, the human crying spectrogram reveals repeated short bursts and irregular high-frequency content, reflecting emotional distress. These differences arise from anatomical and behavioural factors where lions use deep roars for long-range communication, while human crying expresses emotional states through higher pitch and variable timing. This paper documents the methodology and interprets the visual outputs, showing how different spectral analysis techniques can distinguish between these two distinct types of vocalizations.

## TABLE OF CONTENTS

	<b>Page</b>
<b>ABSTRACT</b>	<b>ii</b>
<b>TABLE OF CONTENTS</b>	<b>iii</b>
<b>LIST OF ABBREVIATIONS</b>	<b>iv</b>
<b>LIST OF ILLUSTRATIONS</b>	<b>v</b>
 <b>CHAPTER I     INTRODUCTION</b>	
 <b>CHAPTER II    LITERATURE REVIEW</b>	
 <b>CHAPTER III    METHODOLOGY</b>	
2.1                Data Acquisition	5
2.2                Time-domain Analysis	5
2.3                Frequency-domain Analysis	5
2.4                Time-frequency Domain Analysis	6
 <b>CHAPTER IV    RESULT AND DISCUSSION</b>	
3.1                Time-domain Analysis	7
3.2                Frequency Analysis	9
3.3                Spectrogram Analysis	11
 <b>CHAPTER V     CONCLUSION</b>	
 <b>REFERENCES</b>	<b>14</b>

**LIST OF ABBREVIATIONS**

FFT	Fast Fourier Transform
STFT	Short Time Fourier Transform

**LIST OF ILLUSTRATIONS**

Figure 1	Temporal Segmentation of Lion Roar	7
Figure 2	Temporal Segmentation of Man Crying	8
Figure 3	FFT of Lion's Roar	9
Figure 4	FFT of Man's Crying	10
Figure 5	Spectrogram of Lion Roar	11
Figure 6	Spectrogram of Man Crying	12

## CHAPTER I

### INTRODUCTION

Bioacoustic analysis helps us study animal and human vocalizations by measuring how they change over time (temporal features) and by identifying the different sound frequencies they contain (spectral features). A lion's roar (*Panthera leo*) is one of the loudest and most powerful sounds in the animal world, mainly used for marking territory and social communication. Lion roars are known to have very low fundamental frequencies because of the lion's large vocal folds and long vocal tract, which produce deeper sounds. On the other hand, human crying is an emotional vocalization. When people cry out of distress, the sounds they make tend to have a higher pitch and often include irregular or noisy elements that reflect emotional stress or arousal.

Even though lion roars and human cries serve different communication purposes, both can be studied using similar tools. We begin with time-domain analysis using timer plots to see how the loudness of the sound changes over time. We then use amplitude-threshold segmentation to break the signal into separate vocal parts. After that, we apply frequency-domain analysis using the Fast Fourier Transform (FFT) to find the main frequencies and harmonics. Lastly, we use time-frequency domain analysis with spectrograms that show how the sound frequencies change over time.

Previous studies have used spectrograms to look closely at lion roars. These studies show that lion roars have a strong, repeating structure with most of their energy in low frequencies (the fundamental and harmonics), but sometimes also include higher frequencies around 4 kHz. On the other hand, human crying for both infants and adults has also been studied. For instance, newborn babies usually cry with a pitch around 400–600 Hz. When a cry is caused by pain or high stress, it usually shows an even higher pitch and more instability, such as jitter (changes in pitch) and shimmer (changes in loudness).

The combination of amplitude-threshold segmentation, FFT, and spectrograms allows us to study and compare vocalizations in detail. In this study, we use these methods to analyze and interpret the differences between a lion's roar and a man's crying, linking the findings to their vocal anatomy and communication functions.

