Computational Social Science with Images and Audio

Elliott Ash, **Philine Widmer**

10 November 2023

To move from images to audio, let us see what Dall-E 2 says about audio analysis and human speech.

Figure: "The auditory sense of humans is important, symbolized"

Figure: "Artwork on human speech"

Human hearing is interesting (not only for social scientists).

- Hearing is a critical sense for communication and survival
- As for vision, it enables us to interact with the world without physical contact
- Various areas of the brain are associated with auditory processing
- Humans are "analysts" and "generators" of audio at the same time
- Human hearing has its constraints
  - Limited frequency range: Humans typically hear between 20 Hz to 20 kHz, whereas some animals can perceive sounds beyond these limits
  - Difficulty in locating low-frequency sounds
- Human hearing is fascinating on several levels:
  - Fundamental: How does the auditory system decode sounds? $\rightarrow$ Beyond the scope of this class
  - Aggregate/social: What is the impact of sound on societal outcomes? $\rightarrow$ Central to our class

Audio analysis comes with many definitions.

- Many view it as a signal-processing challenge
- Julius O. Smith III: reconstructing auditory experiences
- In this class, we will explore techniques to analyze auditory data (recordings) in large quantities to study social science questions
  - Again, we begin where the auditory data already exists: we do not cover, from a technical perspective, the mechanisms of capturing this data (e.g., microphone technology)

What is sound?

- A sensory experience created by vibrations traveling through a medium (usually air)
- Like ocean waves move through water, sound waves propagate through the air
- As ocean waves interact with air, they create ripples in the atmosphere, analogous to **sound waves**
  - We experience these as the ambient sounds of the ocean

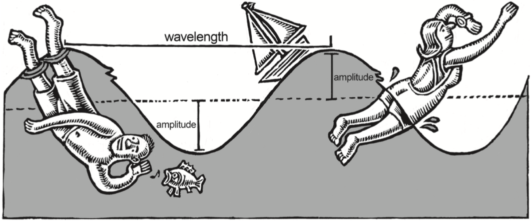

Figure: The analogy of ocean waves to sound waves; Source: Sulzer, (2021)[1]

[1] Sulzer (2021). Music, math, and mind: The Physics and Neuroscience of Music. Columbia University Press.

Humans' hearing range is 20 Hz to 20K Hz.

- We begin to hear low vibrations at around **20 Hz**
- Teenagers can hear frequencies that extend up to about **20,000 Hz**, above which we humans do not perceive sound
- Each time you double the frequency, you move up by one octave
- Hence, with good hearing, we can hear across a ten-octave range:

1. 20 Hz
2. 40 Hz (20 Hz $\times$ 2)
3. 80 Hz (40 Hz $\times$ 2)
4. 160 Hz (80 Hz $\times$ 2)
5. 320 Hz (160 Hz $\times$ 2)
6. 640 Hz (320 Hz $\times$ 2)
7. 1280 Hz (640 Hz $\times$ 2)
8. 2560 Hz (1280 Hz $\times$ 2)
9. 5120 Hz (2560 Hz $\times$ 2)
10. 10240 Hz (5120 Hz $\times$ 2)
11. 20480 Hz (10240 Hz $\times$ 2)

Figure: Human hearing range

Hz, octave, and sampling frequency are important concepts around sound waves.

- Hz (Hertz): Measures the number of cycles per second of a periodic waveform
  - Higher Hertz means a higher-pitched sound
  - If a tone has a frequency of 440 Hz, the sound wave completes 440 cycles every second (A4)
- Octave: The interval between one musical pitch and another with half or double its frequency
  - 440 Hz and 880 Hz
- Sampling frequency: The number of samples per second from a continuous signal to make a digital signal
  - Standard CD sampling rate is 44.1 kHz...
  - ... meaning that the audio waveform is being sampled 44,100 times per second

The wavelength is an important sound wave characteristic.

- What is a sound wave, to begin with?
  - A transfer of sound energy through a medium
  - It is transmitted via the vibration of particles within that medium
- Illustration for non-physicists:
  - When something vibrates (like a speaker's diaphragm), it pushes on the neighboring air particles, increasing the pressure in that region → this push is then transferred from particle to particle → when it reaches someone's ear, the ear interprets the varying pressures as sound
- Wavelength: The physical distance between identical points in consecutive cycles of a sound wave
  - Determines the pitch

The amplitude is another important characteristic.

- Amplitude *in a general wave context*: the magnitude of change in the oscillating variable within the wave
  - $\rightarrow$ Larger amplitudes mean more energy and often translate to louder sounds when talking about sound waves
- Amplitude *in digital audio signal processing*: understood as instantaneous amplitudes that represent the "magnitude of change" in the pressure wave (the sound wave) from its equilibrium position at that particular moment
  - $\rightarrow$ The specific value of the audio waveform at a given sample point
  - What is the equilibrium position?

What does this sound like?

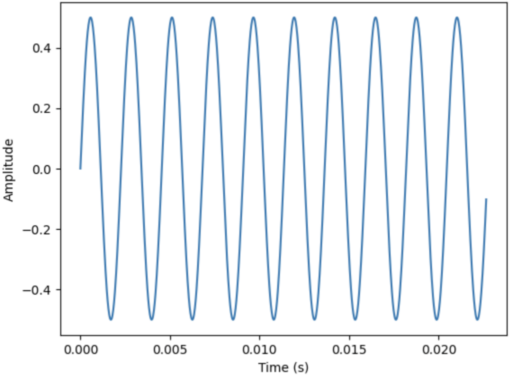

Figure: What sound could this be?

And what could this be?

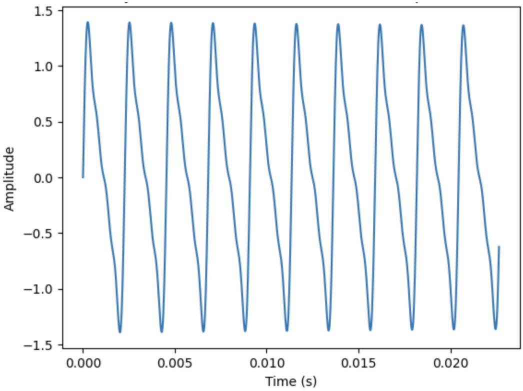

Figure: Any ideas?

This sound wave already looks more complex:

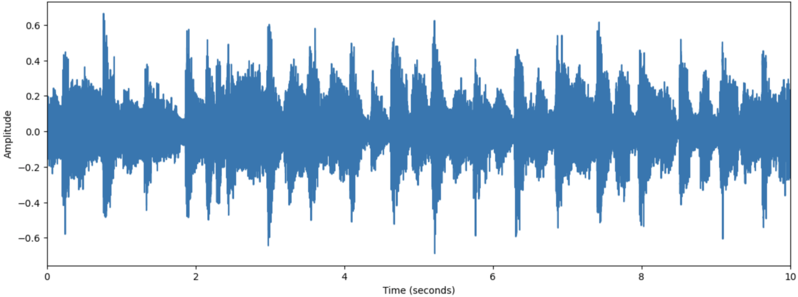

Figure: Hint: You certainly know this sound (too well).