## CHAPTER II

### LITERATURE REVIEW

Bioacoustic studies often use amplitude-threshold methods to divide continuous vocalizations into individual segments. This technique is commonly applied to birdsong and large mammal calls, where uninterrupted sections above a set loudness level are marked as separate syllables or vocal units. For lion roars, past studies have described the acoustic structure as consisting of long-duration segments, multiple pulses, and a fundamental frequency around 180–200 Hz, with harmonic components extending into the kilohertz range. The lion's roar is a typical example of a source-filter vocalization, where the vocal folds generate a low-pitched source sound, and the long vocal tract shapes the final output, resulting in deep formant frequencies (Eklund et al., 2011).

Human crying has also been examined through frequency-based analysis. The main features include the fundamental frequency (F0), formant patterns, and voice quality characteristics such as jitter (small variations in pitch) and shimmer (small variations in amplitude). Studies have shown that cries linked to pain or emotional stress in infants tend to have higher F0 values and increased noise levels. While infant cries typically fall within the 400–600 Hz range, adult male cries are less well documented but are still expected to show elevated F0 and emotional variability (Michelsson & Michelsson, 1999).

To study such complex vocalizations, signal-processing tools like the Fast Fourier Transform (FFT) and Short-Time Fourier Transform (STFT) are commonly used. The FFT provides a full-spectrum view of frequency content in a single segment, while the STFT applies a moving window to capture how frequency changes over time, generating a spectrogram. These tools are widely used in animal and human vocal studies for their ability to reveal both stationary and dynamic spectral features.



In summary, the literature confirms that amplitude-threshold segmentation, FFT, and spectrogram-based STFT are effective methods for analyzing and comparing lion roars and human cries. The lion roar is known for its low pitch and harmonic richness (Eklund et al., 2011), while the human cry is more chaotic and high-pitched, especially during emotional states (Michelsson & Michelsson, 1999). This study builds upon these established techniques to examine both types of vocalizations in detail.

## **CHAPTER III**

### **METHODOLOGY**

#### **2.1 DATA ACQUISITION**

We used two audio clips in this study: one of a lion's roar and one of a man's crying. Both recordings were sourced from publicly available online social media platforms and downloaded in MP3 format.

#### **2.2 TIME-DOMAIN ANALYSIS**

We began by performing a time-domain analysis using timer plots to examine how the amplitude of the signal changes over time. To identify and isolate vocal events where roar pulses for the lion and sob bursts for the crying. We also applied amplitude-threshold segmentation. We computed the envelope of each waveform and defined a fixed threshold (20–50% of the maximum amplitude). Continuous regions exceeding this threshold were marked as vocal segments and annotated with their start and end times. This method allowed us to distinguish active vocalizations from periods of silence and is commonly used in bioacoustic segmentation tasks.

#### **2.3 FREQUENCY-DOMAIN ANALYSIS**

We conducted a frequency-domain analysis using the Fast Fourier Transform (FFT). The FFT converts a time-based waveform into its frequency components, enabling us to identify dominant pitches and harmonic structures. This process revealed key frequency characteristics of each segment, such as fundamental frequency and harmonic content.

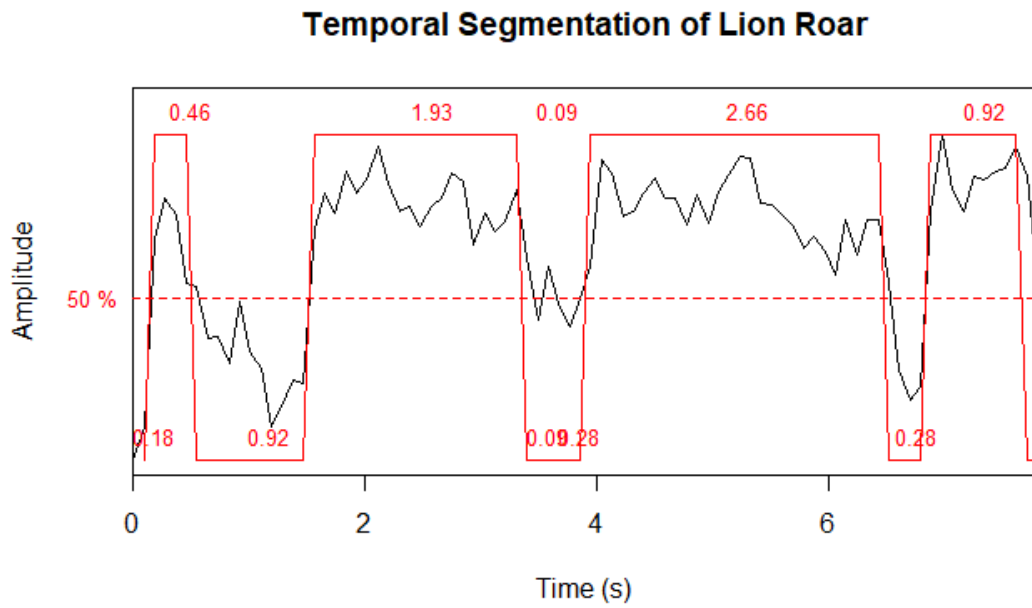
## **2.4 TIME-FREQUENCY DOMAIN ANALYSIS**

To understand how the frequency content of each sound changed over time, we used a time-frequency analysis approach by computing spectrograms. The spectrograms provide a visual map of how spectral energy evolves over time, capturing both steady tones and rapid fluctuations.

## CHAPTER IV

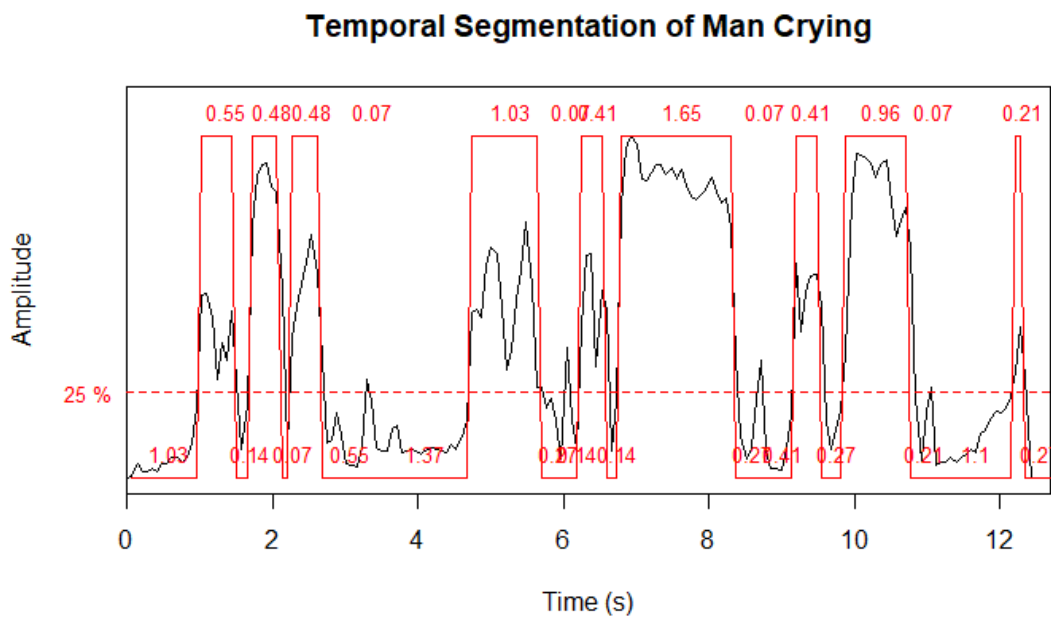
### RESULT AND DISCUSSION

#### 3.1 TIME-DOMAIN ANALYSIS



**Figure 1 Temporal Segmentation of Lion Roar**

Figure 1 shows the time-domain waveform of the lion's roar with amplitude plotted over time in seconds. The red horizontal dashed line marks the 50% amplitude threshold used for segmentation. Any segment of the signal that exceeded this threshold was identified as an active vocal event. The red rectangles highlight the four distinct roar pulses detected throughout the 7-second recording. Each pulse begins and ends where the signal crosses the threshold. For instance, the first roar starts at 0.18 seconds and ends at 0.92 seconds, while the second spans from 0.46 to 1.93 seconds. These segmented events reflect the lion's characteristic vocal pattern of long-duration and high-amplitude pulses spaced by quieter intervals. This segmentation method efficiently isolates individual vocalizations, supporting further analysis in the frequency and time-frequency domains.

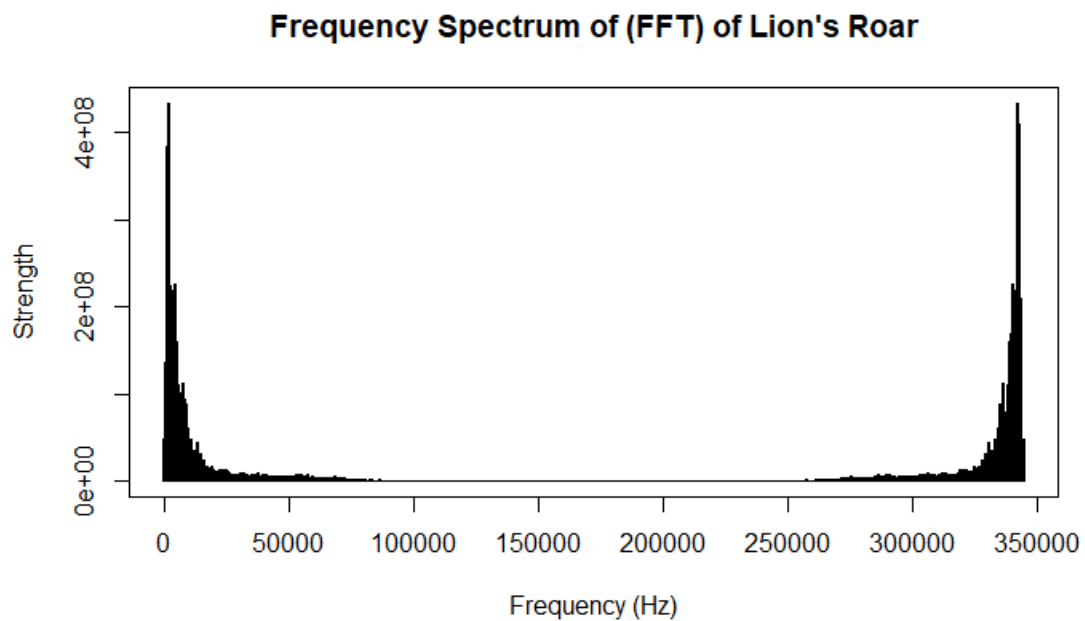


**Figure 2 Temporal Segmentation of Man Crying**

Figure 2 presents the time-domain waveform of the man's crying, with amplitude plotted over a duration of 13 seconds. The red dashed line represents the 25% amplitude threshold used for segmentation. Vocal segments that exceed this threshold are highlighted with red rectangles, marking regions of active crying. Compared to the lion's roar, the man's crying signal shows a much higher frequency of vocal bursts as can be seen from shorter and more numerous throughout the recording.

These vocal bursts shown brief silent pauses, reflecting a sobbing pattern. For instance, multiple segments occur within the first 3 seconds, each lasting less than a second. Additional bursts are seen around 6 to 7 seconds, and again between 9 and 12 seconds. This segmentation pattern highlights the emotional variability in human crying, which is often characterized by rapid, uneven vocal emissions. The fixed amplitude threshold method effectively isolates these crying bursts for further analysis in the frequency and time-frequency domains.

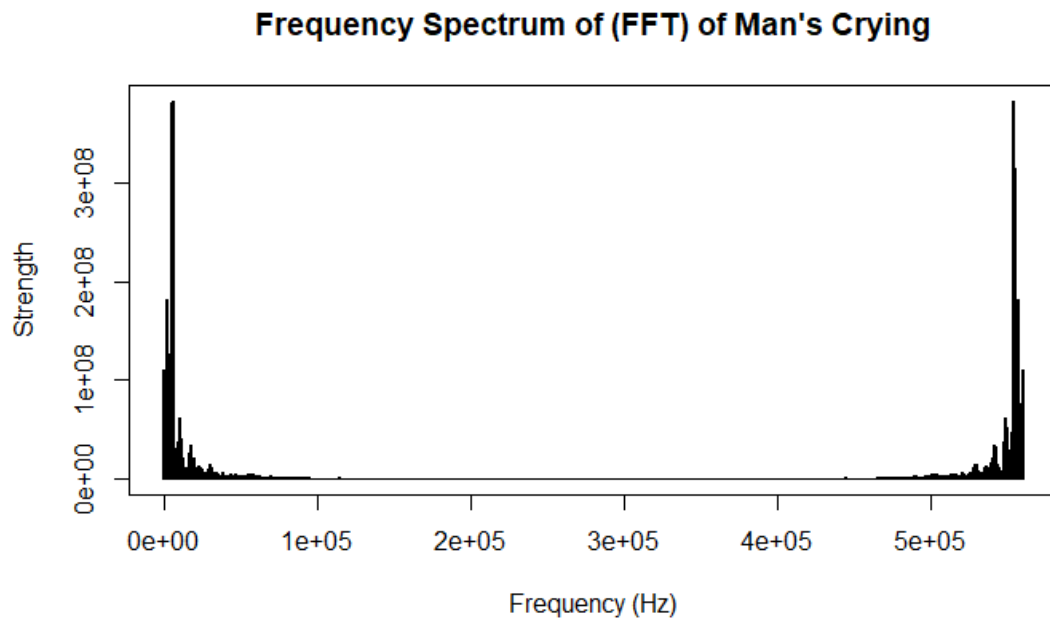
### 3.2 FREQUENCY ANALYSIS



**Figure 3 FFT of Lion's Roar**

Figure 3 shows the frequency spectrum of the lion's roar generated using the Fast Fourier Transform (FFT). The x-axis represents frequency in Hertz (Hz), while the y-axis represents the signal strength or amplitude in the frequency domain. The FFT reveals that the lion's roar contains dominant low-frequency components concentrated below 10,000 Hz, with a particularly strong peak near 200 Hz and consistent with previous findings that lion roars have a fundamental frequency around 180–200 Hz (Eklund et al., 2011).

The spectrum also displays harmonic components, which appear as additional peaks spaced at regular intervals beyond the fundamental frequency. These harmonics result from the periodic nature of the vocal fold vibrations. Beyond 50,000 Hz, the frequency components taper off significantly, indicating that most of the energy in the lion's roar is concentrated in the lower frequency range. This low-frequency concentration contributes to the roar's deep and resonant sound quality, which is effective for long-distance communication.



**Figure 4 FFT of Man's Crying**

Figure 4 displays the frequency spectrum of the man's crying, derived using the Fast Fourier Transform (FFT). The x-axis shows the frequency in Hertz (Hz), while the y-axis shows the corresponding strength or amplitude of those frequencies. The FFT reveals that the crying signal contains a broader range of frequency content compared to the lion's roar, with notable energy in both low and mid-frequency bands.

A strong peak is observed near the fundamental frequency which is most likely within the 300–600 Hz range. The exact value varies due to the unstable, emotional nature of the vocalization. Unlike the lion's roar, which exhibits strong and clear harmonic spacing, the crying signal shows a less uniform harmonic structure.

The spectral energy in the man's crying extends across a wider frequency range, declining gradually after 10,000 Hz but still present up to higher frequencies. This spectral spread is typical of distress vocalizations and highlights the complexity and emotional variability embedded in human crying.

### 3.3 SPECTROGRAM ANALYSIS

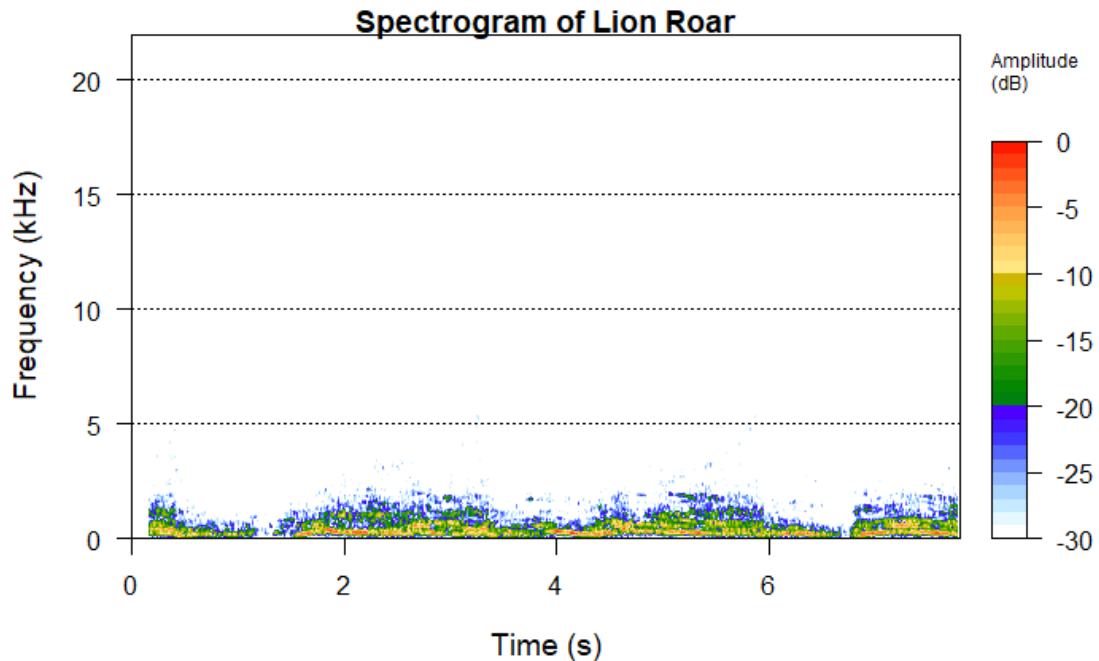


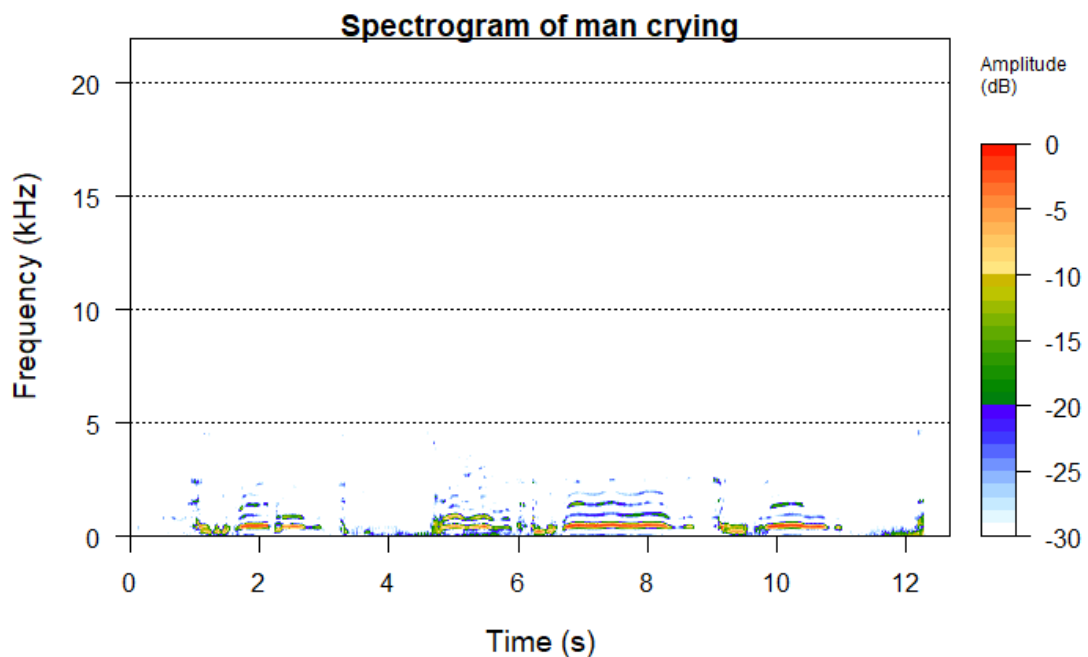
Figure 5 Spectrogram of Lion Roar

Figure 5 illustrates the time-frequency representation of the lion's roar using a spectrogram. The x-axis shows time in seconds, while the y-axis represents frequency in kilohertz (kHz). The colour scale on the right denotes amplitude in decibels (dB), where warmer colours (yellow to red) indicate higher energy, and cooler colours (green to blue) indicate lower energy.

The spectrogram reveals that the lion's vocal energy is heavily concentrated below 1 kHz, with the most intense energy bands appearing between 100 Hz and 500 Hz. This is consistent with the low fundamental frequency of the lion's roar, previously confirmed in the FFT analysis. These bands appear as horizontal streaks across time, representing strong harmonic structure and as evidence of the lion's long vocal folds and source-filter vocal mechanism (Eklund et al., 2011).

Additionally, the periodic nature of the roar is visible through recurring bright sections, corresponding to each roar pulse segmented earlier. Some energy is also visible up to 2–3 kHz, likely due to formants or higher harmonics, but most of the roar's power remains in the lower range. This time-frequency pattern supports the lion's need for low-pitched, far-reaching communication.





**Figure 6 Spectrogram of Man Crying**

Figure 6 shows the spectrogram of the man's crying, representing the time-frequency distribution of the vocal signal. The x-axis indicates time (in seconds), the y-axis represents frequency (in kilohertz), and the colour scale displays amplitude in decibels (dB), with red indicating higher energy and blue representing lower intensity.

Unlike the lion's spectrogram, this spectrogram features a series of short, sharp bursts of vocal activity that correspond to crying episodes segmented earlier in the time-domain analysis. These bursts are visible as narrow vertical bands occurring intermittently throughout the 13-second duration, particularly between 0–3 s, 6–8 s, and 9–12 s. Each band reflects a sobbing vocalization with a short time span and broad spectral spread.

The frequency content shows energy spread more evenly between 500 Hz and 3 kHz, with some harmonics visible as horizontal lines, although they are less clearly defined than in the lion's roar. Overall, the spectrogram of the man's crying demonstrates both the rapid temporal variability and broader frequency distribution typical of human distress vocalizations.

## CHAPTER V

### CONCLUSION

This study performed a comparative bioacoustic analysis of two distinct vocalizations of a lion's roar and a man's crying. Using time-domain, frequency-domain, and time-frequency domain techniques, we explored their structural and spectral differences through timer plots, FFT, and spectrograms.

The lion's roar exhibited a small number of long, low-frequency vocal pulses, with strong harmonic content and concentrated energy below 1 kHz. This pattern reflects its anatomical design with large vocal folds and a long vocal tract which enables deep, far-reaching calls suitable for territorial communication. In contrast, the man's crying consisted of numerous short, high-pitched bursts with greater spectral variability. These cries carried higher frequency components and irregular harmonic structure, which are typical of emotionally expressive human vocalizations.

Overall, the amplitude-threshold segmentation and spectral analysis effectively distinguished the two signals based on their temporal and acoustic signatures. This analysis not only highlighted the physiological differences in sound production but also demonstrated how acoustic features are shaped by communicative purpose.

## REFERENCES

- Eklund, R., Peters, G., Ananthakrishnan, G., & Mabiza, E. (2011). An acoustic analysis of lion roars. I: Data collection and spectrogram and waveform analyses. *Proceedings of Fonetik 2011*, vol. 22.
- Michelsson, K., & Michelsson, O. (1999). Phonation in the newborn, infant cry. *International Journal of Pediatric Otorhinolaryngology*, 49(Suppl. 1), S297–S301.